
Expression, Purification
and Biophysical Characterization
of Pestiviral Autoprotease NS2
and *Sorghum bicolor* Deoxyhypusine/
Homospermidine Synthase

DISSERTATION

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Christina Hopf

Kiel, 2018

Erster Gutacher: Prof. Dr. Axel J. Scheidig
Zweiter Gutachter: Prof. Dr. Dietrich Ober

Tag der mündlichen Prüfung: 05.07.2018

Gez. Prof. Dr. Frank Kempken, Dekan

Eidesstattliche Erklärung:

Hiermit erkläre ich, Christina Hopf, dass ich die vorliegende Dissertation selbständig verfasst und keine weiteren als die angegebenen Quellen und Hilfsmittel verwendet habe. Inhalt und Form dieser Arbeit sind eigenständig erarbeitet und verfasst worden. Die Arbeit entstand unter der Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft. Weder die gesamte Arbeit noch Teile davon habe ich an anderer Stelle im Rahmen eines Prüfungsverfahrens eingereicht. Dies ist mein erster Promotionsversuch.

Kiel, den 08.05.2018

Christina Hopf

Experience is what you get when you didn't get what you wanted.

—RANDY PAUSCH

Meiner Familie.

Abstract

From the early days of life until today, a multifarious variety of unique enzymes has evolved to perform even the most complex tasks in cellular processes. These processes often involve the competing interests of prey and predator, host and pathogen, or plant and herbivore, and they are constantly adapted to keep pace with the challenges of survival. This thesis aims to introduce and characterize three fascinating enzymes from two very different venues of these evolutionary head to head races: the pestiviral autoprotease NS2 (non-structural protein 2) and a pair of paralogous enzymes, the deoxyhypusine synthase (DHS) and the homospermidine synthase (HSS).

The NS2 enzyme is an ER membrane-integrated autoprotease of the bovine viral diarrhea virus (BVDV). This virus is a close relative of the human pathogen hepatitis C virus (HCV) and the causative agent behind two associated, economically very important diseases of livestock breeding: bovine viral diarrhea (BVD) and mucosal disease (MD). In conjunction with a host cell-derived cofactor, the autoprotease NS2 effectively controls the delicate equilibrium between viral reproduction, immune evasion and persistence, and therefore is the key to the successful propagation and impressive worldwide prevalence of BVDV. In addition, the autoprotease acts as the molecular switch that changes the biotype of a BVDV strain from its non-cytopathic to a cytopathic form, thereby triggering the onset of fatal MD. For the treatment and eradication of BVDV, NS2 is consequently a highly valuable drug target. The rational design of antiviral substances, however, demands a precise three-dimensional structure of the target protein. To date, the protein structure of BVDV NS2 still remains to be solved, and even homology modeling failed so far due to the lack of structurally characterized, homologous proteins in the databases.

In this thesis, two common expression systems – prokaryotic *Escherichia coli* and eukaryotic insect cells – were applied to identify optimal conditions for the expression of full-length or truncated variants of the NS2 suitable for future structure determination experiments. Although a variety of strains, expression conditions and purification strategies were tried, the *E. coli* expression system ultimately failed in the production of NS2 protein suitable for further structural characterization. For insect cells, however, the reliable expression of active, ER membrane-integrated NS2 autoprotease could be achieved in the course of this thesis, and even delicate *in vivo* crystals were observed in the expressing insect cells. These results represent the profound base for the detailed biophysical and structural characterization of this unique viral enzyme.

The deoxyhypusine synthase (DHS) and its paralog homospermidine synthase (HSS) from the model plant *Sorghum bicolor* (commonly known as sorghum or great millet) have been intensively studied in the second part of this thesis. The homotetrameric DHS enzyme catalyzes the NAD⁺-dependent deoxyhypusination of a single specific lysine residue of eukaryotic translation initiation factor 5A (eIF-5A) as the first of two modification steps in the synthesis of the unique amino acid modification hypusine. In this reaction, an aminobutyl moiety is transferred from the polyamine substrate spermidine onto the eIF-5A lysine residue. Since the (deoxy)hypusine modification is crucial for the activity of the initiation factor 5A in cellular processes, the modifying DHS enzyme is highly conserved and ubiquitous in eukaryotic cells and archaea. During evolution, several independent duplications of the conserved DHS gene allowed the congruent development of paralogous homospermidine

synthases (HSS) in different lineages of flowering plant. The HSS enzymes are characterized by a loss of the inherited DHS activity and specialization on the homospermidine synthase side reactivity of the original DHS. In many of these plant lineages, the HSS became the first specific enzyme of the pyrrolizidine alkaloid biosynthesis pathway, which produces defensive substances as protection against herbivores. While already many details of the mechanism and dynamics of the deoxyhypusination reaction have been elucidated, the molecular basis for the evolutionary change of function from DHS to HSS still needs to be unraveled.

In this thesis, two *S. bicolor* DHS isoforms and their HSS paralog – which were assumed to represent different stages in the evolution of an HSS enzyme from a DHS duplicate – were characterized with several different biophysical techniques to gain deeper insights into their differences and common features. In contrast to the known basic pH optimum of the deoxyhypusination reaction, the cytosolic DHS and HSS enzymes and the eIF-5A precursor themselves revealed a thermal stability optimum at physiological pH and salt concentrations. However, the HSS was significantly less stable than the two DHS variants, and binding experiments confirmed that the enzyme lost its ability to bind the eIF-5A precursor, which consequently results in the loss of DHS activity. Kinetic analyses of the binding behavior of the two homologous DHS revealed low nanomolar to even sub-nanomolar affinities towards the eIF-5A precursor for both enzymes, which correlates to affinity values determined for human DHS. Distinct differences between the two homologs could not be resolved during these experiments. The DHS/HSS cofactor NAD^+ was demonstrated to affect the stability and aggregation state of the three enzymes and was able to significantly stabilize the DHS–eIF-5A interaction. For the competitive DHS inhibitor and spermidine analog *N*¹-guanyl-1,7-diaminoheptane (GC7), an hitherto unknown feature of the inhibitory mechanism was revealed: Accompanying the competitive displacement of the spermidine molecule from the active site of the DHS enzyme, the inhibitor also interrupts the DHS–eIF-5A interaction. In this thesis, first direct evidence was also obtained for the hypothesized paralog interference in the evolution of HSS and DHS – a concept in which the selective pressure of the original enzyme affects the evolutionary fate of its duplicate due to the formation of heteromeric complexes.

Zusammenfassung

Von den Anfängen des Lebens bis in die heutige Zeit hinein hat sich eine facettenreiche Vielfalt von Enzymen entwickelt, die in der Lage sind, selbst die komplexesten Aufgaben innerhalb zellulärer Prozesse wahrzunehmen. Diese Prozesse beinhalten oft die konkurrierenden Interessen von Beute und Jäger, Wirt und Krankheitserreger oder Pflanze und Pflanzenfresser, und sie werden ständig angepasst, um mit den Herausforderungen des Überlebens Schritt zu halten. Die vorliegende Arbeit möchte drei dieser faszinierenden Enzyme von zwei ganz unterschiedlichen Schauplätzen dieser evolutionären Kopf-an-Kopf-Rennen vorstellen und näher charakterisieren: die pestivirale Autoprotease NS2 und ein Paar paraloger Enzyme, die Desoxyhypusins-Synthase (DHS) und die Homospermidin-Synthase (HSS).

Das NS2-Enzym ist eine in die ER-Membran integrierte Autoprotease des Bovinen Virusdiarrhoe-Virus (BVDV). Dieses Virus ist eng verwandt mit dem humanpathogenen Hepacivirus C (HCV) und ist der Erreger zweier miteinander in Beziehung stehender, ökonomisch bedeutsamer Krankheitsbilder in der Rinderzucht: der Bovinen Virusdiarrhoe (BVD) und der Mucosal Disease (MD). Im Zusammenspiel mit einem Kofaktor aus der Wirtszelle kontrolliert die Autoprotease NS2 effektiv das empfindliche Gleichgewicht zwischen Virusvermehrung, Immunevasion und Persistenz, und spielt deshalb die Schlüsselrolle in der erfolgreichen Verbreitung von BVDV und der damit einhergehenden eindrucksvollen weltweiten Durchseuchung. Außerdem agiert die Autoprotease als molekularer Schalter, der die nicht-zytotoxische Form eines BVDV-Stammes in dessen zytotoxisches Pendant verwandeln kann und damit die tödlich verlaufende MD auszulösen vermag. Die NS2-Autoprotease ist deshalb ein interessantes Ziel für die Entwicklung von Wirkstoffen, die in der Behandlung und Ausrottung des BVD-Virus eingesetzt werden können. Für das gezielte Design von antiviralen Substanzen wird allerdings eine präzise dreidimensionale Struktur des Zielproteins benötigt. Bis jetzt wurde die Struktur der NS2 noch nicht gelöst und auch Homologie-Modellierungen schlugen bislang fehl, da keine strukturell charakterisierten, homologen Proteine in den Datenbanken identifiziert werden konnten.

In der vorliegenden Arbeit wurden zwei beliebte Expressionssysteme eingesetzt – der Prokaryot *Escherichia coli* und eukaryotische Insektenzellen –, um die optimalen Bedingungen für die Expression von verkürzten Varianten und Vollängen-Konstrukten der NS2-Autoprotease zu identifizieren, die anschließend für die Strukturbestimmung eingesetzt werden können. Obwohl eine große Anzahl verschiedener Stämme, Expressionsbedingungen und Reinigungsstrategien getestet wurde, war das bakterielle Expressionssystem letztendlich nicht in der Lage, die Autoprotease in ausreichender Qualität bereitzustellen. In den Insektenzellen allerdings konnte die zuverlässige Expression der aktiven, in die ER-Membran integrierten NS2-Autoprotease erreicht werden, und selbst fragile *in-vivo*-Kristalle wurden in den Insektenzellen beobachtet. Diese Ergebnisse stellen eine gute Grundlage für die weitere detaillierte biophysikalische und strukturelle Charakterisierung dieses einzigartigen viralen Enzyms dar.

Im zweiten Teil der vorliegenden Arbeit wurden die Desoxyhypusins-Synthase (DHS) und die paraloge Homospermidin-Synthase (HSS) des Modellorganismus *Sorghum bicolor* (Mohrenhirse) eingehend biophysikalisch untersucht. Das homotetramere DHS-Enzym katalysiert die NAD⁺-abhängige Desoxyhypusinierung eines einzigen spezifischen Lysinrestes des eukaryotischen Translations-Initiationsfaktors 5A (eIF-5A). In dieser Reaktion wird eine Aminobutyl-Einheit von dem Polyamin-Substrat Spermidin auf diesen Lysinrest des Initiationsfaktors übertragen. Dies stellt den ersten von zwei Schrit-

ten in der Synthese der einzigartigen Aminosäure Hypusin dar. Da die (Desoxy-)Hypusin-Modifikation essentiell ist für die biologische Aktivität des eIF-5A in den entsprechenden zellulären Prozessen, ist die DHS, die diese Modifikation erzeugt, sowohl in eukaryotischen Zellen als auch in Archaeen allgegenwärtig und hochkonserviert. Im Verlauf der Evolution konnte sich aufgrund von Duplikationen des konservierten DHS-Gens in einer Reihe von Blühpflanzen-Familien mehrmals unabhängig voneinander aus dem DHS-Enzym die paraloge Homospermidin-Synthase (HSS) entwickeln. Das HSS-Enzym hat sich auf eine Nebenreaktion der ursprünglichen DHS spezialisiert und besitzt keine DHS-Aktivität mehr. In vielen Blühpflanzen-Familien wurde die HSS zum Schlüsselenzym des darauf aufbauenden Pyrrolizidinalkaloid-Biosyntheseweges, der Abwehrsubstanzen zum Schutz vor Fraßfeinden produziert. Obwohl der Mechanismus und die Dynamiken der Desoxyhypusinerungsreaktion bereits in vielen Einzelheiten bekannt sind, ist die molekulare Grundlage der evolutionär bedingten Funktionsveränderung von der DHS zur HSS noch nicht im Detail aufgeklärt.

In dieser Arbeit wurden deshalb zwei DHS-Isoformen und ihr HSS-Paralog aus der Mohrenhirse – von denen angenommen wird, dass sie unterschiedliche Stadien in der Entwicklung von der DHS zur HSS darstellen – mit verschiedenen biophysikalischen Techniken näher charakterisiert, um einen besseren Einblick in ihre Unterschiede und gemeinsamen Eigenschaften zu erhalten. Im Gegensatz zum basischen pH-Optimum der Desoxyhypusinerungsreaktion zeigten die zytosolisch lokalisierten DHS- und HSS-Enzyme ebenso wie das unmodifizierte eIF-5A-Vorläuferprotein ein Stabilitäts optimum bei physiologischen pH-Werten und Salzkonzentrationen. Allerdings war die HSS im Vergleich zu den beiden DHS-Varianten deutlich weniger stabil. Bindungsexperimente bestätigten zudem, dass die HSS die Fähigkeit zur Bindung des eIF-5A-Vorläuferproteins verloren hat, was den Verlust der DHS-Aktivität zur Folge hat. In Übereinstimmung mit Literaturdaten zur Affinität der humanen DHS zeigten die kinetischen Analysen der Bindungseigenschaften der beiden homologen DHS nanomolare bis sub-nanomolare K_D -Werte. Ausgeprägte Unterschiede zwischen den beiden Homologen konnten im Rahmen dieser Experimente allerdings nicht bestimmt werden. Für den DHS/HSS-Kofaktor NAD^+ konnte ein Einfluss auf die Stabilität und den Aggregationszustand der drei Enzyme festgestellt werden. Zudem stabilisiert er die Interaktion zwischen DHS und eIF-5A deutlich. Eine bislang unbekannte Eigenschaft des inhibitorischen Mechanismus des kompetitiven DHS-Inhibitors und Spermidinanalogs N^1 -Guanyl-1,7-diaminoheptan (GC7) trat im Verlauf weiterer Experimente in Erscheinung: Neben der kompetitiven Verdrängung des Spermidin-Moleküls aus dem aktiven Zentrum des DHS-Enzyms verhindert der Inhibitor auch die Interaktion von DHS und eIF-5A und trennt bereits bestehende Komplexe wieder voneinander. Darüber hinaus wurden in der vorliegenden Arbeit auch erste direkte Hinweise für die Paralog-Interferenz in der Evolution von HSS und DHS gefunden. Dieses Konzept beschreibt ein evolutionäres Phänomen, in dem der Selektionsdruck des ursprünglichen Enzyms die Entwicklung des Duplikates über die Interaktion in gemeinsamen heteromeren Komplexen weiterhin mitbestimmt.

Contents

Abstract	i
Zusammenfassung	iii
Abbreviations	xv
1. Bovine viral diarrhea virus non-structural protein 2	1
1.1. Epidemiology and pathogenesis of BVD/MD	1
1.1.1. The "hit and run" strategy: acute transient and acute severe infections	2
1.1.2. The "infect and persist" strategy: PI animals	3
1.1.3. Mucosal disease	3
1.2. BVDV entry, replication and morphogenesis	5
1.3. Autoprotease non-structural protein 2	7
1.3.1. The role of NS2 in the switch from ncp to cp biotype	8
1.3.2. A cellular chaperone as cofactor of autoprotease NS2	9
1.3.3. The protease and substrate domains of NS2	11
1.3.4. Temporal modulation of NS2 autoproteolytic activity	12
1.3.5. Significance of NS2-3 cleavage for RNA replication and virion morphogenesis	12
1.4. Aim of this thesis	14
2. Results	15
2.1. Expression of BVDV NS2 autoprotease in <i>E. coli</i>	15
2.1.1. Design and cloning of NS2 expression constructs	15
2.1.2. Expression of His ₆ -SUMO constructs of <i>N</i> -terminally truncated NS2 in <i>E. coli</i>	15
2.1.3. Codon-optimization of <i>N</i> -terminally truncated NS2 constructs and expression in different <i>E. coli</i> strains	18
2.1.4. Solubilization of <i>N</i> -terminally truncated NS2 and NS2 _{Jiv90} constructs	23
2.1.5. Affinity purification of <i>N</i> -terminally truncated NS2 and NS2 _{Jiv90} constructs	26
2.1.6. SUMO protease cleavage experiments on His ₆ -SUMO-GFP and His ₆ -SUMO-[1393]NS2 _{Jiv90}	27
2.2. Expression of BVDV NS2 autoprotease in Sf21 and High Five TM insect cells	29
2.2.1. Design and cloning of full-length NS2 constructs for baculovirus-mediated expression in insect cells	30
2.2.2. Expression of full-length NS2 constructs in Sf21 and High Five TM insect cells	32

2.2.3.	In-gel fluorescence of full-length NS2 constructs expressed in High Five™ insect cells	34
2.2.4.	Western blot detection of the <i>N</i> -terminal <i>StreptII</i> -tag in full-length NS2 constructs expressed in Sf21 and High Five™ insect cells	36
2.2.5.	<i>In vivo</i> crystals in High Five™ cells expressing full-length NS2 constructs	38
2.2.6.	Fluorescence microscopy of <i>in vivo</i> -grown crystals	39
2.2.7.	Immunofluorescence microscopy of <i>in vivo</i> crystals	41
2.2.8.	Transmission electron microscopy of <i>in vivo</i> -grown crystals	42
2.2.9.	Serial crystallography/X-ray diffraction experiments with <i>in vivo</i> crystals	44
2.3.	Design and cloning of full-length NS2 constructs for expression in the cell-free BY-2 tobacco cell system	45
3.	Discussion	47
3.1.	<i>E. coli</i> expression of <i>N</i> -terminally truncated NS2 fusion constructs	47
3.1.1.	Codon-optimized, <i>N</i> -terminally truncated NS2 fusion constructs yield inclusion bodies upon expression in <i>E. coli</i>	48
3.1.2.	<i>N</i> -terminally truncated NS2 fusion proteins from inclusion bodies can be affinity-purified after solubilization with <i>N</i> -lauroylsarcosine surfactant	50
3.1.3.	Evaluation of SUMO protease cleavage as means to assess NS2 folding state is impaired by presence of NLS surfactant	51
3.2.	Expression of full-length NS2 fusion proteins in insect cells	52
3.2.1.	Design of functional full-length ncp and cp NS2 fusion constructs and choice of bacmid shuttle vectors	53
3.2.2.	Activable, full-length ncp and cp NS2 fusion protein is successfully expressed in Sf21 and High Five™ insect cells	53
3.2.3.	Full-length ncp NS2 autoprotease might be activated for cleavage by insect cell-derived Jiv homolog	54
3.2.4.	Insect cells expressing ER-targeted NS2 fusion proteins exhibit <i>in vivo</i> crystal growth	55
3.2.5.	Full-length ncp and cp NS2 fusion protein is localized in the insect cell ER, but cannot be detected in <i>in vivo</i> crystals	57
3.2.6.	Delicate <i>in vivo</i> crystals disintegrate upon fixation for immunofluorescence and TEM	58
3.2.7.	<i>In vivo</i> crystals do not diffract micro-focused synchrotron X-radiation <i>in cellulose</i>	59
3.3.	Conclusion and future prospects	60
4.	<i>Sorghum bicolor</i> deoxyhypusine synthase	63
4.1.	Eukaryotic translation initiation factor 5A	64
4.1.1.	The role of eIF-5A and its homologs in protein synthesis	66
4.1.2.	Involvement of eIF-5A in pathologic processes	67
4.2.	Deoxyhypusine synthase	68
4.2.1.	The proposed reaction mechanism of deoxyhypusine formation	68

4.2.2.	Tertiary and quaternary structure of DHS and relevance of specific amino acid residues	70
4.2.3.	Binding sequence of NAD ⁺ , spermidine and eIF-5A	73
4.2.4.	DHS cooperativity effects upon binding of eIF-5A	75
4.2.5.	DHS substrate specificity, inhibitors and side reactions	76
4.3.	DHS paralog homospermidine synthase in angiosperm species: trailhead of the pyrrolizidine alkaloid biosynthesis pathway	79
4.4.	Bacterial HSS and plant HSS – similar, but not the same	83
4.5.	Aim of this thesis	84
5.	Results	85
5.1.	Optimization of <i>SbDHS</i> isoform purification	85
5.1.1.	Scale-up of the <i>E. coli</i> expression and purification of His ₆ -tagged <i>SbDHS</i> isoforms	85
5.1.2.	Optimization of expression and purification conditions	87
5.1.3.	Optimized purification strategy	90
5.2.	Heterologous production of <i>Sorghum bicolor</i> DHS substrate eIF-5A	93
5.2.1.	Design and cloning of <i>SbelF-5A</i> constructs	93
5.2.2.	<i>E. coli</i> expression and purification of <i>SbelF-5A</i> constructs	93
5.3.	Crystallization of <i>SbDHS</i> and <i>SbelF-5A</i> isoforms	95
5.3.1.	Classic crystallization approach	95
5.3.2.	<i>In vivo</i> crystallization in insect cells	98
5.4.	Thermal stability determination of <i>SbDHS</i> and <i>SbelF-5A</i> isoforms	102
5.4.1.	Buffer variation and pH optimum	102
5.4.2.	Effect of NaCl concentration on thermal stability	106
5.4.3.	The stabilizing influence of NAD ⁺ and GC7 on <i>SbDHS</i> variants	107
5.5.	Determination of binding kinetics for the different <i>SbDHS</i> isoforms by bio-layer interferometry	110
5.5.1.	Binding characteristics and kinetics of <i>SbDHS</i> homo/paralogs in basic glycine buffer	113
5.5.2.	Binding characteristics and kinetics of <i>SbDHS</i> homo/paralogs in PBS at physiological pH	117
5.5.3.	Identification of the cause for heterogeneous binding behavior	119
5.5.4.	The influence of NAD ⁺ and the inhibitor GC7 on the interaction of <i>SbelF-5A</i> and DHS	123
5.6.	SEC-MALS characterization of <i>SbDHS1</i> , <i>SbDHS2</i> and <i>SbelF-5A</i>	125
5.7.	Interaction of DHS and HSS subunits	127
6.	Discussion	132
6.1.	Supplementation of NAD ⁺ is necessary to prevent protein loss during purification and aggregation upon storage	133
6.2.	Two <i>Sorghum bicolor</i> eIF-5A isoforms are readily expressed and purified based on previously established protocols	136

6.3.	Crystallization of individual <i>S. bicolor</i> DHS homo/paralogs and in complex with eIF-5A precursor	137
6.3.1.	The classical crystallization approach yielded diffracting crystals with a high number of overlaid crystal lattices for the DHS enzymes	138
6.3.2.	Expression of DHS isoforms in Sf21 insect cells did not yield <i>in vivo</i> crystals	140
6.4.	The DHS isoforms tolerate a broad range of buffer conditions and they are significantly stabilized in the presence of NAD ⁺ and the inhibitor GC7 especially under basic conditions	141
6.5.	Characterization of the binding between the different <i>S. bicolor</i> DHS and eIF-5A isoforms by bio-layer interferometry (BLI)	143
6.5.1.	<i>S. bicolor</i> DHS3 ("HSS") does not bind to either of the two eIF-5A isoforms	144
6.5.2.	Analyses of binding kinetics reveal low nanomolar to sub-nanomolar affinities between <i>S. bicolor</i> DHS and eIF-5A isoforms	147
6.5.3.	Heterogeneity in binding behavior cannot be attributed to a distinct binding model	148
6.5.4.	The inhibitor GC7 prevents binding of eIF-5A to DHS and interrupts existing complexes	150
6.6.	Complex between <i>S. bicolor</i> DHS and eIF-5A precursor does not survive size exclusion chromatography	151
6.7.	Evidence for the existence of paralogous heteromers comprised of DHS and HSS subunits	153
6.8.	Conclusion and future prospects	156
7.	Material and methods	159
7.1.	Bacterial strains and insect cell lines	159
7.2.	BVDV NS2 expression constructs	160
7.3.	<i>Sorghum bicolor</i> eIF-5A and DHS expression constructs	162
7.4.	Oligonucleotides	165
7.5.	Chemicals	166
7.6.	Molecular biology	167
7.6.1.	<i>In silico</i> design of constructs and evaluation of cloning strategies	167
7.6.2.	Codon usage analysis	167
7.6.3.	Sequencing	167
7.6.4.	Oligonucleotide synthesis	167
7.6.5.	Analytical and preparative agarose gel electrophoresis	167
7.6.6.	Purification of DNA fragments	167
7.6.7.	Plasmid DNA preparation	168
7.6.8.	DNA and protein concentration determination	168
7.6.9.	Phosphorylation of oligonucleotides	168
7.6.10.	Dephosphorylation of DNA fragments	168
7.6.11.	Restriction enzyme digestion of DNA	168

7.6.12.	DNA ligation	168
7.6.13.	Polymerase Chain Reaction (PCR)	169
7.6.14.	RF-Cloning and site-directed mutagenesis	169
7.6.15.	Transformation of competent <i>E. coli</i> with plasmid DNA	170
7.7.	Protein-related methods	170
7.7.1.	SDS-PAGE	170
7.7.2.	SDS-PAGE staining	171
7.7.3.	Western blot of protein samples from SDS-PAGE	171
7.7.4.	Chromogenic and fluorescence detection of His ₆ - and <i>StreptII</i> -tagged proteins on western blot	171
7.8.	Cloning of BVDV NS2 constructs	172
7.8.1.	Codon-optimized [1303]NS2 _{Jiv90} and [1303]NS2 for expression in <i>E. coli</i>	172
7.8.2.	Codon-optimized NS2 constructs with <i>N</i> -terminal His ₆ -tagged SUMO fusion	173
7.8.3.	NS2 constructs with <i>N</i> - or <i>C</i> -terminal His ₆ -tagged GFP fusion	173
7.8.4.	NS2 constructs for baculovirus-mediated expression in Sf21 and High Five TM insect cells	173
7.8.5.	NS2 constructs for expression in the cell-free BY2 tobacco cell system	174
7.9.	Cloning of <i>Sorghum bicolor</i> DHS and eIF-5A isoform constructs	175
7.9.1.	Cloning of <i>SbeIF</i> -5A-His ₆ isoform constructs	175
7.9.2.	Cloning of <i>SbeIF</i> -5A-His ₆ - <i>SASTreptII</i> isoform constructs	176
7.9.3.	Cloning of <i>StreptII</i> - <i>SbDHS</i> -His ₆ isoform constructs	176
7.9.4.	Cloning of <i>SASTreptII</i> - <i>SbDHS</i> -His ₆ isoform constructs	176
7.9.5.	Cloning of <i>SbDHS</i> coexpression constructs	176
7.9.6.	Cloning of <i>SbDHS</i> constructs for baculovirus-mediated expression in Sf21 and High Five TM insect cells	177
7.10.	Culture of <i>E. coli</i> for protein expression	178
7.11.	Expression and Purification of <i>SbeIF</i> -5A isoform constructs	178
7.11.1.	Expression of <i>SbeIF</i> -5A-His ₆ isoforms and the <i>SASTreptII</i> -tagged variants	178
7.11.2.	Purification of <i>SbeIF</i> -5A-His ₆ and the <i>SASTreptII</i> -tagged isoforms	179
7.12.	Expression and purification of <i>SbDHS</i> /HSS constructs	180
7.12.1.	Expression of <i>SbDHS</i> /HSS constructs	180
7.12.2.	Purification of <i>SbDHS</i> /HSS constructs	180
7.12.3.	Purification of <i>SbDHS</i> /HSS coexpression and control <i>SbDHS</i> /HSS for the detection of heteromers	181
7.13.	Insect cell culture techniques	182
7.13.1.	Transposition of pFastBac1-cyt, -ER, -KDEL and -SKL constructs into baculovirus shuttle-vectors bMON14272 and EMBAcY	182
7.13.2.	Transfection of Sf21 insect cells with baculovirus shuttle vectors	183
7.13.3.	Baculovirus amplification	184
7.13.4.	<i>In vivo</i> crystallization in Sf21 or High Five TM insect cells	184
7.14.	Biophysical techniques	184
7.14.1.	Fluorescence microscopy of Sf21 or High Five TM insect cells	184

7.14.2. Immunofluorescence staining of NS2-expressing High Five™	184
7.14.3. Transmission electron microscopy (TEM) of insect cells	185
7.14.4. Serial crystallography/X-ray diffraction experiments with <i>in vivo</i> crystals	186
7.14.5. Differential scanning fluorimetry (DSF)	186
7.14.6. Dynamic Light Scattering (DLS)	187
7.14.7. Size exclusion chromatography with multi-angle light scattering (SEC-MALS)	188
7.14.8. Bio-layer interferometry (BLI)	189
A. Appendix	191
A.1. Codon usage analysis of NS2 sequences	192
A.2. Amino acid sequence alignments of Jiv-containing proteins	197
A.3. TEM image of an insect cell infected with Baculovirus heliothis	199
A.4. Codon usage analysis of <i>SbDHS</i> and eIF-5A sequences	200
A.5. Alignments of e/aIF-5A and DHS homo- and paralogs	204

List of Figures

1.1. Overview of BVDV infection strategies	4
1.2. Schematic representation of the BVDV polyprotein	6
1.3. Schematic representation of BVDV NS2 topology	8
1.4. Schematic representation of bovine Jiv topology	10
1.5. Temporal modulation of ncp BVDV NS2 by cellular cofactor Jiv	13
2.1. Constructs for expression of <i>N</i> -terminal truncated NS2 in <i>E. coli</i>	16
2.2. Expression of <i>N</i> -terminal truncated NS2 in <i>E. coli</i> BL21 (DE3)pLysS, BL21 (DE3)RIL, Rosetta TM 2(DE3) and Rosetta TM 2(DE3)pLysS	17
2.3. Constructs for <i>E. coli</i> expression of <i>N</i> -terminal truncated NS2, <i>N</i> - or <i>C</i> -terminally fused to GFP	19
2.4. Codon-optimized His ₆ -SUMO-[1393]NS2 _{Jiv90} expression in BL21 (DE3) and derivatives	20
2.5. Expression of codon-optimized His ₆ -SUMO-NS2 and NS2-GFP-His ₆ constructs in Origami TM 2(DE3) and derivatives	21
2.6. Expression of codon-optimized His ₆ -SUMO-NS2, NS2-GFP-His ₆ and His ₆ -GFP-NS2 constructs in ArcticExpress TM (DE3)	22
2.7. Solubilization experiments with His ₆ -SUMO-[1393]NS2 constructs	24
2.8. Affinity purification of His ₆ -SUMO-[1393]NS2 _{Jiv90} and [1379]NS2 _{Jiv90} -GFP-His ₆	25
2.9. SUMO protease cleavage experiments	29
2.10. Constructs for baculovirus-mediated expression of full-length NS2 in insect cells	31
2.11. Colony PCR results of <i>E. coli</i> DH10EMBacY and DH10Bac TM clones transformed with full-length NS2 pFastBac1 constructs	33
2.12. High Five TM insect cells transfected with full-length NS2 constructs	34
2.13. In-gel fluorescence and western blots of full-length NS2 constructs expressed in Sf21 and High Five TM insect cells	35
2.14. <i>In vivo</i> -grown crystals from full-length NS2 constructs expressed in High Five TM insect cells	38
2.15. Fluorescence microscopy of <i>in vivo</i> crystals in High Five TM insect cells	40
2.16. Immunofluorescence microscopy of High Five TM insect cells infected with the full-length NS2 ER construct	41
2.17. Transmission electron microscopy (TEM) of High Five TM insect cells transfected with full-length NS2 constructs	43
2.18. Constructs for expression of full-length NS2 in BY-2 tobacco cell extract	45
4.1. Representation of <i>A. thaliana</i> eIF-5A tertiary structure and residue conservation	65
4.2. Unique posttranslational modifications on eukaryotic and archaeal eIF-5A and its bacterial counterpart EF-P	67

4.3.	Schematic representation of the putative DHS reaction mechanism	69
4.4.	Representation of human DHS tertiary structure and residue conservation . . .	71
4.5.	Structural representation of DHS substrate spermidine and its inhibitor GC7 . .	72
4.6.	Schematic representation of DHS substrate binding sequence and its catalytic cycle	74
4.7.	Representation of <i>A. thaliana</i> eIF-5A surface electrostatic potential at pH 7.4 . .	76
4.8.	Overview of DHS amino donor/acceptor substrates and their reaction products	78
4.9.	Retronecine	79
4.10.	Representation of the surface electrostatic potential of the active site tunnel entrance of <i>S. vulgaris</i> and <i>S. bicolor</i> DHS and HSS models	80
4.11.	Alignment of the amino acid sequences of two DHS homologs and their HSS paralog from <i>S. bicolor</i>	81
4.12.	Surface representation of the <i>S. bicolor</i> HSS model and its differences com- pared to the DHS model	82
5.1.	Constructs for expression of <i>SbDHS</i> isoforms in <i>E. coli</i>	85
5.2.	IMAC profile and SDS-PAGE result of <i>SbDHS2</i> -His ₆ purification according to the original protocol from up-scaled <i>E. coli</i> expression culture	86
5.3.	DLS results of <i>SbDHS2</i> -His ₆ prior to and after flash freezing/thawing, and the effect of NAD ⁺ and GC7	89
5.4.	SDS-PAGE results of <i>StrepII</i> - <i>SbDHS2</i> -His ₆ purification according to the opti- mized protocol from up-scaled <i>E. coli</i> expression culture	91
5.5.	DLS results of <i>StrepII</i> - <i>SbDHS2</i> -His ₆ purified according to optimized protocol prior to and after flash freezing/thawing	92
5.6.	Constructs for expression of <i>SbeIF</i> -5A isoforms in <i>E. coli</i>	93
5.7.	SDS-PAGE results of <i>SbeIF</i> -5A1 IMAC, SEC and <i>Strep</i> -Tactin [®] purifications	94
5.8.	Crystals and diffraction images of <i>S. bicolor</i> DHS and eIF-5A isoforms	97
5.9.	Constructs for baculovirus-mediated expression of <i>SbDHS1</i> –3 in insect cells . .	98
5.10.	Colony PCR results of <i>E. coli</i> DH10EMBacY and DH10Bac [™] clones trans- formed with <i>SbDHS</i> variant pFastBac1 constructs	100
5.11.	Sf21 insect cells transfected with <i>SbDHS1</i> –3 constructs in bMON14272	101
5.12.	Melting curves of <i>SbDHS3</i> (HSS) under varying buffer and pH conditions determined by DSF	103
5.13.	Melting temperatures of <i>SbDHS</i> and <i>SbeIF</i> -5A homo/paralogs under varying buffer and pH conditions determined by DSF	104
5.14.	Melting curves of <i>SbDHS</i> /HSS determined by DLS	105
5.15.	Melting temperatures of <i>SbDHS</i> and <i>SbeIF</i> -5A homo/paralogs under selected buffer, pH and salt conditions determined by DSF	107
5.16.	Melting curves of <i>SbDHS</i> variants in correlation to NAD and GC7 concentra- tion, determined by DSF	109
5.17.	<i>SbDHS</i> isoforms constructs for BLI measurements	111
5.18.	Schematic representation of the BLI principle and the strategy for the <i>SbDHS</i> /eIF- 5A binding kinetics measurements	112

5.19. Binding behavior of <i>S. bicolor</i> DHS and eIF-5A1 in basic glycine buffer	114
5.20. Binding kinetics of <i>S. bicolor</i> eIF-5A homologs as ligands and DHS homologs as analytes in glycine-NaOH buffer pH 9.0	115
5.21. Binding kinetics of <i>S. bicolor</i> eIF-5A homologs as ligands and DHS homologs as analytes in PBS buffer at pH 7.5	118
5.22. Determination of binding model based on the analysis of a double-exponential fit according to Tiwari et al. 2015 [272]	122
5.23. Association or dissociation of <i>S. bicolor</i> DHS2 to/from eIF-5A2 in the presence of NAD ⁺ and/or the inhibitor GC7	124
5.24. SEC-MALS measurements of <i>S. bicolor</i> DHS2 and eIF-5A	126
5.25. Coexpression and control constructs of <i>SbDHS/HSS</i>	128
5.26. Western blots of coexpressed <i>SbDHS/HSS</i>	130
A.1. Codon usage analysis of BVDV NCP7 [1303]NS2 optimized for bacterial ex- pression compared to <i>E. coli</i> B	192
A.2. Codon usage analysis of BVDV NS2* of strain NCP7 compared to <i>E. coli</i> B .	193
A.3. Codon usage analysis of codon-optimized full-length BVDV NCP7 NS2 com- pared to <i>S. frugiperda</i>	194
A.4. Codon usage analysis of full-length BVDV NCP7 NS2* compared to <i>S. frugiperda</i>	195
A.5. Codon usage analysis of codon-optimized full-length BVDV NCP7 NS2 com- pared to <i>N. tabacum</i>	196
A.6. Alignment of Jiv (<i>Bos taurus</i>) and dnj-5 (<i>Spodoptera litura</i>) amino acid se- quences	197
A.7. Alignment of Jiv90-containing protein sequences from <i>Artiodactyla</i> and <i>Lepi- doptera</i> species	198
A.8. TEM image of an insect cell infected with Baculovirus heliothis	199
A.9. Codon usage analysis of <i>Sorghum bicolor</i> DHS1 compared to <i>E. coli</i> B	200
A.10. Codon usage analysis of <i>Sorghum bicolor</i> DHS2 compared to <i>E. coli</i> B	201
A.11. Codon usage analysis of <i>Sorghum bicolor</i> DHS3 compared to <i>E. coli</i> B	202
A.12. Codon usage analysis of <i>Sorghum bicolor</i> eIF-5A1 and eIF-5A2 compared to <i>E. coli</i> B	203
A.13. Alignment of the amino acid sequences of initiation factor 5A variants from <i>Sorghum bicolor</i>	204
A.14. Alignment of the amino acid sequences of initiation factor 5A homologs from eukaryotes and archae	205
A.15. Alignment of the amino acid sequences of DHS homologs from eukaryotes and archae	206
A.16. Alignment of the amino acid sequences of angiosperm DHS and HSS	207

List of Tables

5.1. Binding kinetics of <i>S. bicolor</i> DHS1 and DHS2 to eIF-5A1 and eIF-5A2 in glycine-NaOH buffer	117
5.2. Binding kinetics of <i>S. bicolor</i> DHS1 and DHS2 to eIF-5A1 and eIF-5A2 in PBS	119
5.3. Relationships of the sums and products of σ and γ and the individual values of σ to the analyte concentration C and the k_a and k_d values of the three biphasic models	120
7.1. Bacterial strains and insect cell lines	159
7.2. NS2 constructs for expression in <i>E. coli</i>	160
7.3. NS2 constructs for insect cell expression	161
7.4. NS2 constructs for cell free expression in tobacco BY-2 cell extract	161
7.4. NS2 constructs for cell free expression in tobacco BY-2 cell extract (<i>continued</i>)	162
7.5. eIF-5A constructs for expression in <i>E. coli</i>	162
7.6. DHS constructs for expression in <i>E. coli</i>	162
7.6. DHS constructs for expression in <i>E. coli</i> (<i>continued</i>)	163
7.7. DHS constructs for expression in insect cells	164
7.8. List of oligonucleotides used for cloning purposes	165
7.8. List of oligonucleotides used for cloning purposes (<i>continued</i>)	166

Abbreviations

aa	amino acid(s)
aIF-5A	archaeal translation initiation factor 5A
AP	alkaline phosphatase
APS	ammonium persulfate
ATP	adenosine triphosphate
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BLI	bio-layer interferometry
BSA	bovine serum albumin
BVD	bovine viral diarrhea
BVDV	bovine viral diarrhea virus
C	core protein
CD	circular dichroism
cDNA	complementary DNA
cp	cytopathic
Cpn10/60	chaperones, GroEL/GroES homologs
cryoEM	cryo-electron microscopy
CV	column volume
cyt	cytosolic localization
DAPI	4',6-diamidino-2-phenylindole
DHS	deoxyhypusine synthase (EC 2.5.1.4.6)
DLS	dynamic light scattering
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DnaJ	J-domain containing chaperone of the Hsp40 family

DnaK	chaperone of the Hsp70 family
DOHH	deoxyhypusine hydroxylase (EC 1.14.99.29)
DSF	differential scanning fluorimetry
DTT	dithiothreitol
E	envelope protein
EDTA	ethylenediaminetetraacetic acid
EF-P	(bacterial) elongation factor P
EYFP	enhanced green fluorescent protein
eIF-5A	eukaryotic translation initiation factor 5A, until 1991 known as eIF-4D
ER(-KDEL)	signal sequence for ER-targeting (and retention)
ER	endoplasmic reticulum
Erns	secreted ribonuclease
EYFP	enhanced yellow fluorescent protein
Fc	fragment crystallizable region of an antibody
FPLC	fast protein liquid chromatography
FRET	Förster resonance energy transfer
GalNAc	<i>N</i> -acetylgalactosamine
GC7	<i>N</i> ¹ -guanyl-1,7-diaminoheptane
GFP	green fluorescent protein
GlcNAc	<i>N</i> -acetylglucosamine
GPCR	G protein-coupled receptor
GroEL	chaperone of the Hsp60 family
GroES	chaperone of the Hsp10 family
GrpE	nucleotide exchange factor of DnaK
HA	affinity tag, containing a part of human influenza hemagglutinin
HCV	hepacivirus C
His₆-tag	hexahistidine affinity tag
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography

Hsp	heat shock protein
HSS	homospermidine synthase (EC 2.5.1.45)
HTX	high throughput crystallization
IgG	immunoglobulin G
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl- β -D-1-thiogalactopyranoside
IRES	internal ribosomal entry site
JBP	Jiv-binding peptide
Jiv	J-domain protein interacting with viral protein
KDEL	see ER(-KDEL)
lacZ	gene coding for β -galactosidase
LB	lysogeny broth medium
LCP	lipid cubic phase
LDAO	lauryldimethylamine- <i>N</i> -oxide
MALS	multi-angle light scattering
MCS	multiple cloning site
MD	mucosal disease
mRNA	messenger RNA
MWCO	molecular weight cut-off
MWM	molecular weight marker
NAD⁺	nicotinamide adenine dinucleotide, oxidized
NADH	nicotinamide adenine dinucleotide, reduced
NBT	nitro-blue tetrazolium chloride
ncp	non-cytopathic
NLS	<i>N</i> -lauroylsarcosine
NMR	nuclear magnetic resonance
Npro	nonstructural autoprotease
NS	nonstructural protein
NTPase	nucleoside-triphosphatase

OD₆₀₀	optical density at 600 nm
ORF	open reading frame
P	virus progeny generation
PA	pyrrolizidine alkaloid
PCR	polymerase chain reaction
p.i.	post infection
PI	persistently infected
PMSF	phenylmethylsulfonyl fluoride
polh	polyhedrin promoter from AcMNPV
P-site	peptidyl site, second binding site for tRNA in the ribosome
PTC	ribosomal peptidyl transferase center
PVDF	polyvinylidene fluoride
RdRp	RNA-dependent RNA polymerase
RF1	release factor 1
RF-Cloning	restriction-free cloning
RNA	ribonucleic acid
RT-PCR	real-time PCR
SAXS	small-angle X-ray scattering
SDC	sodium deoxycholate
SDS	sodium dodecyl sulfate
PAGE	polyacrylamid gel electrophoresis
SEC	size exclusion chromatography
SKL	signal sequence for peroxisomal targeting
SOC	super optimal broth medium with glucose
SP6	promoter, regulatory element for the expression with bacteriophage SP6 RNA polymerase
SPR	surface plasmon resonance
ssRNA(+)	positive-sense single-stranded RNA
StrepII-tag	affinity tag recognized by <i>Strep-Tactin</i> [®]

SUMO	small ubiquitin-like modifier
SVLPA	simian virus 40 late polyA terminator
T7	promoter/terminator, regulatory elements for the expression with bacteriophage T7 RNA polymerase
TB	terrific broth medium
TEM	transmission electron microscopy
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus
TF	trigger factor, ribosome-associated chaperone
TMV	tobacco mosaic virus
Tn7	target site-specific transposon
tRNA	transfer RNA
Ulp1	Ubiquitin-like-specific protease 1 from <i>Saccharomyces cerevisiae</i>
US	United States of America
UTR	untranslated region
XFEL	X-ray free-electron laser
X-gal	5-Bromo-4-chloro-3-indolyl β -D-galactopyranoside
ZnB site	Zn ²⁺ binding zinc finger motif

PART I

**Expression, Purification
and Biophysical Characterization
of Pestiviral Autoprotease NS2**

1. Bovine viral diarrhea virus non-structural protein 2

The bovine viral diarrhea (BVD) and the associated mucosal disease (MD) are listed at the World Organization for Animal Health (OIE, Office International des Epizooties) among the 116 most important animal diseases worldwide. The infectious agent behind these diseases is the bovine viral diarrhea virus (BVDV). Infection of cattle with BVDV causes major economic losses, which are mainly due to increased morbidity and mortality, early abortion, stillbirth, congenital deformities, and reduced fertility and milk yield [107, 173, 232].

The causative agent is an enveloped, spherical virus with a monopartite, positive-sense single-stranded RNA (ssRNA(+)) genome (group IV virus, [14]). Together with two other important animal diseases – the border disease virus (BVD) of sheep and the classical swine fever virus (CSFV) – BVDV belongs to the genus *Pestivirus* in the family of *Flaviviridae*. Based on sequence analyses of the highly conserved 5' untranslated region (UTR) of the ssRNA(+) genome and antigenic characterization, BVDV can be further divided into two distinct species, *Bovine viral diarrhea virus 1* and *2* (BVDV-1 and BVDV-2) – often also referred to as genotypes [222, 233]. In addition to these four classified *Pestivirus* species, putative pestivirus isolates from a growing number of different animals like bovine and buffalo (atypical "HoBi-like" virus [242]), pronghorn antelope ("Pronghorn" virus [282]), giraffe ("Giraffe" virus [18]), pigs ("Bungowannah" virus [128]) and even rats [80] and bats [297] have been described and might represent novel independent species.

The superior *Flaviviridae* family itself has four genera: 1) the *Flaviviruses*, which encompass such prominent human pathogens as the eponymous yellow fever virus (YFV), dengue virus (DENV), West Nile virus (WNV) or Zika virus (ZKV), 2) the *Hepaciviruses* with hepacivirus C (HCV), 3) *Pegiviruses* and 4) *Pestiviruses* [260].

Hepacivirus C is an important human pathogen which causes liver cirrhosis and hepatocellular carcinoma as a result of chronic liver infections in millions of cases worldwide. HCV itself is a challenging study object, since it is not easily cultivatable in cell culture. For this reason, BVDV has been implemented as surrogate system for HCV infections, since it is more closely related to hepacivirus C than members of the other genera of the *Flaviviridae* family [40]. Insights into the molecular and structural background and the underlying biophysical principles of BVDV infections therefore provide also additional clues about the mechanism of HCV infections.

1.1. Epidemiology and pathogenesis of BVD/MD

BVDV infections are restricted to cloven-hoofed animals such as cattle, sheep, goats, wild ruminants and pigs and are passed on from one animal to another oronasally via body secretions or during pregnancy by diaplacental infection [19, 276, 129, 171].

The infection with BVDV can result in a wide range of clinical symptoms and outcomes, depending on several variables: 1) the genotype of the infecting strain, 2) its biotype, 3) the immune status of the infected animal and 4) the time of infection.

The two BVDV species 1 and 2 not only differ in their genetic material and antigenicity but also in their epidemiology and encompass a constantly growing number of strains identified from field isolates. Hitherto, 21 BVDV-1 (1a-1u) and 4 BVDV-2 (2a-2d) subgenotypes have been described. The distribution of both BVDV species and their subgenotypes differs from region to region. While for instance 23.5% of the detected species in the United States of America (USA) belong to BVDV strains of genotype 2, in Europe this only applies to about 3.7% of the strains. However, the diversity of European strains is much higher than in the USA. In Europe, genotypes 1a-l and r-u as well as 2a-c have been detected. In the USA, only genotypes 1a, 1b, 2a and 2b have been found so far [302].

Beside these genetically determined differences, each strain can in addition exist in two so-called biotypes – the cytopathic (cp) and the non-cytopathic (ncp) biotype –, which exhibit significantly deviating clinical courses (see below).

With its global prevalence and a high level of endemic infection BVDV is epidemiologically very successful. This is due to a prosperous combination of two well known pathogenic strategies: "hit and run" (transient infections) and "infect and persist" (persistent infections). Both strategies are explained in detail in the following sections and illustrated in figure 1.1.

1.1.1. The "hit and run" strategy: acute transient and acute severe infections

The "hit and run" strategy takes advantage of a short window of opportunity, during which a naive host population is susceptible for the infection. Thereafter it is self-limiting, since the hosts either develop immunity or die from the infection. In both cases, the virus is cleared from the population eventually.

Both BVDV genotypes can cause "hit and run" type infections, which are differentiated into acute transient and acute severe infections based on their clinical course (see figure 1.1, upper part). BVDV strains of genotype 1 are in general perceived as clinically less aggressive. Upon infection of a healthy, naive host, these strains cause an acute, transient infection in 10 to 30% of the cases, with symptoms like scours, fever, lack of appetite, oculonasal discharge and erosions at the mucosa of muzzle or vagina. However, in most cases the infection does not exhibit specific symptoms but only general immunosuppression. This smooths the path for secondary infections with other diseases circulating in the population, whose symptoms shroud the underlying BVDV infection. The latter therefore often goes unnoticed. In all these scenarios, the virus is typically cleared by the immune system of the host over the course of the infection and the host develops immunity against similar strains [39].

In contrast, healthy, naive animals infected with a BVDV-2 strain often develop severe symptoms, that resemble those of the mucosal disease (MD, see section 1.1.3) with thrombo- and leukocytopenia and high mortality.

1.1.2. The "infect and persist" strategy: PI animals

The "infect and persist" strategy aims to infect a host and persist in its system for the length of its life, preferably while constantly propagating viral progeny and without killing the host. Therefore, it is imperative for the virus to evade or manipulate the host's immune system. Several viruses like herpes simplex virus (HSV) or human immunodeficiency virus (HIV) have developed successful strategies to persist in their hosts and prevent detection by its immune system. In addition to the "hit and run" strategy, BVDV also applies the "infect and persist" strategy, but in a rather unique way: By infecting its host via diaplacental transmission in a certain window of opportunity before it acquires immunological competence, the host develops immunotolerance against the virus and will not be able to discriminate viral antigens from its own from that point on. Thereby, the virus completely circumvents the immune system of the host.

When a naive dam is infected with ncp BVD virus between day 40 and day 120 of her gestation, she can give birth to such a persistently infected, immunotolerant calf (termed PI animal; see figure 1.1, lower part). These calves can appear perfectly normal and healthy, but they might also become stunted and more susceptible to other diseases. Many of these animals will die due to health issues during the first two years of their life. In contrast, infections with either cp or ncp BVDV until day 40 or with a cp biotype until day 120 will in general lead to the resorption of the embryo or miscarriage, while infections between day 120 and day 160 often result in malformation of the respective developing organs at the time of infection. After this period, the immune system of the unborn calf is sufficiently developed to initiate a normal immune response against BVDV and the calf will be born healthy and immune [174].

PI animals are a constant threat to naive animals, as they massively shed ncp BVD virus during their entire life. Approximately 1 to 2% of cattle worldwide is persistently infected and a constant reservoir for ncp BVDV [107]. In addition, the offspring of a PI dam – if there is one at all – will also always be a PI animal [174]. Although PI animals will not show any immune response to a second infection with a very similar, antigenically homologous BVD virus strain, they can be infected with a more distantly related, antigenically heterologous ncp or cp strain and develop symptoms and a potentially eliminating immune response like naive animals. This response is exclusively directed against the heterologous strain and will leave the persisting strain unharmed. Depending on the immunological differences, also a chronic superinfection with the heterologous strain can arise [38].

However, the infection of a PI animal with a homologous cp BVDV strain has much more severe consequences. Since the animal is not able to launch any immunological response against the virus, an infection with the cp biotype runs its course in its most destructive form, the so-called mucosal disease (MD) [227].

1.1.3. Mucosal disease

MD is a special clinical form of BVDV infection and only occurs in PI animals. During their life span, PI animals are very likely to come down with MD eventually. Its eponymous main symptoms arise from the destruction of keratinocytes in the gastrointestinal epithelium by the

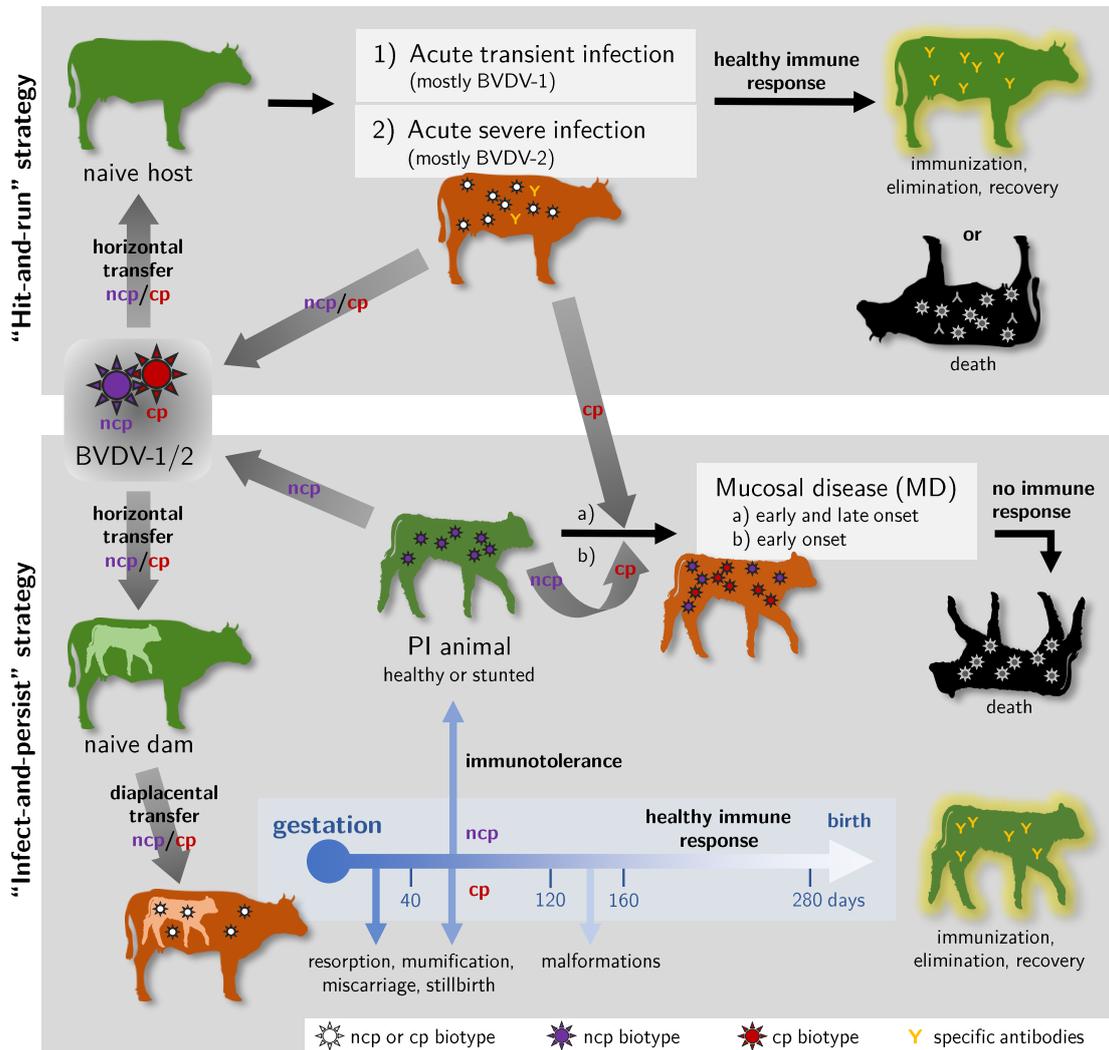


Figure 1.1.: Overview of BVDV infection strategies

BVDV successfully implements a combination of two well-known pathogenic infection strategies. As part of the "hit-and-run" strategy (upper part), naive hosts are infected horizontally via body secretions and develop acute transient or severe infections depending on the pathogenicity of the infecting strain. A healthy immune response results in either full recovery of the immunized host or its death. Both outcomes eliminate the virus from the population. With the "infect-and-persist" strategy (lower part), the virus exploits the developing immune system of a fetal calf diaplacentally infected with a non-cytopathic (ncp) biotype between day 40 and 120 of gestation. The calf (termed persistently infected (PI) animal) is born immunotolerant to the virus and will be a lifelong viral reservoir due to a persistent ncp infection. By superinfection with a cytopathic (cp) strain of a) exogenous or b) endogenous origin, the PI animal will develop lethal mucosal disease eventually. The illustration was prepared with Inkscape™ (version 0.92, <https://inkscape.org/>) based on the information compiled in section 1.1.

cp biotype, which was originally discriminated from the ncp biotype by its visible cytopathic effects on infected cultured cells [13]. The erosion of the gastrointestinal epithelium leads to a drastic loss of fluids, causing non-treatable, often bloody diarrhea and severe dehydration. All mucosal surfaces exhibit more or less pronounced painful lesions and ulceration, the affected animals suffer from anorexia, fever and oculonasal discharge as well as thrombo- and leukocytopenia. Secondary bacterial infections of mucosal lesions are frequent and can give rise to septicemia. Since the animals fail to elicit any specific immune response to eliminate the virus due to their immunotolerant status, the onset of MD in PI animals is always fatal [145].

MD in PI animals results from either exogenic or endogenic superinfection with a cp biotype. If a PI animal – which is persistently infected with and immunotolerant to an ncp BVDV strain – is superinfected with an exogenic, but antigenically homologous cp strain, its immune system will not recognize the foreign cp strain as a threat. After 2 to 3 weeks, this will inevitably lead to so-called early onset MD. However, the cp biotype does not have to be exogenic, it can also develop spontaneously from the already persisting ncp biotype. Since BVDV is an RNA virus, its genome is easily varied by mutations, rearrangements, insertions or deletions. These genomic alterations can give rise to a conversion of the persisting ncp biotype into a cp biotype. The underlying mechanism will be discussed in detail in section 1.3.1. This cp biotype of endogenic origin will likewise cause early onset MD.

Even if the cp strain is antigenically sufficiently different from the persisting ncp strain to elicit an initial immune response, recombination between the persisting ncp strain and the heterologous cp strain can occur at some point. The resulting chimeric virus could exhibit cytopathic behavior while antigenically resembling the persisting ncp biotype and thus evade the host's immune system. Typically, MD onset caused by these recombination events can be delayed for months and is therefore termed late onset MD [83].

1.2. BVDV entry, replication and morphogenesis

The spherical BVDV virion is about 40 to 60 nm in diameter and consists of an enveloped capsid, which encloses the linear ssRNA(+) genome [287]. The lipid bilayer envelope is derived from the ER membrane of the host cell and contains the virally encoded proteins secreted ribonuclease (Erns) and the glycoprotein heterodimer E1/E2, while the capsid is built up from a virally encoded capsid or core protein (C) [47]. Due to its enveloped nature it can easily be inactivated by standard disinfectants.

Upon contact of the virus with epithelium cells of a susceptible animal, an initial contact between cell and virus is established, probably involving the interaction of Erns with the glycosaminoglycans of the target cell surface [112]. Subsequently, E2 in a heterodimer with E1 in the viral envelope binds to the CD46 receptor on the target cell surface [165]. This receptor is a cofactor for the serine protease complement factor I, which normally protects cells against complement-mediated damage. The virion is then internalized by clathrin-mediated endocytosis into intracellular vesicles and delivered to the early endosome [147]. Here, the maturation of the endosome is accompanied by a drastic decrease of the pH, which induces a structural change in the envelope proteins, enabling these proteins to form membrane fusing

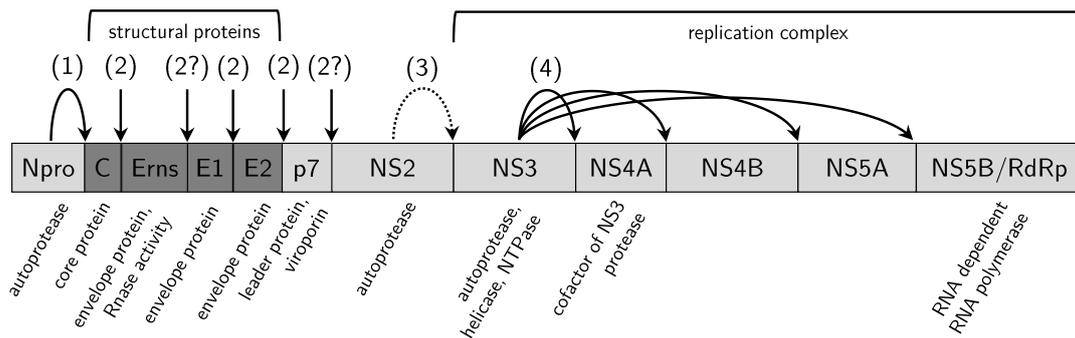


Figure 1.2.: Schematic representation of the BVDV polyprotein

The BVDV polyprotein is processed co- and posttranslationally by (1) Npro, (2) cellular signal peptidases, (3) NS2 protease and (4) NS3 protease. The structural proteins (shaded in dark gray) are part of the infective virion. Several of the non-structural proteins (shaded in light gray) form a genome replicating complex in secluded invaginations of the host's ER membrane. The schematic representation is based on similar illustrations from the web-based knowledge resource ViralZone (Swiss Institute of Bioinformatics, [109]) and the website of Prof. Dr. Norbert Tautz at the Institute of Virology and Cell Biology of Lübeck University. Representation not to scale.

complexes. These complexes mediate the fusion of the virus envelope with the endosomal membrane, thereby releasing the virus capsid and the virus genome into the cytoplasm of the host cell. Since the envelope proteins of BVDV are very resistant to low pH, it is likely that a pre-activating step – presumably involving the breakage of disulfide bonds in the glycoproteins – has to take place during the virion's journey to or through the endosome, which renders the envelope proteins susceptible for a pH-induced structural change [132].

The ssRNA(+) genome released into the cytosol encompasses approximately 12,300 nucleotides. This length varies significantly from strain to strain due to genome duplications or host cell RNA inserts. The genome features a single open reading frame (ORF) enclosed by 5' and 3' UTR and initially acts as messenger RNA. A 5' internal ribosomal entry site (IRES) sequence enables the cap-independent translation of the ORF into a polyprotein by host cell ribosomes at the rough ER [225]. This polyprotein consists of an *N*-terminal nonstructural autoprotease (Npro), which cleaves itself off, 4 structural proteins (C, Erns, E1 and E2) and 6 non-structural proteins (p7, NS2-3, NS4A, NS4B, NS5A and RNA-dependent RNA polymerase (RdRp); see figure 1.2) [57]. The structural proteins are mainly cleaved off the polyprotein by a cellular signal peptidase, while the non-structural proteins are cleaved off by chymotrypsin-like serine protease NS3 in conjunction with its cofactor NS4A [290, 239, 78, 298, 30]. A special case is the cleavage between NS2 and NS3 by autoprotease NS2, which will be explained in detail in section 1.3.

The replication of the ssRNA(+) genome presumably takes place in the cytosol in so-called viral factories. These are clusters of viral proteins involved in genome replication – the replication complex (see figure 1.2) – attached to the surface of rearranged ER membrane

sections. These membrane parts form extended sacs, tubules, spherules or even vacuole-like structures, which are assumed to hide double-stranded RNA intermediates from innate cellular immunity [135, 31, 288, 243]. The genome replication is performed by the virally encoded RNA-dependent RNA polymerase (RdRp, also referred to as NS5B) together with NS3, which also acts as nucleoside-triphosphatase (NTPase) and helicase [306, 91, 94]. The newly synthesized ssRNA(+) can either serve as mRNA for translation, as blueprint for genome replication or is packaged into an assembling virus.

For the virion assembly, the basic C protein binds and condenses the ssRNA(+) to an unordered complex. This complex is then enclosed by ER membrane studded with the glycosylated envelope proteins E1/E2, incorporating the enveloped capsid into the ER lumen. From the ER, the virion is subsequently transported to the *cis*-Golgi system from where it is assumed to bud to be released by exocytosis [93, 158].

1.3. Autoprotease non-structural protein 2

The BVDV ssRNA(+) genome is highly variable and can contain a puzzling number of mutations, rearrangements, deletions and insertions of homologous and heterologous origin. For simplification, the following descriptions refer to the BVDV-1b strain NCP7 (phylogeny based on full-length Npro), unless stated otherwise. NCP7 represents the ncp biotype of the corresponding cp strain CP7 [61, 268] and the amino acid sequence of BVDV non-structural protein 2 (NS2) employed in this thesis originates from this strain. The amino acid numbering throughout this thesis is based on the NCP7 polyprotein, which matches the numbering based on the polyprotein from BVDV-1a strain SD-1 applied in other publications [63].

BVDV NS2 is a 51.5 kDa cysteine autoprotease and encompasses 453 amino acids (aa) of the viral polyprotein [101, 140]. The *N*-terminus is generated at glutamate 1137 (E1137) by a cellular signal peptidase [78]. If activated, NS2 cleaves itself off the following NS3 protein. For the efficient recognition of the cleavage site, the first 7 aa of NS3 are sufficient. The cleavage occurs between arginine 1589 (R1589) and glycine 1590 (G1590). The exchange of either of these two positions to proline (E1589P or G1590P) completely prevents cleavage [140].

The *N*-terminal part of NS2 is very hydrophobic and contains 7 predicted transmembrane helices (see figure 1.3), which are essential for autoprotease activity and embed NS2 in the ER membrane. While an *N*-terminal truncation of 73 aa already affects cleavage efficiency significantly, a truncation of 135 aa prevents cleavage completely [140].

Located in the *C*-terminal part of NS2, histidine 1447 and cysteine 1512 (H1447, C1512; see figure 1.3) have been confirmed as catalytic residues of the autoprotease by mutagenesis analysis. A glutamic acid between these two residues might pose as possible third catalytic residue of a catalytic triad, but has not yet been identified definitely. These findings are in accordance with results obtained from hepaciviruses, where the same conserved histidine and cysteine residues are implicated in catalysis of hepaciviral NS2-3 cleavage [90, 104, 88]. The mutagenesis analysis in BVDV NS2 also revealed four cysteine residues (C1487, C1490, C1500 and C1503; see figure 1.3), which are supposed to be part of a Zn²⁺ binding zinc finger motif (ZnB site) with catalytic relevance [140].

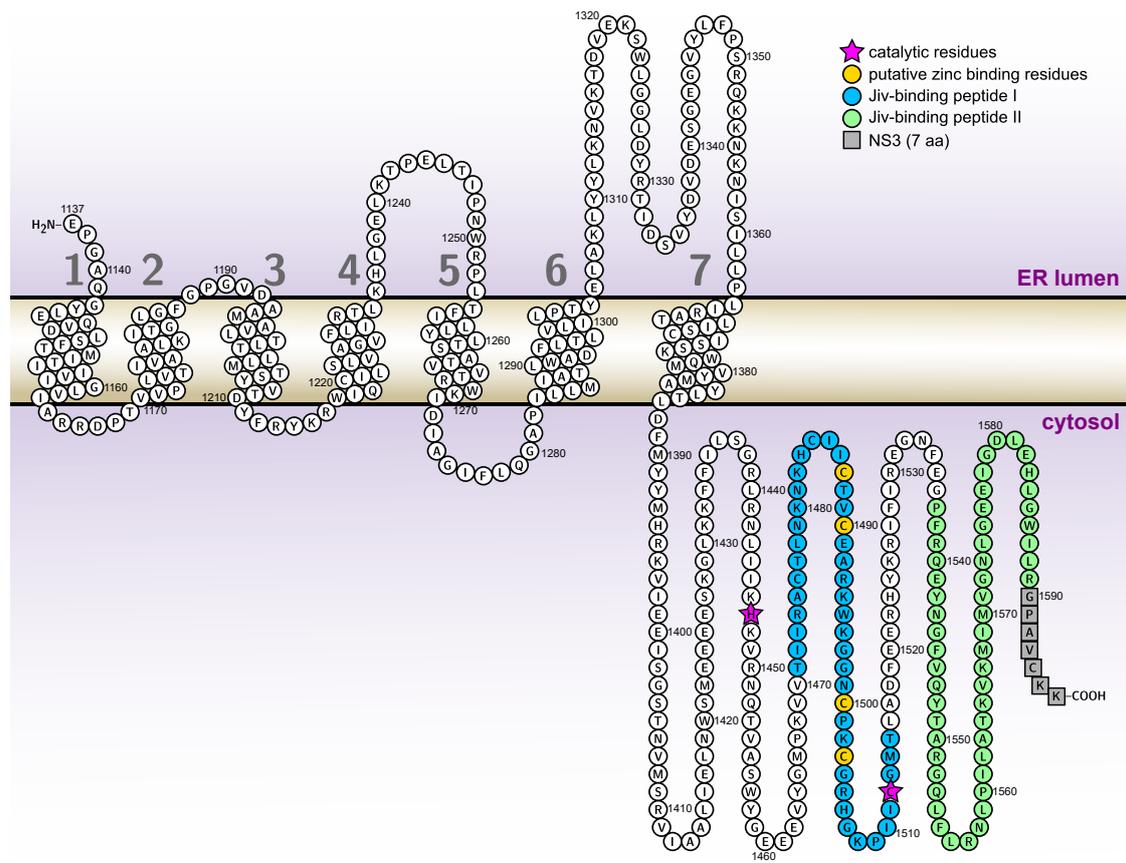


Figure 1.3.: Schematic representation of BVDV NS2 topology

Assumed NS2 topology based on an *in silico* prediction of hydrophobic transmembrane regions and experimental evidence for the location of *N*- and *C*-terminus [78, 305]. The illustration was generated by Protter webserver based on the TExTopo package for L^AT_EX [22, 195]. The visual appearance was optimized with InkscapeTM (version 0.92, <https://inkscape.org/>).

The catalytic residues and the cleavage site as well as their immediate environment are well conserved across pestiviruses, hepaciviruses and pegiviruses. The ZnB site in NS2 is conserved in all known pestiviruses and correlates with a unique hepaciviral ZnB site in NS3, which is involved as accessory domain in the autoproteolytic NS2 cleavage in hepacivirus. This is in contrast to flaviviral NS2, where the NS2A-NS2B and NS2B-NS3 cleavage is mediated by serine protease NS3 without the participation of a ZnB site, and underlines the close functional relation between hepaciviral and pestiviral NS2 [140].

1.3.1. The role of NS2 in the switch from ncp to cp biotype

As outlined in section 1.1, each BVDV strain can exist in two biotypes, a non-cytopathic and a cytopathic one. The ncp biotypes are able to establish persistent infections in hosts that were infected by diaplacental transfer (PI animals, see section 1.1.1). These PI animals

are a lifelong viral reservoir and thereby the cornerstone of BVDV propagation throughout cattle populations worldwide.

Nevertheless, most PI animals will come down with and die from mucosal disease eventually, due to a biotype switch of the persisting ncp BVDV strain or a super-infection with a homologous or heterologous cp strain (see section 1.1.3). In contrast to ncp biotypes, where – after the first hours of infection (see section 1.3.4) – almost exclusively uncleaved NS2-3 can be detected, a high amount of free NS3 is causative for the cytopathic effects and hence the hallmark for an infection with a cp biotype.

These high amounts of NS3 in cp strains can either originate from the expression of separate NS3 from a duplicated genomic region – located in the original RNA genome or in a second, so-called defective interfering RNA [269] – or from efficient cleavage of the existing NS2-3 pool. In the first case, cleavage of the separate NS3 from the polyprotein is mediated by host cell-derived sequences (often ubiquitin) located directly upstream of NS3, which recruit cellular proteases for the task [169]. In the second case, a cp biotype mainly arises from an ncp biotype after the incorporation of mutations or insertions into NS2, which render the autoprotease permanently active. In some strains, like BVDV Oregon, even a few amino acid exchanges can have an activating effect [137]. In others, the region between isoleucine 1529 and glycine 1544 seems to be especially prone to insertions, but they have been reported for other sites as well. The insertions can vary in length significantly. For the CP7 strain for instance, a unique 9 aa peptide inserted after tryptophan 1323 of NCP7 NS2 is sufficient to permanently activate the autoprotease function, resulting in efficient NS2-3 processing and thereby a cp biotype [170, 268]. However, many of these inserts are parts of the coding sequence for the host cell chaperone DnaJ C14 [234]. One of the cp isolates containing a DnaJ C14 fragment is the strain NADL with a 90 aa insert – termed Jiv90 (see below) – after glycine 1536 [167]. In this thesis, constructs with a NADL-like Jiv90 insert in the NCP7 NS2 sequences were employed as representatives for permanently activated NS2.

1.3.2. A cellular chaperone as cofactor of autoprotease NS2

As mentioned above, fragments of the cellular chaperone DnaJ C14, a member of the heat shock protein (Hsp) 40 family, have been identified as inserts in permanently active NS2 correlating with a cp biotype. This is especially interesting, since other studies revealed that the autoprotease NS2 of ncp BVDV is not active on its own, but requires a host cell-derived cofactor for its functionality. This cofactor was found to be the same chaperone DnaJ C14 [235]. It was therefore also termed Jiv. Lackner et al. could demonstrate in 2005, that Jiv (and the Jiv90 fragment derived therefrom, see below) forms stable 1:1 complexes with NS2 and that stoichiometric amounts of this cellular cofactor are necessary for optimal NS2-3 cleavage.

Jiv encompasses 699 aa with a molecular weight of about 78 kDa. In a healthy cell, it functions as a co-chaperone of Hsp70-assisted protein folding. A conserved J-domain in the C-terminal part of Jiv mediates binding to Hsp70. Jiv also features three putative transmembrane helices and resides in the ER membrane, which correlates with the localization of NS2. In 2012, Yi and colleagues could show that actually only two helices span the ER membrane, with both N- and C-terminus located in the cytosol (see figure 1.4) [303]. The human Jiv homolog

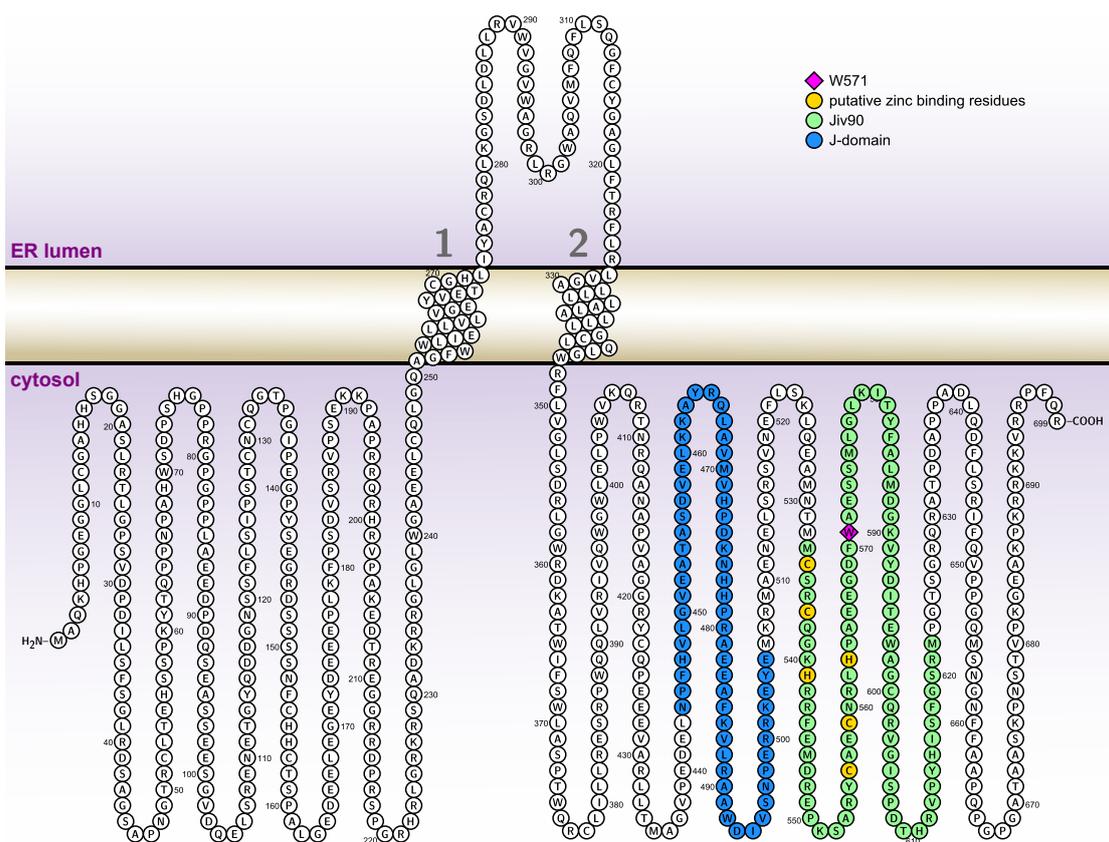


Figure 1.4.: Schematic representation of bovine Jiv topology

Bovine Jiv topology according to Yi et al. 2012, illustration by Protter webserver based on the T_EXtopo package for L^AT_EX[22, 195]. The visual appearance was optimized with Inkscape™ (version 0.92, <https://inkscape.org/>).

is involved in the export of some G protein-coupled receptors (GPCRs) with a C-terminal F(X)₃F(X)₃F motif – like dopamine D1 receptor (DRD1) – from the ER to the cell surface (therefore also termed Dopamine receptor-interacting protein of 78 kDa, DRiP78) [27]. The region, from where the Jiv90 fragment in cp strain NADL is derived, is located downstream of the J-domain from methionine 533 (M533) to methionine 622 (M622) and encompasses two C(X)₂C(X)₃H motifs [235]. These motifs are highly conserved in eukaryotic Jiv and might represent a Zn²⁺ binding domain. Analogous C(X)₂CXGXG motifs in homologous DnaJ from *Escherichia coli* have been shown to bind Zn²⁺ ions and thereby interact with denatured protein substrates [266].

As Rinck and colleagues demonstrated in 2001, the Jiv90 fragment offered in *trans* is sufficient to stimulate ncp BVDV-derived NS2-3 cleavage to a degree comparable to that of full-length Jiv. Further experiments revealed that an N-terminal truncation of Jiv90 of additional 10 aa ($\Delta N10$), which includes the first C(X)₂C(X)₃H motif, reduces cleavage by 50%. An N-terminal truncation of 20 aa as well as a C-terminal truncation of 10 aa abolish

cleavage, while a C-terminal truncation of 5 aa on its own has no detrimental effect. However, $\Delta N10$ and $\Delta C5$ truncations together also abolish cleavage completely [139].

Furthermore, an alanine scan of Jiv90 revealed only one amino acid residue that was crucial to cleavage induction: tryptophan 571 (W571, amino acid numbering according to full-length bovine Jiv). An exchange to alanine at this position abolishes cleavage completely, but phenylalanine, tyrosine and even histidine could restore cleavage efficiency to 86, 78 and 3 %, respectively. Since this was not true for arginine, glutamic acid or proline, an aromatic side chain seems to be required for functionality [139].

The inability of the Jiv90 truncation constructs and the W571A mutation described above to induce NS2 autoproteolysis is not due to impaired binding to NS2, as co-precipitation assays confirmed. All alanine mutants as well as Jiv90 with truncations as large as 50 N-terminal or 30 C-terminal amino acids still co-precipitated with NS2 [139].

1.3.3. The protease and substrate domains of NS2

In 2006, Lackner et al. were able to describe the different functional sections of the NS2 autoprotease in more detail. Truncation and co-precipitation analyses in combination with a cleavage efficiency assessment revealed that the N-terminal half of NS2 encompasses two regions, which independently interact with the cofactor fragment Jiv90. These regions are termed Jiv-binding peptides I (JBP-I, T1471-T1515) and II (JBP-II, P1537-R1589; see figure 1.3). As already mentioned above (see section 1.3.1), the region from I1529 to G1544 located between JBP-I and JBP-II can contain inserts with a varying length from 16 to 157 aa [234, 168]. It is therefore presumed to be a flexible linker between the N-terminal protease part – encompassing JBP-I, the putative JBP and the catalytic residues – and the C-terminal substrate domain containing JBP-II [138].

Full-length NS2 is only active in *cis*, producing its own C-terminus. However, truncation experiments with different parts of the autoprotease confirmed that protease and substrate part can be separated into two independent constructs. As a result, the protease part is able to act in *trans* on the separate substrate peptide in the presence of Jiv90 [138].

The necessary substrate peptide for *trans* cleavage encompasses the last 53 aa of NS2 (P1537-R1589) and the first two amino acids of NS3 (G1590-P1591). A further N-terminal truncation of 3 aa or the deletion of both NS3-derived amino acids abolishes *trans* cleavage. While an N-terminal elongation to R1505 – covering the linker region and an N-terminal portion of JBP-I – has no detrimental effect, further elongation to S1402 and the concomitant incorporation of the putative ZnB site of JBP-II into the substrate prevents cleavage in *trans* [138].

The separate protease is active in *trans* with C-terminal truncations between 4 and 58 aa. Deletion of another six amino acids leads to inaction. Interestingly, truncations shorter than 4 aa interfered with *trans* cleavage, so that full-length NS2 is also not able to cleave the separated substrate part offered in *trans*. Co-precipitation experiments showed that these protease derivatives lost the ability to bind the substrate peptide in the presence of Jiv90. Based on these findings, an intramolecular product inhibition mechanism was proposed, in which the C-terminal 4 aa of NS2 remain in the substrate-binding pocket of the protease after autoproteolytic cleavage and thereby prevent binding and cleavage of further substrates in

trans [138]. A similar inactivating position of the C-terminus has also been reported for the crystal structure of NS2 protease from closely related HCV [156].

Co-precipitation assays with *trans*-active protease and the substrate peptide revealed that Jiv90 interacts with both fragments independently as an essential member of a stable protease/substrate complex required for cleavage. Based on these findings, the abolished *trans* cleavage observed with substrate peptides N-terminally elongated to S1402 or truncated to Q1540 mentioned above could be traced back to the inability of Jiv90 to form a stable protease/substrate complex with these fragments [138].

In conclusion, the host cell-derived cofactor Jiv controls the NS2-3 cleavage in *trans* by promoting the assembly of a stable and functional complex composed of protease, substrate and cofactor. Conferred to the situation in the authentic viral NS2-3 protein, Jiv is responsible for the repositioning of active site and substrate peptide into a cleavage-competent conformation [138].

1.3.4. Temporal modulation of NS2 autoproteolytic activity

The autoproteolytic activity of NS2 in ncp biotypes relies on the presence of the host cell-derived cofactor Jiv, as summarized in the previous section. In contrast to the prevalent assumption that only cells infected with a cp biotype feature free NS3, in 2004, Lackner et al. were able to demonstrate that efficient NS2-3 cleavage and therefore free NS3 could also be detected in cells infected with ncp biotypes in the first 5 to 6 h p.i. (post infection). After this period, cleavage efficiency decreased significantly to a barely detectable level at 8 to 9 h p.i. and below the detection limit eventually. In cells infected with the corresponding cp biotype, cleavage efficiency stably remains at a high level at the same time.

Since the host cell-derived cofactor Jiv is required 1:1 in stoichiometric amounts for autoproteolytic cleavage of NS2-3 and stays bound to NS2 afterwards, the limited cellular Jiv pool was the likely reason for the drastic decrease of cleavage efficiency after 6 h p.i.. This hypothesis is in line with the observation that overexpression of cellular Jiv induces cp characteristics in cells infected with an ncp biotype [235, 177]. In addition, a direct dependency of NS2-3 cleavage efficiency on intracellular Jiv levels could be proved [139].

1.3.5. Significance of NS2-3 cleavage for RNA replication and virion morphogenesis

The consumption of cellular Jiv over the first hours of infection by NS2-3 cleavage does not only result in an obvious decrease of NS3 release, but thereby also causes a significant down-regulation of viral RNA replication [139]. Experiments with cleavage deficient active-site mutants (C1512A, H1447A) of the CP7 strain revealed that – while the authentic CP7 generated high virus titers in cell culture – no viral progeny could be detected for the cleavage deficient mutants [140]. As earlier studies already implied, RNA replication requires free NS3 as essential component of the BVDV replication complex, which cannot be functionally replaced by either NS2 or NS2-3 [20, 140]. However, while free NS3 is vital for RNA replication, virion morphogenesis depends on uncleaved NS2-3 [3]. Therefore, a delicate balance between uncleaved and cleaved NS2-3 has to be established, since both proteins are essential for the generation of infectious progeny.

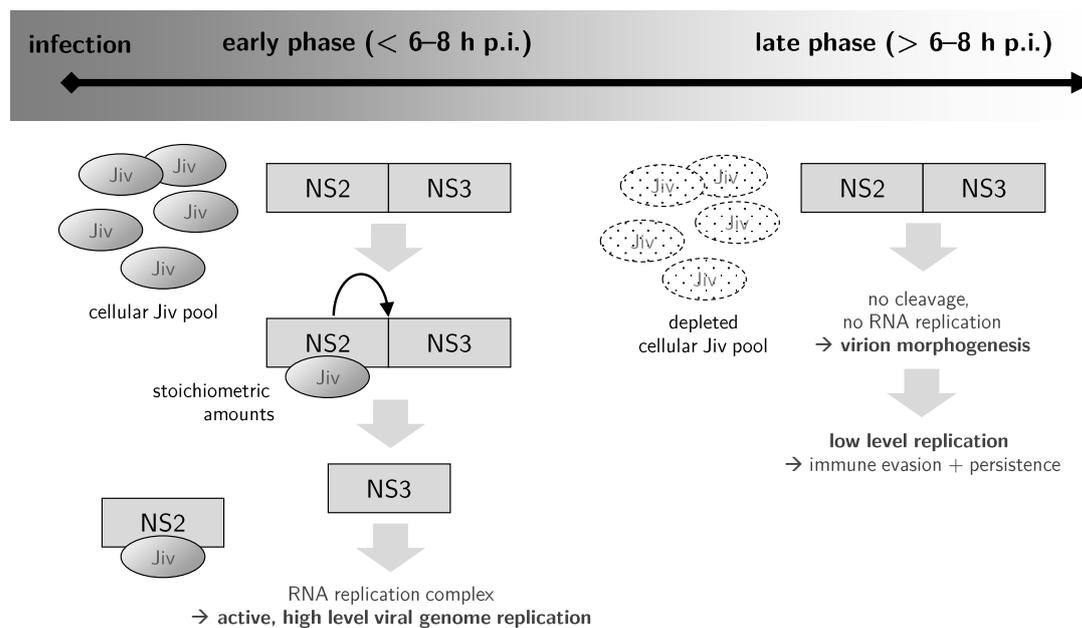


Figure 1.5.: Temporal modulation of ncp BVDV NS2 by cellular cofactor Jiv

The presence of cellular cofactor Jiv in the early phase of an ncp BVDV infection enables autoproteolytic NS2-NS3 cleavage by stable, stoichiometric binding of Jiv to NS2. Free NS3 as part of the RNA replication complex promotes high level genome replication. Without cofactor, NS2 is not active and with the depletion of the cellular Jiv pool the autoproteolytic NS2 activity is diminished. Thereby, while uncleaved NS2-NS3 is available for virion morphogenesis, RNA replication is regulated down, leading to a low replication level – a crucial prerequisite for immune evasion and persistence. The illustration was prepared based on the information compiled in sections 1.3.4 and 1.3.5.

In conclusion, the autoprotease NS2 is the pivotal element in the establishment and maintenance of lifelong persistent BVDV infections in cattle, which represent the key to an impressively successful virus propagation on a worldwide scale. The autoproteolytic activity of NS2 is tightly controlled by the endogenous amount of the cofactor Jiv, since NS2-3 cleavage requires the formation of stable 1:1 complexes between NS2 and Jiv (see figure 1.5). In addition, due to an intramolecular product inhibition mechanism NS2 is restricted to *cis* cleavage and can thus only induce the release of exactly one NS3 molecule, which further tightens the control of Jiv. When the amount of free cellular Jiv decreases in the course of the first hours of infection, so does NS2 cleavage efficiency and thereby the amount of released NS3. This results in a down-regulation of viral RNA replication, which is a prerequisite for the establishment and maintenance of a persistent infection (see figure 1.5 for a schematic summary).

cp biotypes derived from persisting ncp biotypes with a permanently active NS2 autoprotease (or a separately expressed NS3) no longer depend on cellular Jiv for NS3 release. They exhibit high NS3 and RNA replication levels, which is accompanied by cytopathic effects and the onset of lethal MD in the afflicted animals [169, 167, 177, 140]. Since a dead host is

epidemiologically a dead end for the virus, the elaborate control mechanism explained above is an essential prerequisite for viral survival and propagation.

1.4. Aim of this thesis

The NS2 cysteine autoprotease of bovine viral diarrhea virus (BVDV), which is a close relative of the human pathogen hepatitis C virus (HCV), plays a key role in the clinical course of two associated, economically important diseases of livestock breeding: bovine viral diarrhea (BVD) and the corresponding mucosal disease (MD). In conjunction with the host cell-derived cofactor Jiv, NS2 autoprotease controls viral persistence and represents the molecular switch that changes the biotype of a BVDV strain from its non-cytopathic to a cytopathic form. Since this is a unique pestiviral feature, NS2 therefore represents a promising drug target.

Although already many details of the prerequisites for its autoproteolytic activity have been elucidated, the three-dimensional structure has not been determined yet. Reliable homology modeling based on already existing structures also failed so far, due to the lack of structurally characterized, homologous proteins in the databases, denying a deeper insight into the underlying molecular mechanisms. To date, all analytical experiments have only been performed in mammalian cell culture, which is not suitable for the production of larger amounts of heterologous protein. However, especially biophysical characterization methods and techniques for the determination of protein structures – like X-ray diffraction experiments based on protein crystals – require considerable amounts of pure protein.

In this thesis, the heterologous expression of full-length and truncated variants of BVDV NS2 was intensively studied in two popular non-mammalian expression systems – prokaryotic *E. coli* and eukaryotic insect cells. A variety of expression constructs and – in the case of *E. coli* expression – a broad range of strains and expression conditions have been screened in order to establish successful expression of NS2 autoprotease in a non-mammalian host. The immediate aim was to identify the optimal combination of expression host and NS2 expression construct, as well as the best conditions for the production of NS2 autoprotease with a provable native tertiary structure, since this is a crucial prerequisite for all further biophysical and structural characterization experiments.

2. Results

2.1. Expression of BVDV NS2 autoprotease in *E. coli*

2.1.1. Design and cloning of NS2 expression constructs

The BVDV NS2 autoprotease is a transmembrane protein, which most likely comprises 7 *N*-terminal transmembrane helices. After expression from the viral ssRNA(+) genome and further processing of the resulting polyprotein, NS2 is localized in the ER membrane of the bovine host cell. All putative catalytically relevant residues, namely H1447 and C1512, as well as the putative zinc binding site with residues C1487, C1490, C1500 and C1503 are located in the *C*-terminal part of the protein, which resides in the cytosol and does not contain any transmembrane helices (see section 1.3).

Since the expression of a eukaryotic protein with several transmembrane helices in *E. coli* is not a trivial matter, three *N*-terminally truncated constructs were designed, omitting several of these helices. To improve solubility and enable affinity purification, an *N*-terminal His₆-tag followed by SUMO (small ubiquitin-like modifier, *Saccharomyces cerevisiae*) was introduced. The resulting constructs were named for the number of the *N*-terminal amino acid they started with, based on the NCP7 polyprotein (see section 1.3): pET-His₆-SUMO-[1286]NS2* with two, pET-His₆-SUMO-[1351]NS2* with one and pET-His₆-SUMO-[1393]NS2* with none of the putative transmembrane helices. These constructs code for the NS2 amino acid sequence from a non-cytopathic biotype. To reflect the situation in cytopathic biotypes, the coding sequence for the so-called Jiv90 insert, a 90 amino acid portion of the bovine DnaJ C14 protein which was found in a cytopathic biotype, was inserted at the respective site. This led to three additional constructs, termed pET-His₆-SUMO-[1286]NS2_{Jiv90}*, pET-His₆-SUMO-[1359]NS2_{Jiv90}* and pET-His₆-SUMO-[1393]NS2_{Jiv90}*. All 6 constructs were kindly provided by the Institute of Virology and Cell Biology of Prof. Dr. Norbert Tautz at the Lübeck University. Figure 2.1 provides a schematic overview of these constructs.

2.1.2. Expression of His₆-SUMO constructs of *N*-terminally truncated NS2 in *E. coli*

First expression trials under a range of standard expression conditions of these 6 constructs in *E. coli* BL21 (DE3) and BL21 (DE3)pLysS – a strain that reduces basal expression of fusion proteins under control of the T7 promoter and therefore is especially suited to express heterologous proteins, which are toxic for *E. coli*, – did neither yield any overexpression bands on SDS-PAGE nor specific signal on a western blot probed for the His₆-tag (data not shown). Since the coding sequences were not codon optimized with respect to the expression host, an unfavorable codon usage could not be ruled out. Indeed, an *in silico* codon usage analysis (see section 7.6.2) revealed several critical codons which could be responsible for the lack of expression (see section A.1 figure A.2).

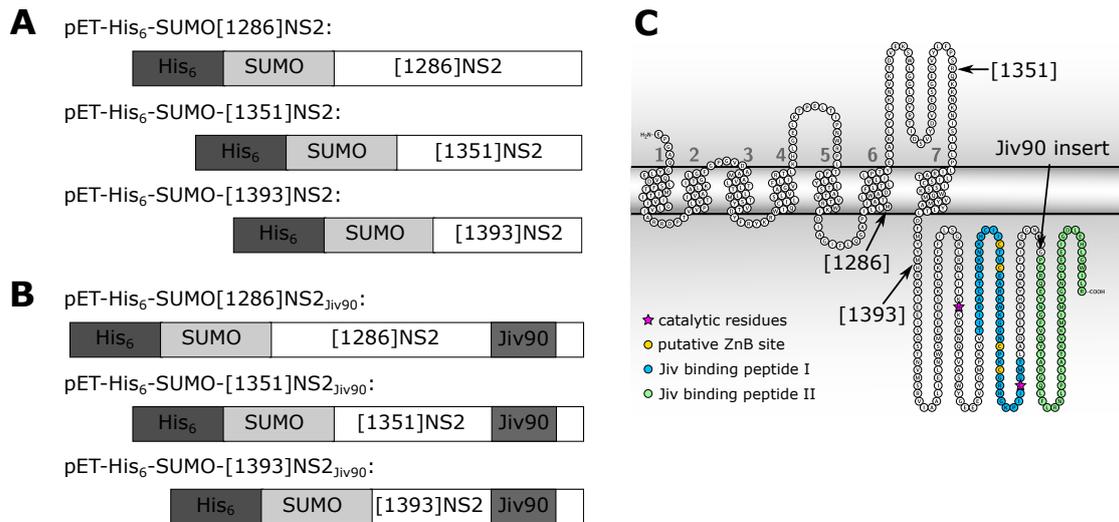


Figure 2.1.: Constructs for expression of *N*-terminal truncated NS2 in *E. coli*

A: NCP7 NS2, *N*-terminally truncated at the amino acid position denoted in square brackets and fused to His₆-tagged SUMO. **B:** NCP7 NS2 with NADL-like Jiv90 insert at G1536, *N*-terminally truncated at the amino acid position denoted in square brackets and fused to His₆-tagged SUMO. All constructs not to scale. **C:** Overview of full-length NCP7 NS2, black arrows indicate the position of the *N*-terminal truncations denoted in square brackets and the location of the Jiv90 insert.

To overcome this problem, the constructs were transformed into BL21 (DE3)RIL and Rosetta™ 2(DE3), which provide additional tRNAs for the most critical codons in *E. coli* on a separate, chloramphenicol resistance-conferring plasmid. The constructs were expressed under a range of standard conditions and whole cell lysate was analyzed on SDS-PAGE and western blot. Example results are depicted in figure 2.2.

In **A**, whole-cell lysates from *E. coli* BL21 (DE3)pLysS and RIL containing the pET-His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* constructs are shown exemplarily. Here, the cells were either induced with 0.25 mM IPTG at an OD₆₀₀ of 0.6 (lanes 5–8) or grown as uninduced control (lanes 1–4). The expression was performed at 37 °C for 3 h. In lanes 7 and 8, distinct overexpression bands are visible at apparent molecular weights of about 35 and 50 kDa, respectively. This is in accordance with the expected molecular weights of 34.8 and 45.2 kDa. The same constructs expressed in BL21 (DE3)pLysS (lanes 5 and 6) did not yield any visible overexpression. For the other 4 constructs also no significant overexpression could be achieved in neither BL21 (DE3) or BL21 (DE3)pLysS, nor RIL.

To verify the identity of the overexpression bands from **A** lanes 7 and 8, a western blot detecting the His₆-tag with an AP-conjugated antibody was performed (see sections 7.7.3 and 7.7.4). The result is depicted in figure 2.2 **B**. Unfortunately, the employed anti-His₆-tag antibody usually exhibits a lot of unspecific reactivity, as is visible in lanes 1 and 3. Nevertheless, the bands in question are most intensively stained, allowing the conclusion that the detected proteins indeed possess a His₆-tag.

2.1. Expression of BVDV NS2 autoprotease in *E. coli*

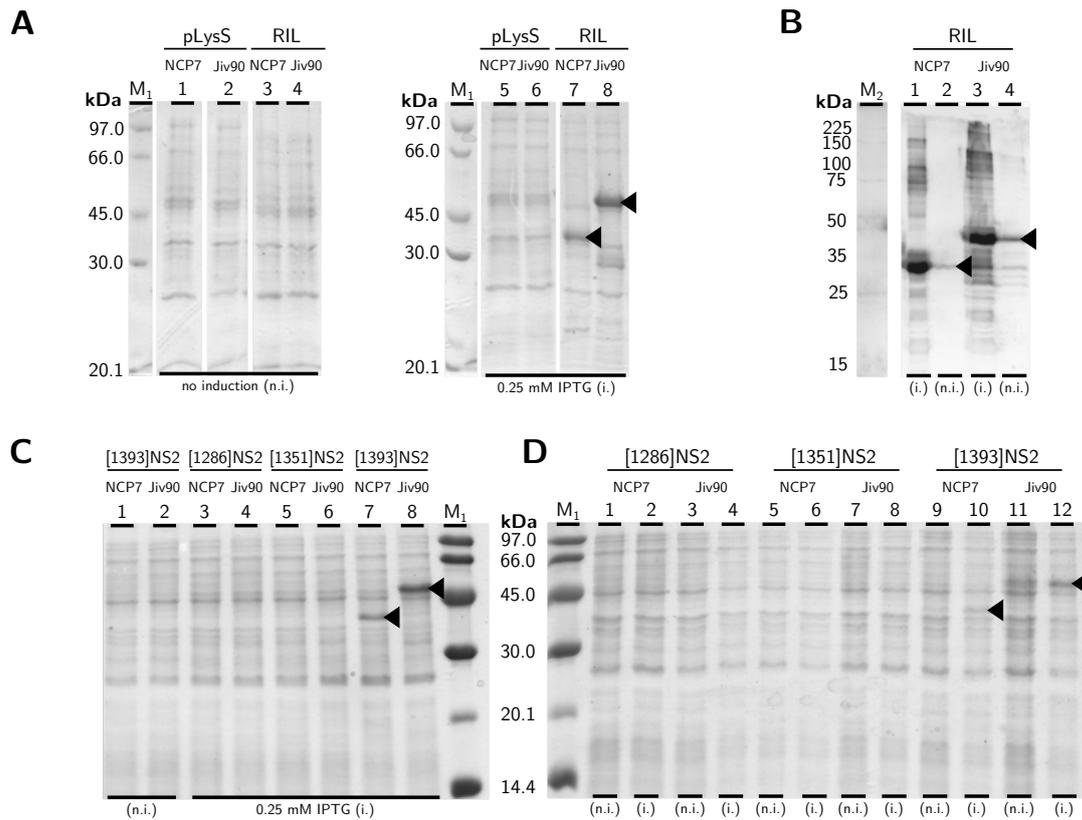


Figure 2.2.: Expression of *N*-terminal truncated NS2 in *E. coli* BL21 (DE3)pLysS, BL21 (DE3)RIL, RosettaTM 2(DE3) and RosettaTM 2(DE3)pLysS

A: Whole-cell lysate from His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* expression in *E. coli* BL21 (DE3)pLysS and RIL, with and without induction. Expression was performed for 3 h at 37 °C. **B:** Western blot of samples from **A** lanes 3, 4, 7 and 8. Detection of His₆-tag with primary anti-His₆-tag mouse IgG antibody and AP-conjugated goat-anti-mouse-IgG secondary antibody. **C** and **D:** Whole-cell lysate from expression of His₆-SUMO-[1286]NS2*, -[1351]NS2* and -[1393]NS2* and their corresponding Jiv90 counterparts in *E. coli* RosettaTM 2(DE3) (**C**) or RosettaTM 2(DE3)pLysS (**C**), with and without induction. Expression was performed for 4.5 h at 37 °C. The expected molecular weights of His₆-SUMO-[1286]NS2*, -[1351]NS2* and -[1393]NS2* are 47.2, 39.8 and 34.8 kDa, respectively. The expected molecular weights of the corresponding Jiv90-containing constructs are 57.7, 50.3 and 45.2 kDa. (i.): Induction with 0.25 mM IPTG at an OD₆₀₀ of 0.6. (n.i.): No induction. Black arrowheads indicate the overexpression bands. M₁: Molecular weight marker (MWM; Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare). M₂: MWM (Trail MixTM Western Markers, Merck-Millipore).

In figure 2.2 **C**, whole-cell lysates from *E. coli* RosettaTM 2(DE3) containing the pET-His₆-SUMO-[1286]NS2*, -[1351]NS2* and -[1393]NS2* as well as the corresponding Jiv90-containing constructs are shown. Again, the cells were either induced with 0.25 mM IPTG at an OD₆₀₀ of 0.6 (lanes 3–8) or grown as uninduced control (lanes 1 and 2). The expression was performed at 37 °C for 4.5 h. As for the induced BL21 (DE3)RIL from **A**, there are

overexpression bands visible for the His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* constructs at about 40 and 50 kDa, respectively (lanes 7 and 8). All other samples from the larger constructs did not exhibit overexpression. Expression in RosettaTM 2(DE3)pLysS (see C) did not improve the protein yield. Again, only the His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* constructs exhibited visible overexpression.

As described above, the strains BL21 (DE3)RIL, RosettaTM 2(DE3) and -pLysS, which all supply additional, rare tRNAs, enable the overexpression of His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* constructs. This supports the assumption that unfavorable codon usage is indeed preventing successful overexpression in BL21 (DE3) and BL21 (DE3)pLysS – at least for the two constructs without any putative transmembrane helix.

2.1.3. Codon-optimization of N-terminally truncated NS2 constructs and expression in different *E. coli* strains

First small scale purification trials from protein expressed in BL21 (DE3)RIL and RosettaTM 2 (DE3) led to the conclusion that the expressed protein was exclusively localized in inclusion bodies (data not shown). This normally occurs when the folding machinery of *E. coli* is not able to process the expressed protein properly, so that unfolded or partly folded protein is collected into inclusion bodies to save the cell from detrimental effects of free, aggregating protein. The fast expression of huge amounts of heterologous protein can be one reason for this behavior. In addition, it is possible that the prokaryotic expression host is not fit to handle the complex folding of a eukaryotic protein, especially if necessary posttranslational modifications cannot be realized by its own enzymes.

To address the speed and amount of expressed protein, the expression conditions were varied in the direction of less intensive induction (decreasing IPTG concentration, auto-induction medium) and slower overall expression (decreasing expression temperature). Nevertheless, the expressed protein was still localized in inclusion bodies.

To assist *E. coli* with the folding of difficult target proteins, a set of helper plasmids, which code for additional chaperones, can be applied (Chaperone Plasmid Set, TaKaRa). The set consists of the following 5 plasmids: (1) pG-KJE8, (2) pGro7, (3) pKJE7, (4) pG-Tf2 and (5) pTf16. Upon L-arabinose- and/or tetracycline-induction, these plasmids express different combinations of the chaperone systems DnaK/DnaJ/GrpE (1, 3), GroEL/GroES (1, 2, 4) and trigger factor (TF; 4, 5).

The helper plasmids are retained in *E. coli* via a chloramphenicol resistance. Since the helper plasmid carrying the additional tRNAs in BL21 (DE3)RIL and RosettaTM 2(DE3) also confers a chloramphenicol resistance, these systems cannot be used in conjunction. Nevertheless, the 6 constructs described above (see figure 2.1) were transformed into BL21 (DE3), supplemented with either of the five helper plasmids. Visible overexpression bands could not even be detected for the His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* constructs on SDS-PAGE (data not shown). This supports the above-mentioned assumption that the supplementation of rare tRNAs is essential for the expression of these constructs.

To enable the expression of NS2 not only alongside additional chaperones, but also in a number of other specialized *E. coli* expression strains that do not supply rare tRNAs – like OrigamiTM 2(DE3) or ArcticExpressTM (DE3) – the C-terminal NS2 portion (starting

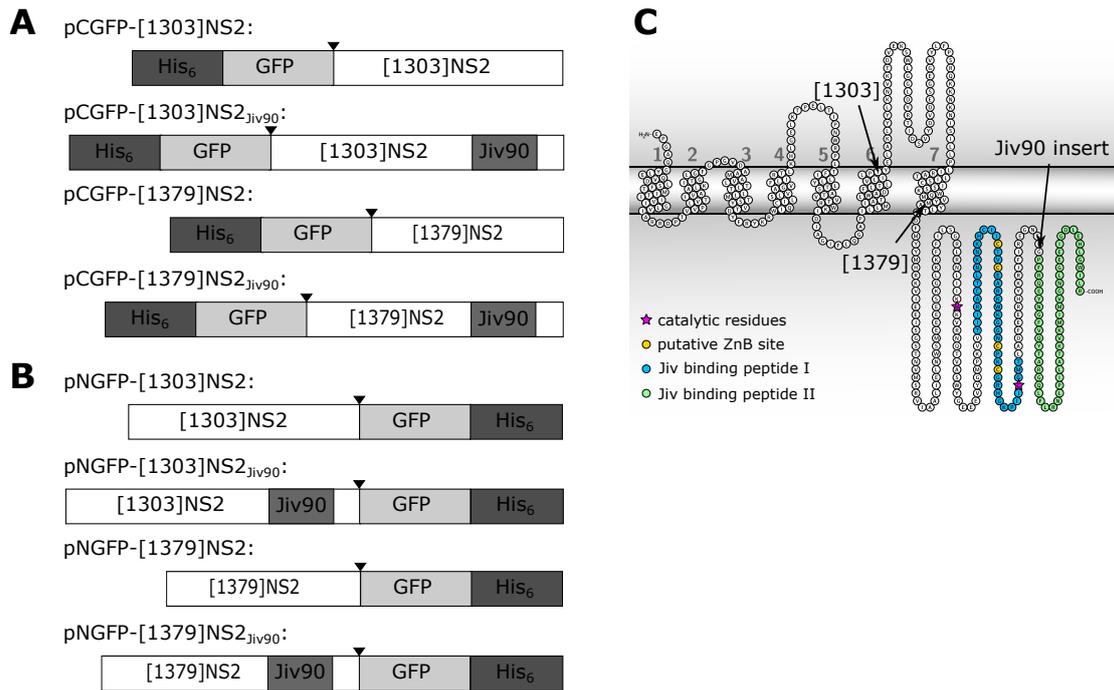


Figure 2.3.: Constructs for *E. coli* expression of *N*-terminal truncated NS2, *N*- or *C*-terminally fused to GFP

A: NCP7 NS2 and NS2 with NADL-like Jiv90 insert, *N*-terminally truncated at the amino acid position denoted in square brackets and *N*-terminally fused to His₆-tagged GFP. **B:** NCP7 NS2 and NS2 with NADL-like Jiv90 insert, *N*-terminally truncated at the amino acid position denoted in square brackets and *C*-terminally fused to His₆-tagged GFP. All constructs not to scale. Black arrowheads indicate the position of a thrombin cleavage site. **C:** Overview of full-length NCP7 NS2, black arrows indicate the position of the *N*-terminal truncations denoted in square brackets and the location of the Jiv90 insert.

at threonine 1303 and containing the Jiv90 insert) was ordered as codon-optimized coding sequence for cloning purposes from GenScript. The codon-optimized sequence without Jiv90 insert was generated from this sequence by deletion PCR (see section 7.8.1). Both *C*-terminal NS2 portions – with and without Jiv90 insert – were used to substitute the corresponding sequences in the His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}* constructs (see section 7.8.2).

In addition to the His₆-SUMO NS2 constructs, a set of 8 new constructs based on the codon-optimized NS2 sequence was designed (see section 7.8.3). These new constructs contained His₆-tagged GFP *N*- or *C*-terminally fused to NS2/NS2_{Jiv90}, which was *N*-terminally truncated at either threonine 1303 (one putative transmembrane helix) or methionine 1379 (no transmembrane helix; see figure 2.3). The GFP fusion is supposed to perform three tasks: First of all, since GFP has a stable fold and is highly soluble, it takes over the role of the solubility mediator SUMO from the first constructs. Secondly, its fluorescence has already been applied successfully as an indicator for correct protein folding – not only of itself, but also of the adjacent fusion protein. Lastly, the GFP fluorescence would simplify the localiza-

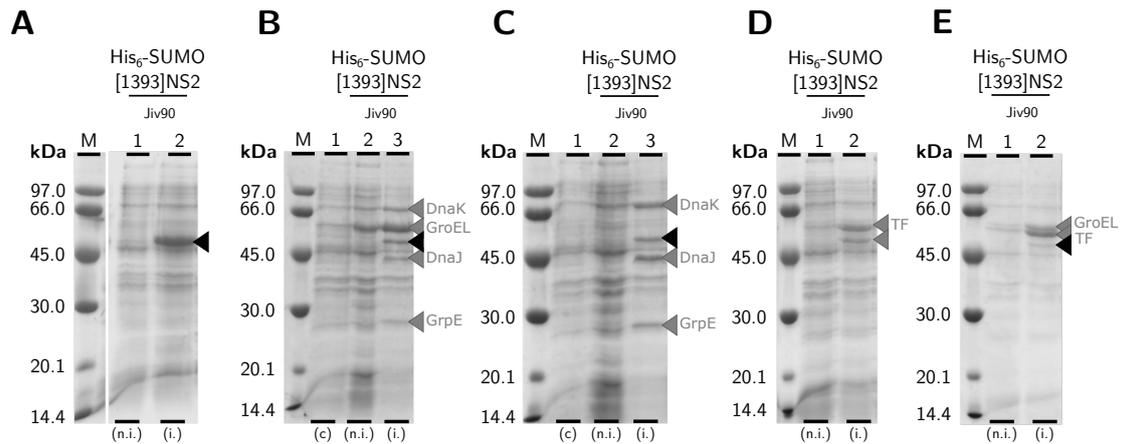


Figure 2.4.: Codon-optimized His₆-SUMO-[1393]NS2_{Jiv90} expression in BL21 (DE3) and derivatives

A: Whole cell lysate from BL21 (DE3) expression. **B:** Whole cell lysate from BL21 (DE3)pG-KJE8 expression. **C:** Whole cell lysate from BL21 (DE3)pKJE7 expression. **D:** Whole cell lysate from BL21 (DE3)pTf16 expression. **E:** Whole cell lysate from BL21 (DE3)pG-Tf2 expression. The expression was performed for 20 h at 20 °C. The construct type and the presence of the Jiv90 insert in the respective clones is denoted above the lanes. Black arrowheads indicate the overexpression bands from His₆-SUMO-[1393]NS2_{Jiv90} with an expected molecular weight of 45.2 kDa. Gray arrowheads indicate the overexpression bands from additional chaperones and are labeled accordingly. (c): Control; the respective strain without construct. (i.): Induction with 0.5 mM IPTG. (n.i.): No induction. M: Molecular weight marker (Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare).

tion of the fusion protein in the sub-cellular fractions. A terminal His₆-tag enables affinity purification of all constructs.

Both His₆-SUMO constructs and the new GFP constructs were transformed into a series of different *E. coli* strains, namely BL21 (DE3) and its chaperone-supplying derivatives containing the plasmids pG-KJE8, pKJE7, pTf16 or pG-Tf2, OrigamiTM 2(DE3) and its derivatives containing the plasmids pKJE7, pTf16 or pGro7, and ArcticExpressTM (DE3). While the OrigamiTM 2(DE3) strains feature mutations in thioredoxin and glutathione reductase genes – thereby enhancing disulfide bond formation in the cytoplasm of *E. coli* –, the ArcticExpressTM (DE3) strain co-expresses the cold-adapted chaperones Cpn10 and Cpn60 (GroEL and GroES homologs) from *Oleispira antarctica*. These chaperones have their activity optimum at 4 to 12 °C and support the folding of slowly expressed protein at low temperatures much better than the corresponding *E. coli* chaperones.

With this set of constructs and expression strains trial expressions were performed under varying conditions. As expected, both codon-optimized His₆-SUMO constructs were now expressible in BL21 (DE3) – and its derivatives – without the supplementation of tRNAs for rare codons (see figure 2.4 A, lane 2). In figure 2.4, example SDS-PAGE results from whole cell lysates of expression trials with His₆-SUMO-[1393]NS2_{Jiv90} in different BL21 (DE3) strains are compiled. Here, the expression was performed at 20 °C for 20 h. All induced samples show

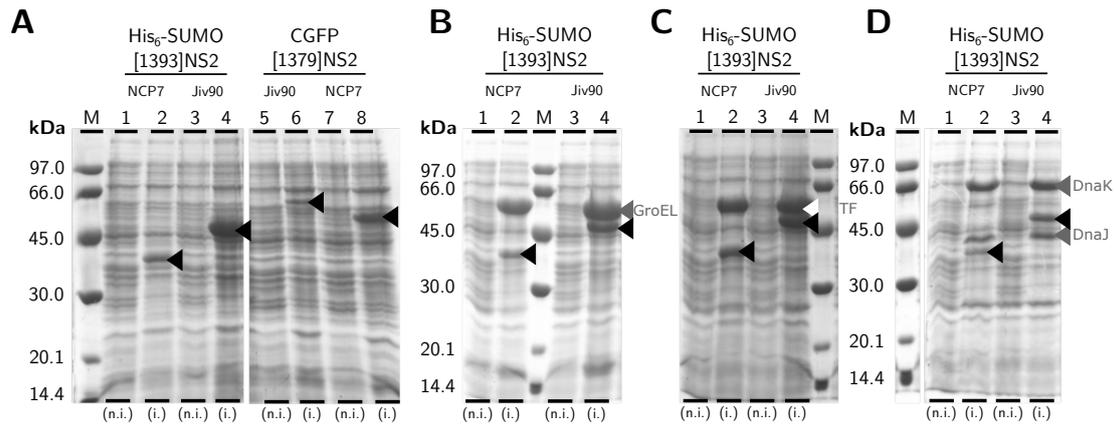


Figure 2.5.: Expression of codon-optimized His₆-SUMO-NS2 and NS2-GFP-His₆ constructs in OrigamiTM 2(DE3) and derivatives

A: Whole cell lysate from codon-optimized His₆-SUMO-[1393]NS2 and [1379]NS2-GFP-His₆ constructs with and without Jiv90 insert expressed in OrigamiTM 2(DE3). **B:** Whole cell lysate from His₆-SUMO-[1393]NS2 constructs with and without Jiv90 expressed in OrigamiTM 2(DE3)pGro7. **C:** Whole cell lysate from His₆-SUMO-[1393]NS2 constructs with and without Jiv90 expressed in OrigamiTM 2(DE3)pTf16. **D:** Whole cell lysate from His₆-SUMO-[1393]NS2 constructs with and without Jiv90 expressed in OrigamiTM 2(DE3)pKJE7. The construct type and the presence of the Jiv90 insert in the respective clones is denoted above the lanes. The expression was performed for 20 h at 30 °C. Black arrowheads indicate the overexpression bands from NS2 constructs. The expected molecular weights of His₆-SUMO-[1393]NS2 and the corresponding Jiv90 construct are 34.8 and 45.2 kDa, respectively. [1379]NS2-GFP-His₆ and the corresponding Jiv90 construct have an expected molecular weight of 53.4 and 63.8 kDa, respectively. Gray/white arrowheads indicate the overexpression bands from additional chaperones and are labeled accordingly. (i.): Induction with 0.5 mM IPTG. (n.i.): No induction. M: Molecular weight marker (Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare).

an overexpression band at 45 to 50 kDa, which is in accordance with the expected molecular weight of 45.2 kDa for the His₆-SUMO-[1393]NS2_{Jiv90} construct (see **B** and **C** lanes 3; **D** and **E** lanes 2). In the strains supplying additional chaperones, their overexpression bands are mostly visible as well. Overall, both His₆-SUMO constructs were expressible in all tested BL21 (DE3) strains, although the overexpression was less pronounced in strains providing additional chaperones. In contrast, all four NS2/NS2_{Jiv90}-GFP-His₆ constructs did not exhibit any visible overexpression on SDS-PAGE (data not shown). The His₆-GFP-NS2/NS2_{Jiv90} constructs were not tested.

OrigamiTM 2(DE3) and its chaperone-supplying derivatives containing pKJE7, pTf16 or pGro7 did all support the overexpression of both codon-optimized His₆-SUMO constructs as depicted in figure 2.5. Here, whole-cell lysates from expressions performed at 30 °C for 20 h were analyzed. In the induced lanes (**A-D** lanes 2 and 4), an overexpression band is visible between the 30 and 45 kDa bands of the molecular weight marker (MWM) for the NS2 construct and slightly above the 45 kDa band of the MWM for the NS2_{Jiv90} construct. This is in accordance with their expected molecular weights of 34.8 and 45.2 kDa, respectively. The His₆-tagged N- and C-terminal GFP fusion proteins did only yield detectable overexpression on

2.1. Expression of BVDV NS2 autoprotease in *E. coli*

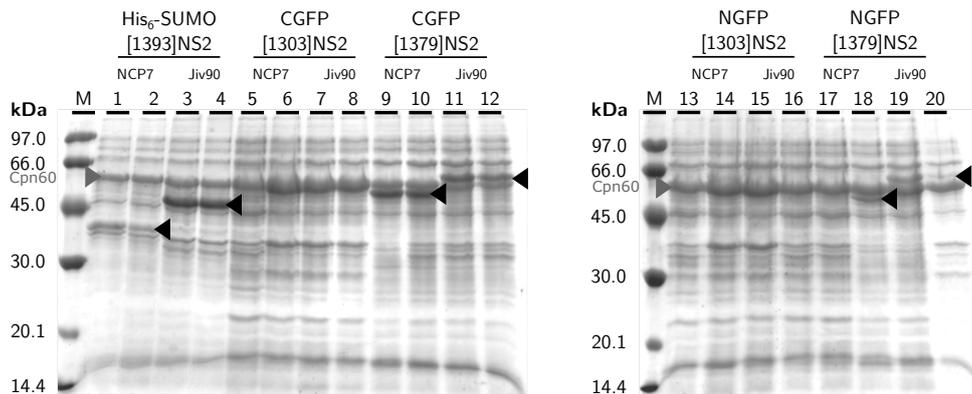


Figure 2.6.: Expression of codon-optimized His₆-SUMO-NS2, NS2-GFP-His₆ and His₆-GFP-NS2 constructs in ArcticExpressTM (DE3)

The construct type and the presence of the Jiv90 insert in the respective clones is denoted above the lanes. The expression was performed for 24 h at 10 °C in autoinduction medium (ZYP-AIM, [265]). Black arrowheads indicate the overexpression bands from NS2 constructs. The expected molecular weights of His₆-SUMO-[1393]NS2 and the corresponding Jiv90 construct are 34.8 and 45.2 kDa, respectively. [1379]NS2-GFP-His₆, His₆-GFP-[1379]NS2 and the corresponding Jiv90-containing constructs have an expected molecular weight of about 54 and 64 kDa, respectively. The expected molecular weights of [1303]NS2-GFP-His₆, His₆-GFP-[1303]NS2 and the corresponding Jiv90-containing constructs are about 64 and 74 kDa. Gray arrowheads indicate the overexpression band from chaperone Cpn60 and are labeled accordingly. M: Molecular weight marker (Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare).

SDS-PAGE for the [1379]NS2/[1379]NS2_{Jiv90}-GFP-His₆ constructs (see figure 2.5 B, lanes 4 and 8). Here, overexpression bands are visible at about 50 and 60 kDa, respectively. This correlates with the expected molecular weight of these constructs of 53.4 and 63.8 kDa. As for the BL21 (DE3) derivatives, also the OrigamiTM 2(DE3) derivatives exhibit overexpression bands corresponding to the additional chaperones GroEL, TF, DnaK and DnaJ.

Since the overexpression of the His₆-SUMO and GFP fusion constructs in the different strains described above still resulted in the production of insoluble protein, the Arctic-ExpressTM (DE3) strain was tried. It is engineered for expression at low temperatures, providing the cold-adapted chaperones Cpn10 and Cpn60 (see above). The low expression temperature results in a very slow expression, and assisted by the highly active chaperones the cellular folding machinery should have enough time to properly fold heterologously expressed protein – hopefully resulting in higher amounts of soluble protein. In figure 2.6, example SDS-PAGE results from whole cell lysates of ArcticExpressTM (DE3) expressing His₆-SUMO and GFP fusion constructs are shown. The expression was performed at 10 °C for 24 h in autoinduction medium (ZYP-AIM [265]). Besides the overexpression band of the Cpn60 chaperone at about 60 kDa, overexpression bands for both His₆-SUMO constructs (lanes 1-4), as well as for [1379]NS2/NS2_{Jiv90}-GFP-His₆ (lanes 18-20) and His₆-GFP-[1379]NS2/NS2_{Jiv90} constructs (lanes 9-12) are visible. The whole-cell lysates of ArcticExpressTM (DE3) expressing

His₆-tagged *N*- or *C*-terminal GFP fusion proteins of [1303]NS2 or [1303]NS2_{Jiv90} did not exhibit any overexpression bands on SDS-PAGE.

In conclusion, both codon-optimized His₆-SUMO constructs were expressible in all tested BL21 (DE3) strains, in OrigamiTM 2(DE3) and its chaperone-supplying derivatives and in ArcticExpressTM (DE3). The shorter variant of the His₆-tagged *C*-terminal GFP fusion protein of NS2 with and without Jiv90 insert did not yield any detectable overexpression in BL21 (DE3) or OrigamiTM 2(DE3) derivative strains, except for unaltered OrigamiTM 2(DE3). Expression for these constructs as well as for the shorter variant of the His₆-tagged *N*-terminal GFP fusion protein of NS2/NS2_{Jiv90} could also be detected in ArcticExpressTM (DE3). The longer *C*- and *N*-terminal GFP fusions did not exhibit any visible overexpression in the tested strains. This correlates with the results from the first expression trials of the two longer *N*-terminal truncations of NS2*/NS2_{Jiv90}* – [1286] and [1351] – in rare tRNA-supplying strains, which also did not show any signs of expression.

2.1.4. Solubilization of *N*-terminally truncated NS2 and NS2_{Jiv90} constructs

Fusion protein from successful expressions (see section 2.1.3) was used to determine optimal buffer conditions for cell lysis, protein solubilization and subsequent affinity purification. A crucial prerequisite for all following experiments is soluble fusion protein, therefore a broad range of lysis buffer conditions and various lysis methods with samples of different expression origin were tried. The success of these experiments was evaluated on SDS-PAGE by comparing the amount of soluble fusion protein from supernatants after centrifugation with the amount of insoluble fusion protein in the pellet fraction.

In figure 2.7, example SDS-PAGE results of solubilization trials are compiled. In **A**, five different buffer chemicals covering a pH range from pH 5.0 to pH 10.0 and two different salt concentrations (100 and 500 mM NaCl) were tested on fusion protein His₆-SUMO-[1393]NS2_{Jiv90} from BL21 (DE3). Unfortunately, the fusion protein cannot be detected in distinct amounts in either of the supernatant samples, although the overexpression band is clearly visible in the pellet samples. This result is representative for both His₆-SUMO-[1393]NS2 fusion constructs from different expression sources (data not shown). In all cases, varying lysis buffer compositions without detergents did not yield soluble fusion protein.

Commercial protein solubilization formulations such as B-PERTM II (Pierce) promise effective cell disruption and solubilization of recombinant proteins without denaturation, based on a proprietary nonionic detergent. The mentioned reagent was therefore tried as well, according to manufacturer's recommendations. In figure 2.7 **B**, SDS-PAGE results from solubilization trials with B-PERTM II on His₆-SUMO-[1393]NS2 and -[1393]NS2_{Jiv90} fusion proteins from OrigamiTM 2(DE3) and its chaperone-supplying derivatives are shown. The overexpression bands of both fusion proteins are clearly visible in the pellet fractions of all samples at about 40 and 50 kDa, respectively – alongside the overexpression bands of the additional chaperones in strains OrigamiTM 2(DE3)pGro7, -pTf16 and -pKJE7. Although the B-PERTM II reagent was able to solubilize some of the additional chaperones (see lanes 10, 12, 14 and 16), the NS2 fusion proteins still remained in the pellet fractions.

Despite all efforts to facilitate expression and folding in *E. coli*, the overexpressed fusion protein still assembles into insoluble inclusion bodies. Since all solubilization trials without

2.1. Expression of BVDV NS2 autoprotease in *E. coli*

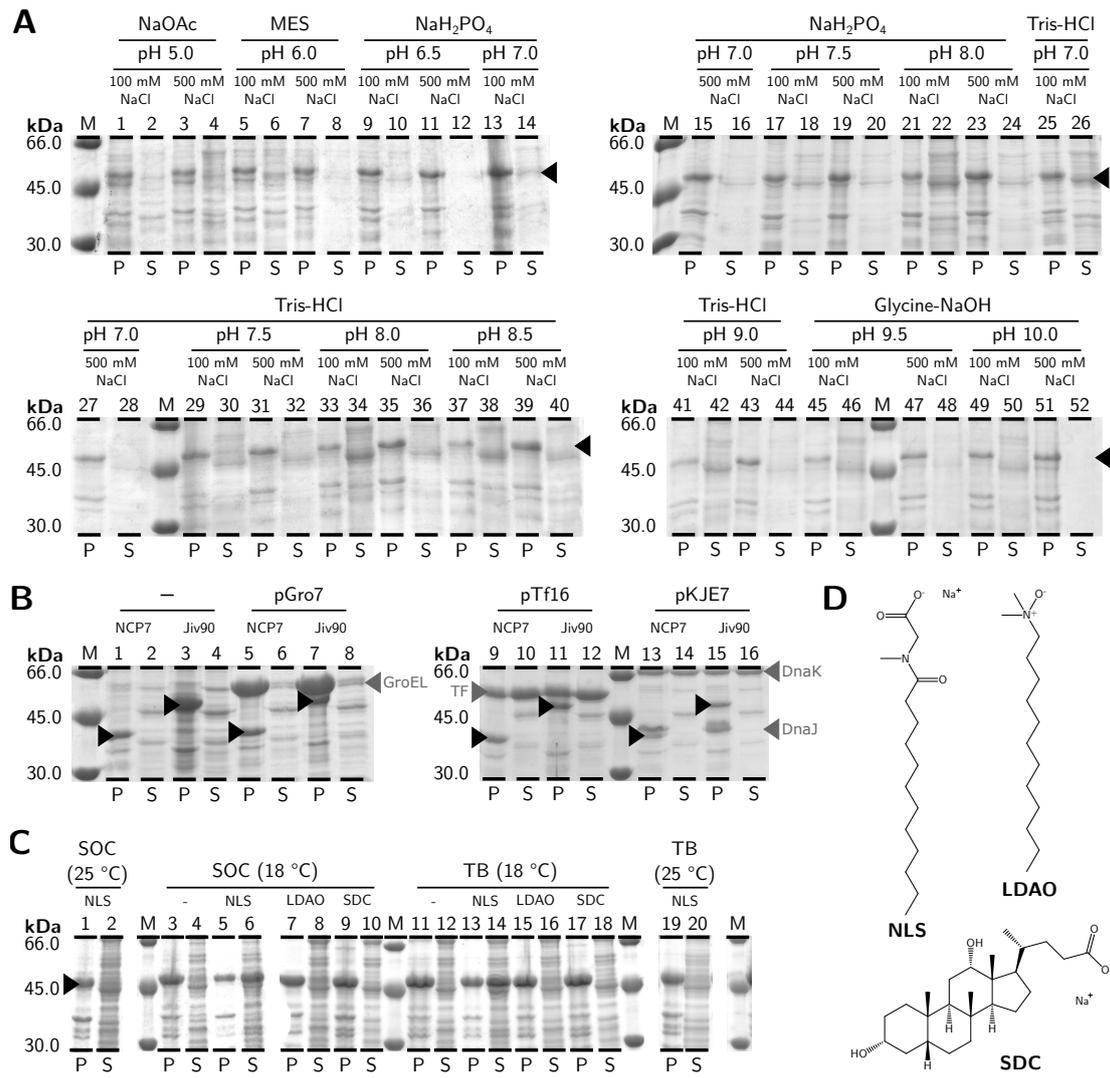


Figure 2.7.: Solubilization experiments with His₆-SUMO-[1393]NS2 constructs

A: Samples from His₆-SUMO-[1393]NS2_{Jiv90} expressed in BL21 (DE3) at 18 °C for 18 h. Solubilization was performed as a combination of lysozyme and freeze-thaw lysis in a variety of buffers at several pH values and with two different salt concentrations (specified above the corresponding lanes). **B:** Samples from His₆-SUMO-[1393]NS2 and -[1393]NS2_{Jiv90} expressed in OrigamiTM 2(DE3) or its derivatives at 20 °C for 20 h. The strain type and the presence of the Jiv90 insert in the respective clones is denoted above the lanes. Gray arrowheads indicate the overexpression bands from additional chaperones and are labeled accordingly. Solubilization was performed with the commercially available B-PERTM II reagent (Pierce). **C:** Samples from His₆-SUMO-[1393]NS2_{Jiv90}* expressed in BL21 (DE3)RIL in SOC or TB medium at 18 or 25 °C for 18 h, as indicated above the lanes. Solubilization was performed using ultrasonication in a phosphate buffer supplied with 0.3 % of the detergent denoted above the corresponding lanes (-: no detergent; NLS: *N*-lauroylsarcosine; LDAO: lauryldimethylamine-*N*-oxide; SDC: sodium deoxycholate). Black asterisks mark the supernatant samples with solubilized fusion protein. Black arrowheads indicate the overexpression bands from NS2 constructs. P: Pellet sample. S: Supernatant sample. M: Molecular weight marker (Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare). **D:** Molecular structure of the detergents NLS, LDAO and SDC.

2.1. Expression of BVDV NS2 autoprotease in *E. coli*

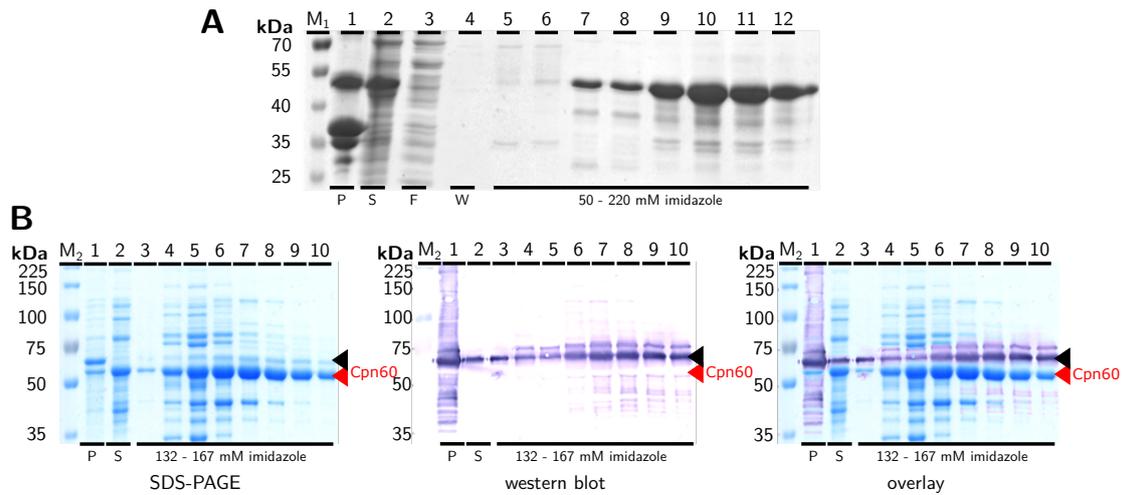


Figure 2.8.: Affinity purification of His₆-SUMO-[1393]NS2_{Jiv90} and [1379]NS2_{Jiv90}-GFP-His₆

A: Samples from His₆-SUMO-[1393]NS2_{Jiv90} expressed in BL21 (DE3) and affinity purified on a HisTrapTM HP FPLC column (GE healthcare) in a phosphate buffer at pH 8.0 supplied with 0.3 % *N*-lauroylsarcosine (NLS). Black arrowhead indicates the overexpression bands from His₆-SUMO-[1393]NS2_{Jiv90}. **B:** Samples from [1379]NS2_{Jiv90}-GFP-His₆ expressed in ArcticExpressTM (DE3) and affinity purified on a HisTrapTM HP FPLC column (GE Healthcare) in a Tris-HCl buffer at pH 7.5 without detergents. Red arrowheads indicate the overexpression bands from the additional chaperone Cpn60 and are labeled accordingly. Black arrowheads indicate the overexpression bands from [1379]NS2_{Jiv90}-GFP-His₆. P: Pellet. S: Supernatant. F: Flow-through. W: Wash. M₁: Molecular weight marker (MWM; PageRuler Prestained Protein Ladder, Thermo Scientific). M₂: MWM (Trail MixTM Western Markers, Merck-Millipore).

any detergent did not yield significant amounts of soluble fusion protein, different detergents were applied to help get the fusion protein into solution. In figure 2.7 C, example SDS-PAGE results are shown. In this experiment, the ionic detergent sodium *N*-lauroylsarcosine (NLS), the nonionic detergent lauryldimethylamine-*N*-oxide (LDAO) and the bile acid salt sodium deoxycholate (SDC; molecular structures see D) were added to a phosphate lysis buffer at a concentration of 0.3 %. With these buffers, samples from His₆-SUMO-[1393]NS2_{Jiv90}* expressed in BL21 (DE3)RIL in SOC or TB medium at 18 or 25 °C were lysed by ultrasonication. In lanes 6 and 14, a fusion protein band can be detected in the supernatant samples. Both samples were lysed with buffer containing 0.3 % NLS. LDAO and SDC were not able to solubilize the fusion protein. Interestingly, only fusion protein from the expression at 18 °C irrespective of the expression medium could be brought into solution (see lanes 2 and 20).

All data from experiments performed in the course of this thesis taken together, only buffers containing NLS were able to transfer a part of the insoluble fusion protein into the supernatant for subsequent purification.

2.1.5. Affinity purification of *N*-terminally truncated NS2 and NS2_{Jiv90} constructs

With NLS-solubilized NS2 and NS2_{Jiv90} fusion protein, affinity purification experiments on HisTrap™ HP FPLC columns (GE healthcare) were performed.

In figure 2.8 A, example SDS-PAGE results of an affinity purification of His₆-SUMO-[1393]NS2_{Jiv90} from BL21 (DE3) are shown. Here, the cells were lysed by ultrasonication in a phosphate buffer at pH 8.0 containing 0.3 % NLS. The cleared supernatant (see lanes 1 and 2 for pellet and supernatant fractions, respectively) was applied onto a HisTrap™ HP FPLC column and washed with 15 CV wash buffer (0.3 % NLS and 20 mM imidazole). By comparing the flow-through sample (lane 3) with the sample from the applied supernatant (lane 2), it is obvious that nearly all solubilized fusion protein – which forms a clearly visible band at about 50 kDa in the pellet and supernatant samples – bound to the Ni²⁺ matrix of the affinity column. The interaction was also stable enough to withstand the 20 mM imidazole (see lane 4). The elution was performed with an imidazole gradient from 20 to 375 mM over 30 CV. The main elution peak of the fusion protein ranged from about 80 to 300 mM imidazole (see A lanes 5 to 12). The elution fractions do not only contain the fusion protein, but also minor impurities in a molecular weight range below the fusion protein. Since the His₆-tag is localized *N*-terminally, these impurities might represent *C*-terminally truncated fragments of the fusion protein, but a simple co-purification of proteins of *E. coli* origin is also possible. Nevertheless, these results prove that the NLS-solubilized fusion protein can be purified by affinity chromatography via the *N*-terminal His₆-tag.

Further affinity purification experiments revealed that the NLS-solubilized fusion protein, which was immobilized on the Ni²⁺ matrix of an affinity column, could not even be eluted with as much as 500 mM imidazole in the elution buffer, if the NLS was gradually removed beforehand from the environment of the fusion protein to enable on-column-refolding of putatively not yet correctly folded protein. A subsequent stepwise increase of the NLS concentration from 0 to 0.3 % enabled elution of the fusion protein at a concentration of about 0.1 % NLS (data not shown). Following experiments supported these results and confirmed that a concentration of 0.08 % NLS was suitable to keep the His₆-SUMO-[1393]NS2 and -[1303]NS2_{Jiv90} fusion proteins in solution.

In general, [1379]NS2-GFP-His₆ and [1379]NS2_{Jiv90}-GFP-His₆ fusion proteins behaved similar to the His₆-SUMO ones with respect to their solubility (data not shown). Nevertheless, since it might be possible that small amounts of the overexpressed protein are soluble by themselves, but do not yield significant bands in the supernatant samples of an SDS-PAGE, solubilization and affinity purification experiments in the absence of NLS were also performed. In figure 2.8 B, the SDS-PAGE of samples from an affinity purification of [1379]NS2_{Jiv90}-GFP-His₆ from ArcticExpress™ (DE3) is shown. This was chosen since ArcticExpress™ (DE3) provides best conditions for very slow expression and a high folding activity due to the cold-adapted Cpn60 chaperone. In addition to the SDS-PAGE, a western blot detecting the His₆-tag with an AP-conjugated antibody was performed (see sections 7.7.3 and 7.7.4) to clarify the identity of the purified protein bands. In this experiment, the bacteria were lysed in a Tris-HCL buffer at pH 7.5 by two passages through a high pressure homogenizer at 1200 bar (EmulsiFlex®-C3, Avestin). By comparing lanes 1 and 2 of SDS-PAGE and corresponding western blot, which contain pellet and supernatant samples, it seems as if only a very small

amount of the fusion protein can be found in the supernatant. Nevertheless, the supernatant was applied onto a HisTrapTM HP FPLC column and washed with 10 CV wash buffer (20 mM imidazole). Subsequently, the bound fusion protein was eluted with an imidazole gradient from 20 to 250 mM over 10 CV. The main elution peak ranged from about 130 to 170 mM imidazole (see **B**, lanes 3 to 10) and the collected fractions exhibited visible GFP fluorescence. The SDS-PAGE samples from this elution peak exhibit a very prominent band at about 60 kDa. As can already be deduced from the SDS-PAGE, this band does not represent the desired fusion protein, but the chaperone Cpn60. This is confirmed in the western blot, where a thin band above the Cpn60 band is colored most intensively and even more obvious in the overlay of both images. Images from SDS-PAGE and western blot are derived from two separate gels run in parallel with identical samples and were aligned via the molecular weight marker.

In conclusion, although a minor fraction of the [1379]NS2_{Jiv90}-GFP-His₆ fusion protein can be found in the supernatant without added detergents, it seems that this fraction is bound to the chaperone Cpn60, which mediates its solubility. Further experiments with Mg²⁺ and ATP in the buffers – which are supposed to help GroEL-like chaperones to release their substrates – did not liberate significant amounts of fusion protein (data not shown).

2.1.6. SUMO protease cleavage experiments on His₆-SUMO-GFP and His₆-SUMO-[1393]NS2_{Jiv90}

For subsequent biophysical characterization and structure determination experiments, the integrity of the native fold of the *N*-terminally truncated NS2 fusion proteins is of utmost importance. Final proof of the correct tertiary structure of the NS2 portion in the expressed and purified fusion proteins would be the detection of its autoproteolytic activity. However, the autoproteolytic activity of native NS2 in bovine cells depends on the presence of stoichiometric amounts of the cellular cofactor Jiv (see section 1.3.2 for further details). For the assessment of the integrity of the expressible [1379]NS2 and [1393]NS2 fusion proteins, the establishment of a reliable cofactor-assisted cleavage assay would be a prerequisite. In contrast, [1379]NS2_{Jiv90} and [1393]NS2_{Jiv90} fusion proteins already contain the activating cofactor fragment in their amino acid sequence. Nevertheless, the original His₆-SUMO constructs do not encompass any protein portion *C*-terminal of the NS2 autoprotease part, which could be cleaved off by active NS2. The [1379]NS2_{Jiv90}-GFP-His₆ fusion protein on the other hand contains the His₆-tagged GFP portion *C*-terminal of the protease cleavage site, but lacks the 7 *N*-terminal amino acids of NS3, that are necessary for the cleavage site recognition by NS2.

However, while all these obstacles could be addressed by optimizations of the expression constructs or the laborious development of a reliable cofactor-assisted cleavage assay, all efforts would probably be futile since a fundamental problem still remains: As Lackner et al. found out in 2004, *N*-terminally truncated NS2 exhibits significantly reduced ([1210]NS2) or even no cleavage activity at all ([1272]NS2; see section 1.3). In the light of these findings, it can be assumed that neither of the NS2 truncations realized in the expression constructs described above is capable of autoproteolytic cleavage.

Therefore, only indirect methods can be applied to assess the folding state of the truncated NS2 autoprotease. In the original constructs the *N*-terminal SUMO is supposed to play the role of a solubility mediator and "draft horse" for the folding process of the following NS2. SUMO is the substrate for the specific SUMO protease, which would only cleave correctly folded SUMO off the following protein. The confirmation of the correct tertiary structure of SUMO in the fusion proteins after the expression and purification procedure would at least allow the assumption that a stably folded protein is not harmed by the chosen conditions. In this case, especially the detergent NLS poses a certain risk, since it is assumed to possess denaturing potential.

To determine the folding status of SUMO in the NS2 fusion proteins, cleavage experiments with Ulp1(C219) SUMO protease (an active, truncated version of Ubiquitin-like-specific protease 1 from *S. cerevisiae*, heterologously expressed in *E. coli*) were performed on His₆-SUMO-[1393]NS2_{Jiv90} in buffers containing 0.08 % NLS. This was the minimum NLS concentration determined in prior experiments to keep the His₆-SUMO fusion protein in solution. As a positive control, His₆-SUMO-GFP was tested under the same conditions and additionally in conditions without NLS. The SDS-PAGE results of this cleavage assay are compiled in figure 2.9 A. Samples were taken over the course of 20 h. To ensure protein stability, Ulp1(C219) SUMO protease (lanes 1 to 4), His₆-SUMO-[1393]NS2_{Jiv90} (lanes 11 and 12) and His₆-SUMO-GFP (lanes 13 to 16) were incubated individually at cleavage conditions. As can be seen in the mentioned lanes, no unwanted protein degradation takes place during the experiment. Unfortunately, Ulp1(C219) SUMO protease exhibits no activity on the His₆-SUMO-[1393]NS2_{Jiv90} substrate (see lanes 9 and 10) in the presence of 0.08 % NLS. This is also true for the control substrate His₆-SUMO-GFP (lanes 7 and 8). Small amounts of cleavage products His₆-SUMO and GFP are visible at about 20 and 30 kDa already immediately after addition of protease to the mixture. However, the amount does not increase anymore during the next 20 h. This implies that the chosen NLS concentration interferes with the cleavage - either by preventing the interaction with the substrate by adhering to relevant surfaces or by actually inactivating the protease or denaturing SUMO. In contrast, Ulp1(C219) SUMO protease is highly active on control substrate His₆-SUMO-GFP in the absence of NLS, as can be seen in lanes 5 and 6. Here, already immediately after addition of the protease to the mixture nearly all substrate is cleaved, and after 1 h, the cleavage is complete (data not shown).

To determine the maximal tolerated NLS concentration for Ulp1(C219) SUMO protease, further cleavage experiments were performed with a range of NLS concentrations. In figure 2.9 B, the SDS-PAGE result of one of these experiments is shown. Prior experiments already narrowed the tolerated concentration down to a range between 0 and 0.025 % NLS (data not shown). In this experiment, it could be specified even further to concentrations below 0.02 % NLS (see lanes 1 to 9). Higher concentrations significantly impaired cleavage (lanes 10 to 13).

Taking these results together, the assessment of the integrity of the NS2 tertiary structure based on the cleavage of SUMO by Ulp1(C219) SUMO protease is not an option in buffers that contain NLS in amounts necessary for the solubilization of the His₆-SUMO-[1393]NS2 and -[1393]NS2_{Jiv90} fusion proteins. Whether NLS actually denatures SUMO, the SUMO

2.2. Expression of BVDV NS2 autoprotease in Sf21 and High Five™ insect cells

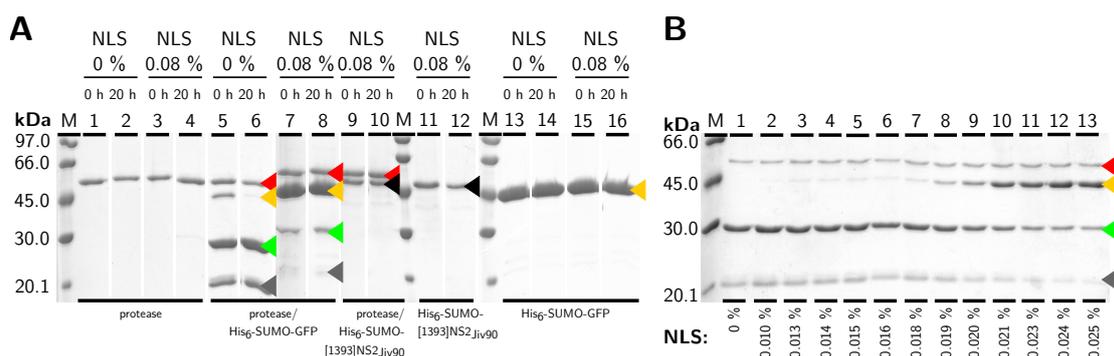


Figure 2.9.: SUMO protease cleavage experiments

A: Cleavage test of Ulp1(C219) SUMO protease without and with 0.08 % NLS on His₆-SUMO-[1393]NS2_{Jiv90} and the control substrate His₆-SUMO-GFP. Samples were taken directly after addition of the protease to the cleavage mixture (0 h), at 1 and 3 h (data not shown), and after 20 h at 30 °C. NLS concentration and sampling time are denoted above the corresponding lanes. **B:** Cleavage test of Ulp1(C219) SUMO protease on control substrate His₆-SUMO-GFP in the presence of a range of NLS concentrations (0–0.025 %, denoted below the lanes), samples taken after 20 h at 30 °C. Red arrowheads indicate the bands from Ulp1(C219) SUMO protease. Yellow and black arrowheads mark the bands from substrate proteins His₆-SUMO-GFP and His₆-SUMO-[1393]NS2_{Jiv90}, respectively. Green and gray arrowheads indicate the bands from cleavage products GFP and His₆-SUMO, respectively. M: Molecular weight marker (Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare).

protease or the truncated NS2 – or just prevents the interaction of the protease and its substrate – cannot be concluded with certainty from these results.

In the light of all described experimental results, the expression and purification of *N*-terminally truncated NS2 in and from *E. coli* poses a number of problems, which cannot easily be bypassed. After all optimization efforts, the greatest drawback is still the impossibility to prove the integrity and native state of the final purification product. This would be a crucial prerequisite as stable starting point for further experiments like biophysical characterizations and crystallization trials for X-ray structure determination.

2.2. Expression of BVDV NS2 autoprotease in Sf21 and High Five™ insect cells

The expression of *N*-terminally truncated constructs of BVDV NS2 autoprotease in *E. coli* proved to be complicated and although several different constructs, strains and expression and purification strategies were tried, no soluble protein suitable for further characterization could be obtained. In addition, *N*-terminally truncated NS2 expressed in baby hamster kidney cells (BHK-21) exhibited significantly reduced ([1210]NS2) or no cleavage activity ([1272]NS2), as described in section 1.3 [140]. This implies that even if the truncated constructs could be obtained in a soluble form, they would not represent the active autoprotease, which reduces the scientific value of subsequent experiments. The expression of full-length NS2 would in contrast ensure full functionality of the autoprotease. Nevertheless, since the *N*-terminal

portion of NS2 encompasses 7 transmembrane helices (see section 1.3) and after the failed soluble expression of truncated constructs, the expression of full-length NS2 in *E. coli* was not an option. The expression system of choice for full-length constructs in this thesis was the baculovirus-mediated expression in insect cells, namely the cell lines Sf21 and High FiveTM.

2.2.1. Design and cloning of full-length NS2 constructs for baculovirus-mediated expression in insect cells

For the baculovirus-mediated expression of proteins in Sf21 and High FiveTM insect cells, several compatible vectors for the transposition of the desired construct into a suitable artificial baculovirus genome (bacmid) have been designed to date. For this thesis, four modified pFastBacTM 1 vectors from the Bac-to-Bac[®] baculovirus system (Thermo Scientific) were applied. These four pFastBacTM 1 vectors were prepared for the expression of the desired construct into different cellular compartments: pFastBac1-cyt for cytosolic localization, pFastBac1-ER with an *N*-terminal signal sequence for ER targeting, pFastBac1-ER-KDEL with the *N*-terminal ER targeting sequence and a *C*-terminal retention signal for permanent ER localization and pFastBac1-SKL with a *C*-terminal signal sequence for peroxisomal targeting. All vectors feature an *Sfo*I restriction site which would allow in-frame blunt-end cloning of the desired construct. The vectors were obtained from the work group Structural Infection Biology applying new Radiation Sources of Prof. Dr. Lars Redecke from Hamburg and Lübeck University.

The sequence of full-length NS2 was constructed from two sources. The *N*-terminal portion (amino acids 1137 to 1302) was derived from the vector pCITE-BVDV-CP7-E2-p7-hNS2-NS3-NS4a, which was a kind gift from the work group of Prof. Dr. Norbert Tautz at the Institute of Virology and Cell Biology of Lübeck University. The sequence in this vector codes for full-length NS2 from the cytopathic strain CP7 and was codon-optimized for the expression in human cell lines. According to a codon usage analysis performed on the authentic NS2 and this sequence in correlation to the codon usage in *S. frugiperda* (see section A.1 figure A.3 (amino acids 1 to 166) and figure A.4), the humanized sequence seemed superior. Nevertheless, since the CP7 NS2 sequence contains an unwanted 9 aa insert at position 1323, the *C*-terminal part of NS2 (amino acids 1303 to 1589) was derived from vectors pUC57-[1303]NS2 or pUC57-[1303]NS2_{Jiv90}, which both contain an NS2 sequence optimized for *E. coli* expression.

A *Strept*II-tag directly fused to the *N*-terminus of the NS2 sequence allows the affinity purification of the expressed protein for further characterization experiments.

To enable easy detection of the expressed protein in cell culture, the full-length NS2 was *C*-terminally fused to mCherry, a bright-red fluorescing monomeric derivative of DsRed from the sea anemone *Discosoma sp.*. mCherry matures fast and is very photostable. Its excitation and emission maxima are located at 587 nm and 610 nm, respectively [255].

Since the final proof for a well behaved, correctly folded NS2 fusion protein would be its autoproteolytic activity, all constructs were provided with the first 7 aa of NS3 between NS2 and mCherry to create a junction which is cleavable by activated NS2 (see section 1.3). Unless the insect host cells are able to provide a Jiv-like cofactor for the activation of NS2, only constructs containing the NADL-like Jiv90 insert should be able to cleave off the mCherry.

2.2. Expression of BVDV NS2 autoprotease in Sf21 and High FiveTM insect cells

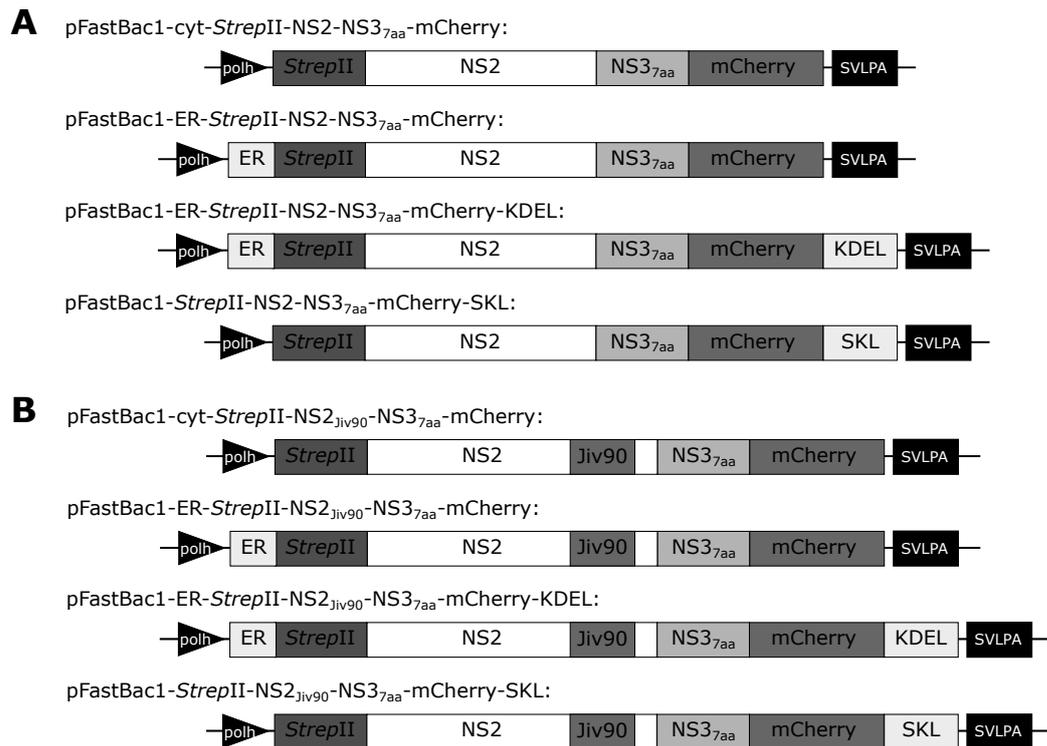


Figure 2.10.: Constructs for baculovirus-mediated expression of full-length NS2 in insect cells

A: Constructs for cytosolic, ER (ER, ER+KDEL) and peroxisomal (SKL) localization. Full-length NS2. **B:** Constructs for cytosolic, ER (ER, ER+KDEL) and peroxisomal (SKL) localization. Full-length NS2 encompassing the Jiv90 fragment should be active and perform the cleavage between NS2 and NS3_{7aa} upon expression. Constructs not to scale. polh: Polyhedrin promoter. SVLPA: SV40 late polyA terminator.

Taken together, 8 constructs with different cellular localizations of the expression product were designed: four with the amino acid sequence of NS2 from the non-cytopathic strain NCP7 and four corresponding constructs with a NADL-like Jiv90 insert (see figure 2.10). The cloning strategy for these constructs is outlined in further detail in section 7.8.4. All pFastBac1 constructs were sequenced to verify the integrity of the expression cassette (data not shown). The expression of the NS2 fusion protein in these constructs is under the control of the strong polyhedrin promoter from the baculovirus *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV).

Subsequently, the vectors were transformed into *E. coli* DH10BacTM or DH10EMBacY. Both strains harbor the baculovirus shuttle-vector bMON14272 or EMBacY, respectively, and a helper plasmid, pMON7124. The transfer of the expression cassette and adjacent regions into the respective bacmid is mediated by the site-specific mini-Tn7 transposon. The bacmid confers a kanamycine resistance and comprises the mini-Tn7 transposon, which contains the coding sequence for the lacZ α fragment. This complements the lacZ Δ M15 mutation in the

β -galactosidase gene of *E. coli* DH10 and allows the expression of functional β -galactosidase. The helper plasmid provides all necessary transposase genes and can be selected for with tetracycline. The expression cassette in the pFastBac1 vectors contains a gentamycin resistance and is flanked by specific sequences, that mediate the transposition of the expression cassette from the pFastBac1 vector into the bacmid – thereby substituting the coding sequence for the lacZ α fragment. The transformed bacteria are streaked onto selective agar plates containing X-gal and IPTG for blue/white screening. Bacteria harboring the pFastBac1 vector, which are resistant to gentamycin but have not undergone transposition yet, will express the lacZ α fragment due to IPTG induction – and thereby functional β -galactosidase. The β -galactosidase converts the X-gal into the blue 5,5'-dibromo-4,4'-dichloro-indigo product, which turns the bacterial colony blue. After a successful transposition, however, the coding sequence for the lacZ α fragment is substituted by the expression cassette, the defective β -galactosidase is inactive and the bacterial colony will appear white. To ensure transposition homogeneity of the clone of choice for all 8 constructs, the clones were re-streaked twice onto fresh agar plates supplemented with the necessary antibiotics, X-gal and IPTG and checked for their color.

Afterwards, the completeness of the expression cassette was verified by colony PCR with primers M13for and M13rev. The result is shown in figure 2.11. Positive clones feature a PCR product in the range from 4.417 bp – for the cytosolic construct without insert – to 4.762 bp for the ER-KDEL construct with Jiv90 insert. Clones without a successful transposition of the expression cassette from the pFastBac1 vector into the bacmid show a PCR product of about 350 bp, which can be seen in figure 2.11 B, lane 12.

Sometimes colonies contain a mixture of bacteria harboring the original and the already transposed bacmid. In these cases, in addition to the 350 bp fragment, the PCR product of the transposed expression cassette would be detectable, too. These colonies should be avoided. According to the results of the colony PCRs depicted in figure 2.11, 44 out of 48 tested colonies show a single band for a successfully transposed expression cassette. Three colonies have no visible band, while one colony contains a single band representing the original empty bacmid. For each expression construct in either of the bacmids one positive clone (indicated by an asterisk in figure 2.11) was chosen and used for bacmid preparation and subsequent transfection into Sf21 insect cells.

2.2.2. Expression of full-length NS2 constructs in Sf21 and High FiveTM insect cells

All 8 bacmids from the clones specified above were transfected into Sf21 insect cells, as elaborated in section 7.13.2. The transfected insect cells were incubated for 7 d in 6-well plates, then resuspended in the medium and harvested by gentle centrifugation. Virus of the first generation (P1) was obtained between 48 to 72 h after transfection by removing the cell culture supernatant and replacing it with fresh medium.

P1 supernatant was used to infect fresh Sf21 insect cells and generate a second virus generation (P2) with a higher virus titer. Likewise, the third virus generation (P3) was derived from fresh Sf21 insect cells infected with P2 supernatant (see section 7.13.3 for details). P3 virus (and in some cases P4) was used for the actual protein expression experiments in Sf21 and High FiveTM insect cells.

2.2. Expression of BVDV NS2 autoprotease in Sf21 and High Five™ insect cells

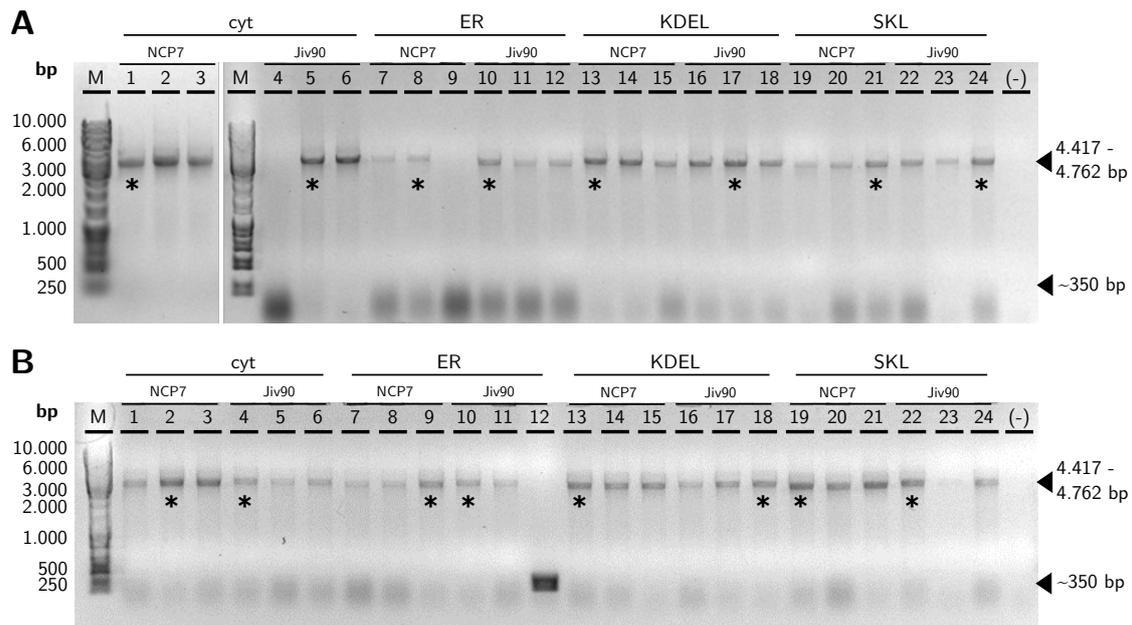


Figure 2.11.: Colony PCR results of *E. coli* DH10EMBacY and DH10Bac™ clones transformed with full-length NS2 pFastBac1 constructs

A: DH10EMBacY clones. **B:** DH10Bac™ clones. The constructs all contain *N*-terminally *StrepII*-tagged full-length NCP7 NS2 (with and without NADL-like Jiv90 insert), fused to the first 7 aa of NS3, followed by mCherry. The type of localization signal and the presence of the Jiv90 insert in the respective clones is denoted above the lanes. The expected PCR product sizes for successfully transposed expression cassettes range from 4.417 bp for the cytosolic construct without insert to 4.762 bp for the ER-KDEL construct with Jiv90 insert, as indicated above. Empty bacmid produces a PCR product with a size of about 350 bp (see **B**, lane 12). M: Molecular weight marker (GeneRuler™ 1 kb DNA Ladder, Thermo Scientific). (-) negative control: PCR reaction without template. Asterisks indicate clones chosen for bacmid preparation and subsequent insect cell transfection.

To check for expression of the desired fusion proteins, the EMBacY bacmid provides the coding sequence for EYFP under the polyhedrin promoter from AcMNPV – the same promoter that also controls the expression of the NS2 constructs. This causes the simultaneous expression of EYFP and the corresponding fusion protein and allows an immediate estimation of the expression rate of the target protein, even if it does not feature an easily detectable marker itself. Here, the fusion protein contains its own fluorescent marker, a C-terminal mCherry, which indicates successful translation and at least partly successful folding with its fluorescent presence. In addition to the expression marker EYFP, EMBacY possesses some other advancements over the older bMON14272 bacmid, like protease and chitinase deficiencies, extended cell viability and reduced proteolysis, which are supposed to support the expression of high amounts of heterologous protein. Nevertheless, since advancements which are beneficial for the expression of most proteins might be detrimental to the expression of others, and since the presence of EYFP in the cells might be a disadvantage in subsequent experiments, EMBacY as well as bMON14272 were applied in parallel in this thesis.

2.2. Expression of BVDV NS2 autoprotease in Sf21 and High FiveTM insect cells

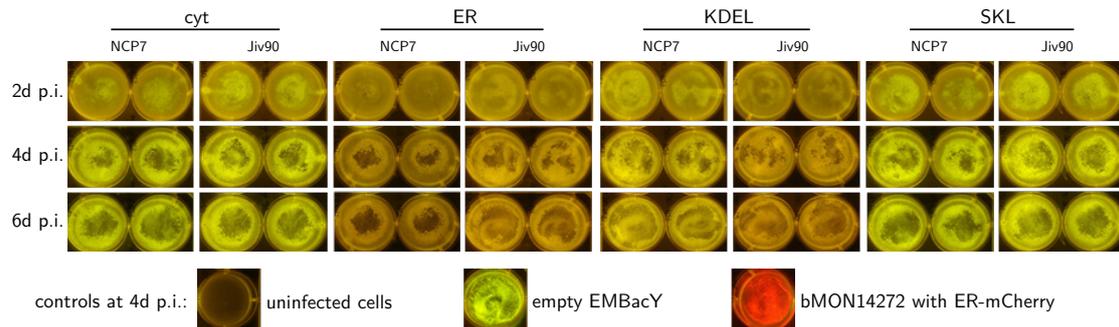


Figure 2.12.: High FiveTM insect cells transfected with full-length NS2 constructs

Images of High FiveTM insect cells infected with P3 EMBacY virus containing the 8 different full-length NS2 constructs. Uninfected cells (no fluorescence), empty EMBacY (only EYFP fluorescence) and bMON14272 with ER-mCherry (only mCherry fluorescence) were chosen as controls. The images were taken with a consumer camera on a Blue/Green LED Transilluminator (Nippon Genetics) with an Amber filter at day 4 p.i. for the controls and at day 2, 4 and 7 p.i. for the NS2 constructs. The type of localization signal and the presence of the Jiv90 insert is denoted above the images of the NS2 constructs.

The expression of the different NS2 constructs in Sf21 and High FiveTM insect cells was monitored during the 7 d incubation period with a Blue/Green LED Transilluminator (Nippon Genetics) and the corresponding Amber filter by visually checking for the fluorescence of mCherry and EYFP (in insect cells infected with EMBacY-derived recombinant baculovirus). In figure 2.12, images taken at day 2, 4 and 7 after infection of High FiveTM insect cells with P3 EMBacY virus from all 8 full-length NS2 constructs are depicted exemplarily. In all cases, first signs of infection – namely EYFP and mCherry fluorescence as well as small disruptions of the confluent cell layer – are already visible at day 2 p.i.. In cells infected with cytosolic and SKL constructs, the EYFP fluorescence is more apparent, while in cells infected with ER and ER-KDEL constructs the expression of the complete fusion protein is proven by the mCherry fluorescence, giving the corresponding wells an orange tint.

2.2.3. In-gel fluorescence of full-length NS2 constructs expressed in High FiveTM insect cells

The insect cells were harvested 7 d after the infection to check for protein expression by SDS-PAGE (see section 7.7.1). They were resuspended in their growth medium and carefully centrifuged to sediment the cells. The medium supernatant was concentrated in 3 kDa MWCO centrifugal filters (Merck-Millipore) to check for protein from already lysed cells. The samples were mixed with a Laemmli buffer variant and heated for 30 min at 37 °C to preserve the integrity and thereby fluorescence of mCherry and EYFP. After the electrophoresis of the samples, the gel was photographed with a consumer camera on a Blue/Green LED Transilluminator (Nippon Genetics) equipped with an Amber filter. Figure 2.13 A and B show the resulting images from High FiveTM insect cells infected with all 8 NS2 constructs in EMBacY bacmid. As expected, all lanes – except the negative control (-) with a sample

2.2. Expression of BVDV NS2 autoprotease in Sf21 and High FiveTM insect cells

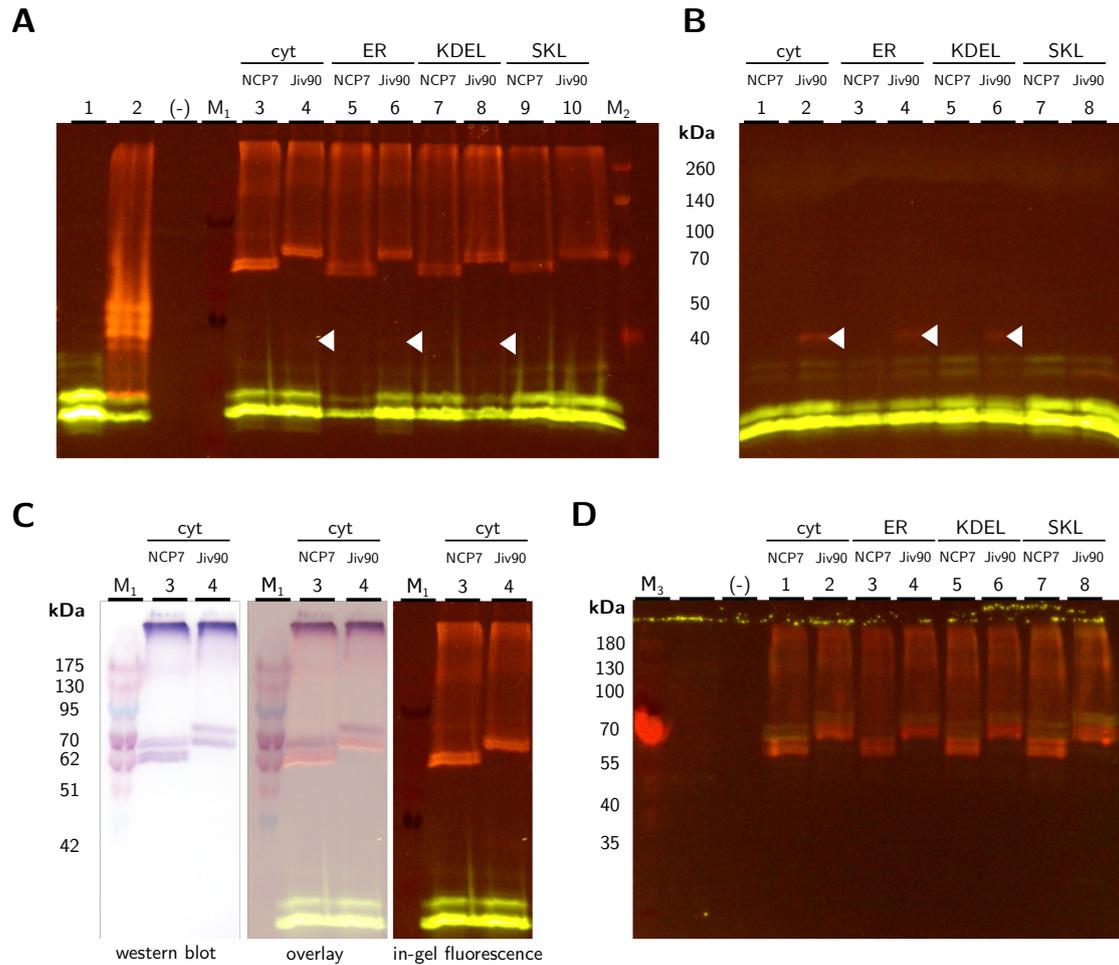


Figure 2.13.: In-gel fluorescence and western blots of full-length NS2 constructs expressed in Sf21 and High FiveTM insect cells

A: Whole cell lysates of High FiveTM cells infected with P3 baculovirus. (-) negative control: uninfected cells; lanes 1 and 2: cells infected with empty EMBacY and ER-mCherry in EMBacY, respectively; lanes 3 to 10: cells infected with NS2 fusion constructs. **B:** Concentrated medium supernatant from High FiveTM cells from **A**. **A** and **B:** White arrowheads indicate the position of the hardly visible mCherry fluorescence. **C:** Western blot of the gel from **A** and overlay of both images. Depicted are only the cytosolic constructs in lanes 3 and 4 from **A**. The *StrepII*-tag of NS2 was detected with a color reaction of AP-conjugated *Strep*-Tactin[®]. **D:** Western blot of whole cell lysates from Sf21 cells infected with P3 baculovirus. The *StrepII*-tag of NS2 was detected with a combination of anti-*Strep* IgG and fluorescent anti-IgG Alexa Fluor[®] 532 antibody. All fluorescent images were taken with a consumer camera on a Blue/Green LED Transilluminator (Nippon Genetics) equipped with an Amber filter. The localization signal and the presence of the Jiv90 insert is denoted above the lanes. The expected molecular weights of the fusion proteins range from 80.5 kDa for the complete cytosolic construct without insert to 93.6 kDa for the complete ER-KDEL construct with Jiv90 insert. After cleavage, the molecular weight of the NS2 portion ranges from 53.7 to 65.3 kDa, while free mCherry is about 28 kDa. M: Molecular weight marker (M₁: Roti[®]-Mark BI-PINK, Roth; M₂: SpectraTM Multicolor Broad Range Protein Ladder, Thermo Scientific; M₃: PageRuler Prestained Protein Ladder, Thermo Scientific).

from uninfected cells – exhibit strong EYFP fluorescence in two bright and two less visible bands in the range below the 40 kDa band of the molecular weight marker (MWM). This is in accordance with the result from High FiveTM cells infected with empty EMBacY bacmid (see figure 2.13 A lane 1), where the same yellow fluorescent bands can be seen. EYFP has a calculated molecular weight of 26.8 kDa and seems to run much faster than could be expected from this estimation. This is due to the mild sample preparation conditions, which did not denature the sample proteins completely to preserve fluorescence. The protein is still folded and therefore much more compact in large parts and can migrate faster through the pores of the PA gel than the completely denatured protein would.

In addition to the EYFP fluorescence, mCherry fluorescence can also be detected in the gels. As a reference for the migration behavior of mCherry, figure 2.13 A lane 2 contains a sample from High FiveTM cells infected with ER-mCherry in the EMBacY bacmid. As expected, the two bright EYFP bands can be detected here as well. The mCherry fluorescence is streaked from the beginning of the separating PA gel to the higher one of the two bright EYFP bands and shows three bright bands and one less visible band around the same height as the 40 kDa band of the MWM. The calculated molecular weight of mCherry is about 28 kDa.

In the NS2 construct samples in image A lanes 3 to 10, the main mCherry fluorescence is located at the 70 kDa band of the MWM for the four constructs containing the Jiv90 insert and at about 60 kDa for the four constructs without insert. Considering the incomplete denaturation status of the protein samples, this is in accordance with the calculated molecular weights of the fusion proteins with and without Jiv90 insert of 90.9 to 93.6 kDa and 80.5 to 83.1 kDa, respectively. In addition to the mCherry fluorescence corresponding to full-length fusion protein, a very faint band with mCherry fluorescence is visible in samples from cells infected with cytosolic, ER and ER-KDEL full-length NS2 constructs with Jiv90 insert (see figure 2.13 A lanes 4, 6 and 8; band indicated by white arrowhead). This band is located directly below the 40 kDa band of the MWM, which correlates with the location of free mCherry (see A lane 2). This band is more prominent in the samples from the concentrated supernatants depicted in figure 2.13 B. Here, no bands corresponding to full-length fusion proteins are visible. This is probably because the fusion protein would be bound to cellular membrane fragments of the lysed cells via the transmembrane helices of NS2. These fragments were very likely centrifuged down with the intact cells or adhered to the membrane of the concentration device, so that the full-length fusion protein only appears in gel A. Nevertheless, the presence of this free mCherry only in samples derived from Jiv90-containing constructs strongly suggests that the NS2 fusion construct is expressed not only complete, but also in an authentic, functional state. The Jiv90 insert was obviously able to render the protease active and induce the cleavage between NS2 and the adjacent 7 aa of NS3 with the following mCherry.

2.2.4. Western blot detection of the N-terminal *StreptII*-tag in full-length NS2 constructs expressed in Sf21 and High FiveTM insect cells

If the NS2 autoprotease is active in Jiv90-containing constructs expressed in Sf21 or High FiveTM insect cells, besides free mCherry also NS2 without NS3_{7aa}-mCherry fusion should be detectable. Therefore, the PA gel from figure 2.13 A was blotted onto a PVDF membrane

(see section 7.7.3) and probed for the *N*-terminal *StreptII*-tag of the NS2 fusion construct with *Strep*-Tactin[®] coupled to alkaline phosphatase (AP; see section 7.7.4). The presence of AP on the blot was detected with a chromogenic reaction, yielding a permanent purple color at the location of the *StreptII*-tag. In figure 2.13 C, lanes 3 and 4 of the developed blot are shown. For better comparability, the blot is overlaid with the corresponding fluorescence image from A (also depicted in C on the right). Intense purple color is visible at the area which represents the beginning of the separating gel. This is very likely due to the hydrophobic, membrane-associated nature of the NS2 fusion protein, which favors aggregation of the fusion protein over the migration into the PA gel under the applied conditions. Interestingly, both cytosolic constructs – with and without Jiv90 insert – exhibit two distinct purple bands at about 70 and 80 kDa and 65 and 70 kDa, respectively. In contrast, the fluorescence image has only one band for each construct at about 70 and 65 kDa, respectively. In the overlay image, it becomes clear that only the lower band of the blot of each construct corresponds to the mCherry band from the fluorescence image. This band contains the *StreptII*-tag as well as the mCherry fusion and therefore must represent the complete fusion protein. In contrast, the upper band on the blot does not show any distinct fluorescence in the fluorescence image, leading to the assumption that this protein indeed possesses a *StreptII*-tag, but misses the C-terminal mCherry fusion. In samples derived from Jiv90-containing constructs, this correlates with the presence of free mCherry as illustrated in figure 2.13 A and B, but for NS2 without Jiv90 the existence of a band which contains a *StreptII*-tag but no mCherry fluorescence is not as easily explained.

In a second western blot experiment, samples from Sf21 insect cells expressing the 8 different NS2 fusion constructs were analyzed. The samples were prepared as described for in-gel fluorescence (see above) and separated by SDS-PAGE. Subsequently, they were blotted onto a PVDF membrane (see section 7.7.3) and probed for the *N*-terminal *StreptII*-tag of the NS2 fusion construct with a primary IgG antibody against the *StreptII*-tag. A secondary Alexa Fluor[®] 532-conjugated anti-IgG antibody was used to enable fluorescence detection. The excitation and emission maxima of Alexa Fluor[®] 532 are at 532 and 554 nm, respectively, so that the labeled antibody exhibits a yellow fluorescence. The blot was photographed with a consumer camera on a Blue/Green LED Transilluminator (Nippon Genetics) equipped with an Amber filter. The resulting blot is shown in figure 2.13 D. Due to the mild experiment conditions and its high stability, mCherry still shows its red fluorescence. As for the samples from High FiveTM cells, the mCherry fluorescence is streaked from the beginning of the separating gel to a fluorescent band at about 70 kDa for the constructs containing Jiv90 and at about 60 kDa for the ones without. A fluorescent band at about 40 kDa for the free mCherry is not visible. The yellow fluorescence of Alexa Fluor[®] 532 is located in a band above the mCherry band, corresponding to the upper band of the blot developed with AP-conjugated *Strep*-Tactin[®] (see C). Since the imaging setup with Blue/Green LED Transilluminator and consumer camera is only makeshift, it cannot be differentiated between mCherry and Alexa Fluor[®] 532 fluorescence. Therefore it is not possible to say if the mCherry band also exhibits Alexa Fluor[®] 532 fluorescence, as can be expected from the results of the blot developed with AP-conjugated *Strep*-Tactin[®]. Nevertheless, these results confirm the presence of a second

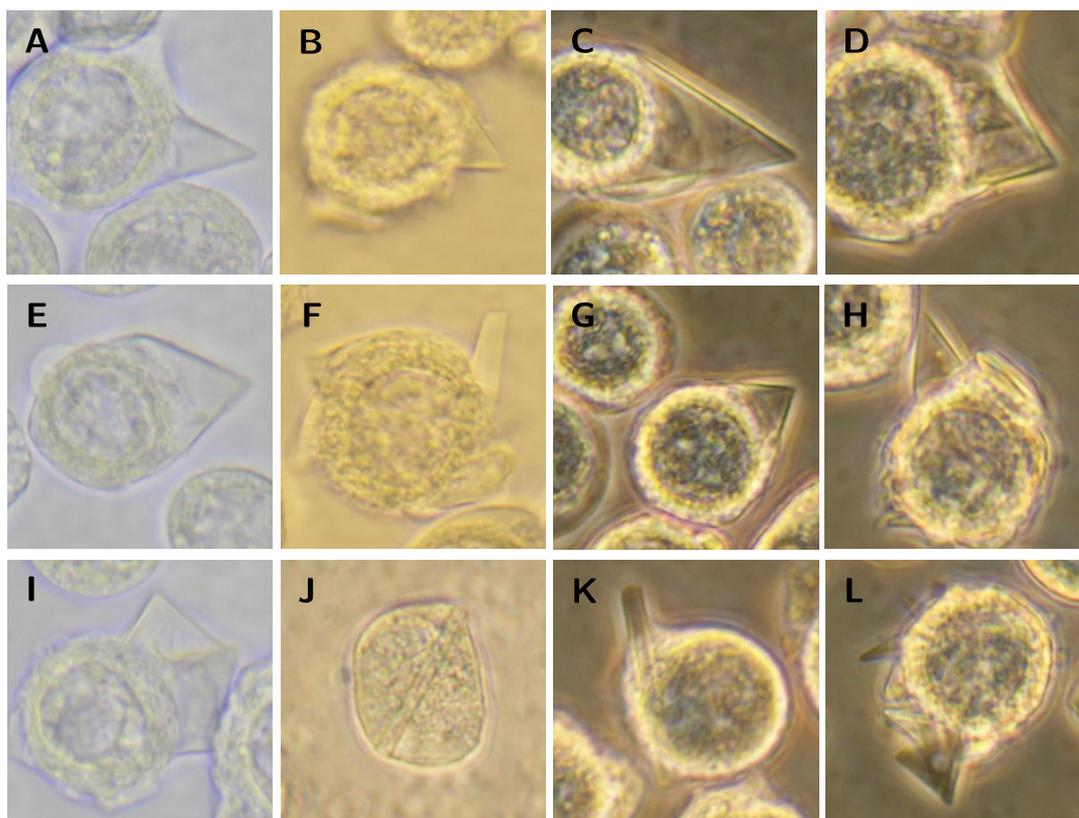


Figure 2.14.: *In vivo*-grown crystals from full-length NS2 constructs expressed in High Five™ insect cells

Different shapes of *in vivo* crystals from full-length NS2 constructs grown in High Five™ insect cells. The images represent crystal examples from cells infected with ER and ER-KDEL NS2 constructs, both with and without Jiv90 insert. Plate-like crystals in the shape of triangles (A–E, G, H, L), rectangles (I) and less frequently small rods (F, J, K) could be found.

StreptII-tagged protein band without mCherry fluorescence directly above the red fluorescing band corresponding to the complete fusion protein in the samples from all 8 constructs.

2.2.5. *In vivo* crystals in High Five™ cells expressing full-length NS2 constructs

During the expression of the full-length NS2 constructs in High Five™ insect cells, unusual crystal-like structures could be detected under the inverted microscope rather frequently (see figure 2.14). Spontaneous crystallization of heterologously expressed protein in insect cells has been reported already for different proteins. In conjunction with recent developments like the X-ray free-electron laser (XFEL) and serial crystallography, this phenomenon can be exploited for their structural characterization [229].

The crystal-like structures in NS2-expressing High FiveTM cells could be reproduced reliably by infecting freshly seeded cells with P3 virus generation (see section 7.13.4). The cells were checked for signs of *in vivo* crystal growth for the following 6 to 8 days after infection at 40x magnification under an inverted light microscope (Olympus CKX41 equipped with an Olympus C-5060 camera). *In vivo*-grown crystal-like structures could only be observed in High FiveTM cells infected with full-length NS2 fusion constructs targeted for ER (ER and ER-KDEL), and much less frequently in those targeted for peroxisomes (SKL). In cells infected with cytosolic constructs, none of the structures could be detected. In High FiveTM cells, fusion protein expressed from both bacmids (bMON14272 and EMBacY) yielded crystals in comparable amounts and shapes. In contrast, individual mCherry expressed as a control from cytosolic, ER, ER-KDEL and SKL constructs in both bacmids did not exhibit any sign of *in vivo* crystal growth. The same applies to mCherry control constructs and all 8 full-length NS2 fusion constructs expressed in Sf21 insect cells, where also no crystals could be detected. Most crystals appeared at day 5 to 7 after infection with P3 or P4 baculovirus generation, and occasionally single crystals were already visible as early as day 3.

In figure 2.14, examples for different shapes and morphologies of *in vivo*-grown crystals from High FiveTM insect cells are compiled. All crystals appear colorless and translucent. In most detected cases, the visible crystal part sticks out of the normal cellular boundaries, but in some cases, the crystal shape is also traceable inside the cell (see figure 2.14, J and K). Frequently, several crystals grow from a single cell (L). The crystals detected so far are most often of triangular (A–E, G, H, L) or even rectangular shape (I), sometimes rod-like (F, J, K). Deduced from microscopic inspection, they appear to consist of very thin plates, occasionally in several layers on top of each other (C and D). They range in size from barely visible to more than 20 µm, about the average diameter of an infected High FiveTM cell (C, E and I). The crystal shapes and sizes do not correlate with the crystal origin, all shapes and sizes could be detected in all crystal-yielding conditions.

2.2.6. Fluorescence microscopy of *in vivo*-grown crystals

In vivo-grown crystals appeared in High FiveTM cells independently of the used bacmid, while the corresponding mCherry control constructs did not exhibit any sign of crystal growth. Therefore, it seemed reasonable to attribute the *in vivo* crystals to the NS2 fusion constructs. To check this hypothesis, the location of the mCherry fluorescence of the NS2 fusion constructs in the infected cells was determined by fluorescence microscopy. 5 to 8 day old infected High FiveTM cells grown in 6-well plates were examined with an inverted fluorescence microscope (AxioVert 200 M, Zeiss) equipped with filters 38HE and 43HE (Zeiss).

Representative fluorescence images are compiled in figure 2.15. In A, High FiveTM insect cells infected with the ER localized NS2 construct without Jiv90 insert in EMBacY bacmid are shown. As expected, the cells exhibit both EYFP and mCherry fluorescence, although not all cells with EYFP fluorescence also feature visible mCherry fluorescence. While the fluorescence of the soluble EYFP is distributed almost uniformly throughout the whole cell (A: EYFP), the mCherry fluorescence is located in a well-defined shell around the presumed nucleus (A, B: mCherry). Considering the shape of the cellular structure in which mCherry fluorescence can be detected, it is likely that mCherry and the corresponding fusion protein

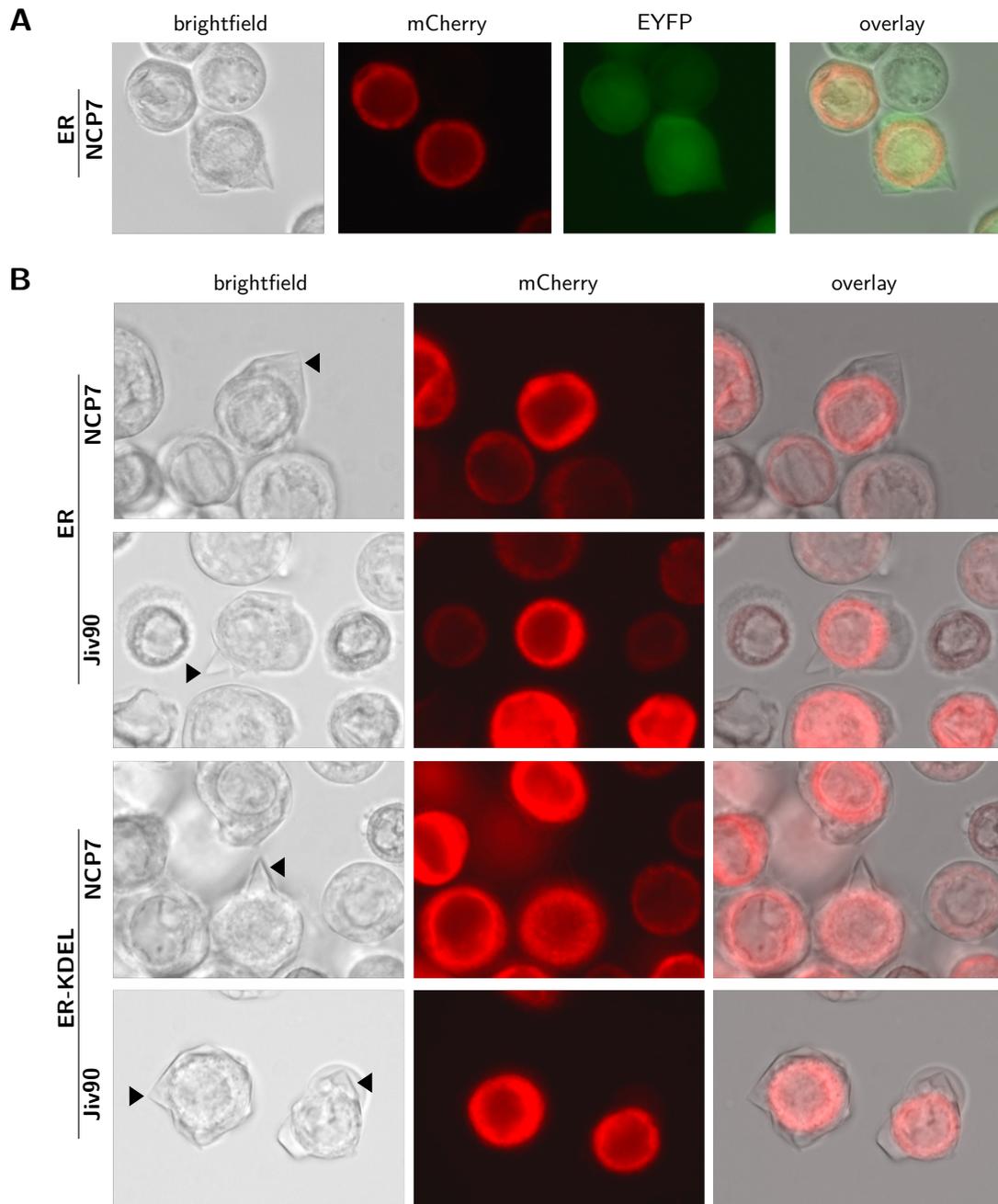


Figure 2.15.: Fluorescence microscopy of *in vivo* crystals in High Five™ insect cells

A: *In vivo* crystal grown in High Five™ cells from the ER NS2 construct in EMBacY bacmid. **B:** *In vivo* crystals grown in High Five™ insect cells from full-length NS2 constructs in bMON14272 bacmid. The images represent crystal examples from cells infected with ER and ER-KDEL NS2 constructs, both with and without Jiv90 insert. The type of localization signal and the presence of the Jiv90 insert is denoted beside the images. Images were taken with an inverted fluorescence microscope (AxioVert 200 M, Zeiss) and processed with the AxioVision SE64 software (Zeiss, Rel. 4.9). Black arrowheads: *In vivo* crystals.

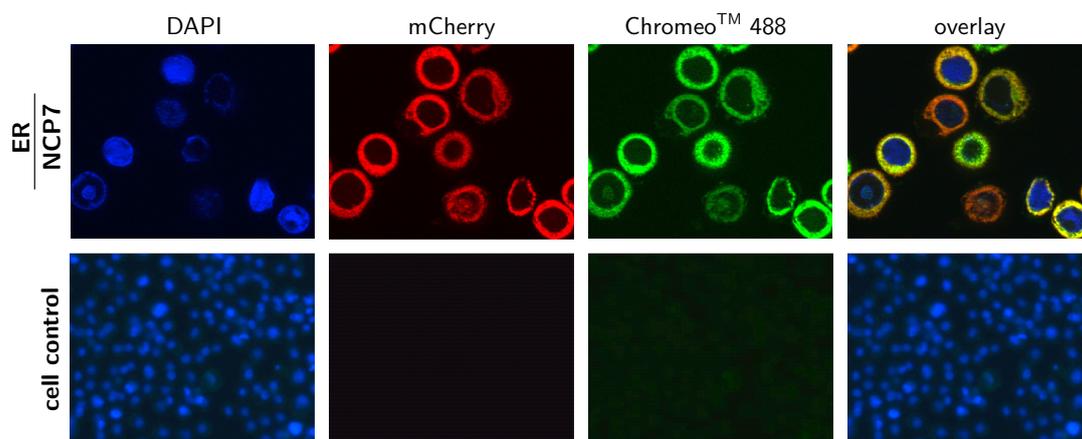


Figure 2.16.: Immunofluorescence microscopy of High FiveTM insect cells infected with the full-length NS2 ER construct

Uninfected High FiveTM insect cells (cell control; 10x magnification) and cells infected with the ER construct of full-length NCP7 NS2 (ER NCP7; 40x magnification) stained with DAPI and StrepMAB-Classic antibody coupled to the green fluorescent ChromeoTM 488. The cells were prepared on poly-D-lysine coated cover slips. The images were taken with an inverted fluorescence microscope (AxioVert 200 M, Zeiss) equipped with filters 38HE, 43HE and 49 (Zeiss).

resides in the ER. This is in accordance with the cellular localization of full-length NS2 in its natural bovine host cells, where it is integrated into the ER membrane (see section 1.3).

2.2.7. Immunofluorescence microscopy of *in vivo* crystals

Crystals grown in High FiveTM cells exhibit neither mCherry nor EYFP fluorescence, as depicted in figure 2.15. For constructs with Jiv90 insert, it could be argued that the crystals consist of already autoproteolytically cleaved NS2 without C-terminal mCherry. This would correlate with the presence of free mCherry in Jiv90-containing samples on the in-gel fluorescence gels from section 2.2.3 figure 2.13 **A** and **B** and the two *StreptII*-tagged bands from the western blot in figure 2.13 **C**, from which only the lower one exhibits mCherry fluorescence. However, the non-fluorescent crystals from cells infected with NS2 constructs without Jiv90 cannot unequivocally be explained by this line of argument. Nevertheless, although no free mCherry could be detected in samples derived from NS2 constructs without Jiv90, they also exhibit a band in the western blot, which exclusively contains a *StreptII*-tag, but no mCherry (see figure 2.13 **C**, lane 3).

Considering these contradictory results, the detection of the *StreptII*-tag in the *in vivo* crystals would help to clarify the identity of the crystallizing matter. Therefore, an immunostaining of the infected, crystal-containing High FiveTM cells with a fluorescence-labeled anti-*StreptII*-tag antibody (StrepMAB-Classic, ChromeoTM 488 conjugate; iba) was performed on cells either grown, fixed and stained on poly-D-lysine coated glass cover slips or grown in plastic

6-well plates and prepared for immunofluorescence in suspension (see section 7.14.2). The cells grown in the 6-well plates contained larger, more prominent crystals in contrast to those grown on cover slips, where the crystalline structures were often difficult to distinguish from the cellular boundaries. This might be due to the tighter attachment of the cellular surface to the poly-D-lysine coated cover slips than to the plastic surface of the 6-well plates, which might render the cellular envelope less flexible.

In figure 2.16 the results from uninfected High FiveTM cells (cell control) and cells infected with the ER construct of NCP7 NS2 grown on coated cover slips are shown representatively. The cell control was prepared and stained in the same way as the infected cells. For infected and uninfected control cells, the DAPI staining of the nucleus was successful. As expected, the cell control exhibits only DAPI (4',6-diamidino-2-phenylindole) fluorescence, verifying that the ChromeoTM 488-conjugated StrepMAB-Classic shows no unspecific binding to cellular structures other than the *StreptII*-tag. Infected cells still exhibit mCherry fluorescence from the NS2 fusion construct, which confirms that the chosen conditions for immunofluorescence staining keep the original cellular and protein structures mostly intact. The distribution of the mCherry fluorescence – and therefore the localization of the NS2 fusion protein – in a defined shell around the nucleus is in accordance with earlier results from fluorescence microscopy experiments on living cells (see section 2.2.7) and is therefore assumed authentic. The ChromeoTM 488 fluorescence completely co-localizes with the mCherry fluorescence, as can be seen in the image from the ChromeoTM 488 channel, and even more obvious in the overlay image, where the addition of mCherry and ChromeoTM 488 fluorescence yield an orange coloration of the corresponding areas. However, not all cells are stained equally well, so that cells with a lower ChromeoTM 488-to-mCherry fluorescence ratio appear a deeper shade of orange. Nevertheless, the co-localization confirms that *StreptII*-tag and mCherry are both located in the ER of the insect cell as part of the NS2 fusion protein.

Although the immunofluorescence staining of infected and uninfected High FiveTM cells was overall successful, in all prepared samples no crystal-like structures could be detected. Control preparations of crystal-containing insect cells under surveillance with an inverted light microscope revealed that the crystals are very unstable and already disappear during the fixation step. Therefore, no conclusions concerning the identity of the crystallizing matter can be drawn from these immunofluorescence experiments to date.

2.2.8. Transmission electron microscopy of *in vivo*-grown crystals

In vivo crystals from other protein constructs have already been successfully examined by transmission electron microscopy (TEM) [131], revealing the highly ordered features of a crystal on a high resolution scale and even some membranous structures encompassing the crystals. To gain further insights into the morphology of crystalline structures grown in High FiveTM insect cells infected with full-length NS2 fusion constructs, the crystal-containing insect cells infected with ER and ER-KDEL NS2 constructs (with and without Jiv90 insert) as well as uninfected control cells were prepared for TEM imaging as described in section 7.14.3. Subsequently, the samples were imaged with a JEM-1400 Plus TEM (JEOL; see section 7.14.3).

In figure 2.17, representative examples of the resulting TEM images and details thereof are shown. In contrast to uninfected control cells, the infected cells show the common signs

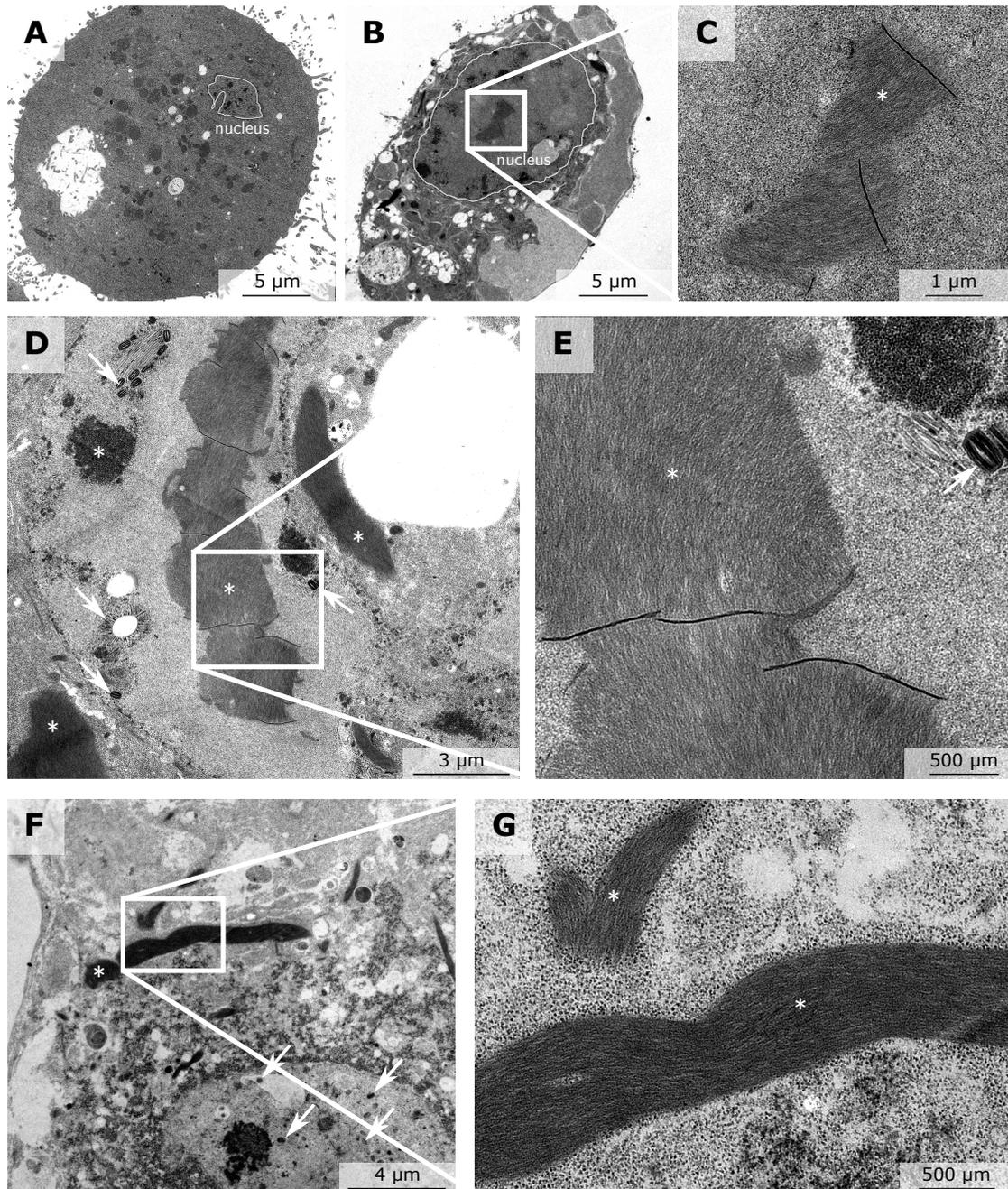


Figure 2.17.: Transmission electron microscopy (TEM) of High Five™ insect cells transfected with full-length NS2 constructs

A: Uninfected High Five™ cell. **B-E:** High Five™ cells infected with ER-*StrepII*-NS2-NS3_{7aa}-mCherry-KDEL. **F, G:** High Five™ cells infected with ER-*StrepII*-NS2_{Jiv90}-NS3_{7aa}-mCherry. The TEM images of the infected cells show fur-like structures (white asterisks), which cannot be found in uninfected cells (see **A**). White arrows: Cross-sections of baculoviruses. Images were obtained with a JEM-1400 Plus transmission electron microscope (JEOL).

of baculovirus infection: an increased size, an enlarged nucleus, a more granular appearance and even baculovirus cross-sections can be found (**A**, **B** and **C**). Unfortunately, no crystal-like structures could be detected in any of the analyzed samples. Nevertheless, prominent, sometimes quite large structures with a "furry" appearance (**B-G**: white asterisks) are exclusively visible in nucleus and cytoplasm of infected cells. Since no baculovirus-infected control cells were examined in parallel, these structures cannot unambiguously attributed to the NS2 fusion construct.

To determine if the sample preparation is responsible for the lack of detectable crystal structures – as was observed for the cells prepared for immunofluorescence (see section 2.16) –, the first fixation and washing steps were controlled under an inverted light microscope and the sample cells checked for crystals. Since the cells for TEM preparation were grown on poly-D-lysine coated cover slips, the crystal structures were already not very prominent and only protruded marginally from the cellular boundaries. Unfortunately, these structures merged with the cell and disappeared from view in the course of the fixation and washing steps.

2.2.9. Serial crystallography/X-ray diffraction experiments with *in vivo* crystals

Against the background of recent developments like serial crystallography with micro-focused synchrotron X-radiation or X-ray free-electron lasers (XFEL), the occurrence of *in vivo*-grown crystals is not only a fascinating phenomenon, but can finally be exploited for the structural characterization of heterologously expressed protein – without the need for the time- and labor-intensive, classical crystallization pathway.

At present, there are two options for X-ray diffraction experiments on *in vivo* crystals with a micro-focused X-ray beam: The crystals are either measured as suspension in an adequate, cryo-preserving buffer after gentle extraction from the host cells, or directly *in cellulo* by serial crystallography after brief cryo-preservation of the complete cells [76, 34].

As already described for the immunofluorescence and TEM experiments (see sections 2.16 and 2.17), the crystal structures are not very stable, when the cells are manipulated. Therefore, the *in cellulo* approach coupled with serial crystallography was chosen to check whether the crystals show any X-ray diffraction.

High FiveTM cells infected with ER and ER-KDEL NS2 constructs (with and without Jiv90 insert) were grown in 6-well plates for about 7 days to allow for *in vivo* crystallization. The cells were applied onto a mesh loop and scanned for diffraction with a synchrotron-generated X-ray beam (beamline P14 of EMBL Hamburg at the PETRA III storage ring, DESY, Hamburg, Germany) as described in section 7.14.4. Although crystals were still visible after sample preparation immediately prior to flash-freezing in the nitrogen gas cryo stream of the beamline, no detectable diffraction could be obtained from any of the samples.

2.3. Design and cloning of full-length NS2 constructs for expression in the cell-free BY-2 tobacco cell system

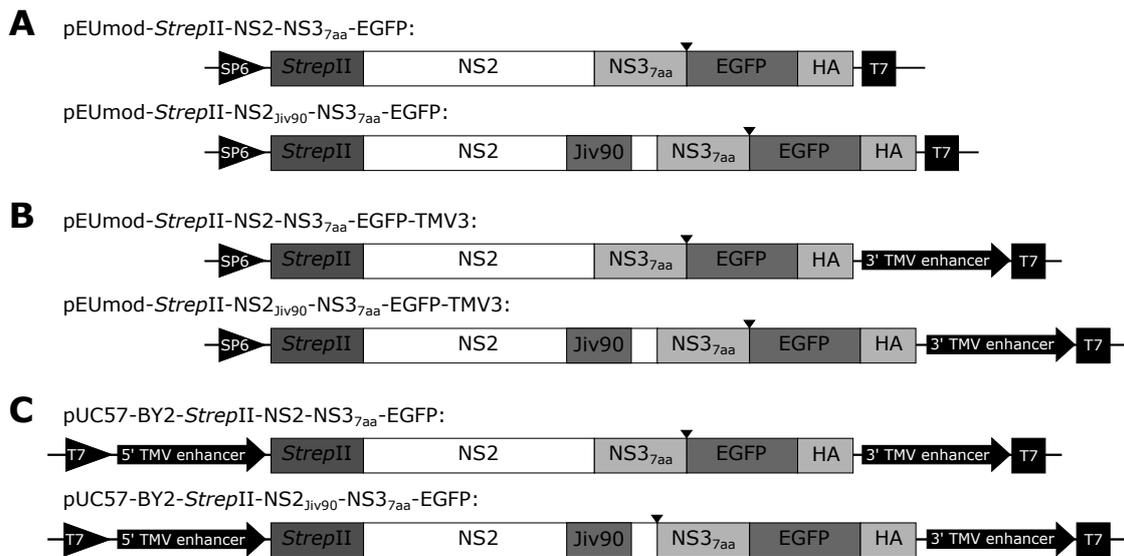


Figure 2.18.: Constructs for expression of full-length NS2 in BY-2 tobacco cell extract

All constructs contain either full-length NCP7 NS2 or NS2 with the NADL-like Jiv90 insert. NS2 is *N*-terminally fused to a *StrepII*-tag and *C*-terminally to the first 7 aa of NS3, followed by HA-tagged EGFP. **A:** The constructs are under the control of an SP6 promoter, combined with a T7 terminator. **B:** The constructs are under the control of an SP6 promoter and T7 terminator, combined with a 3' translational enhancer sequence from tobacco mosaic virus (TMV). **C:** T7 promoter and terminator are combined with 5' and 3' translational enhancer sequences from TMV. All constructs not to scale. Black forms symbolize non-coding elements. Black arrowheads indicate the position of tobacco etch virus (TEV) protease cleavage sites.

2.3. Design and cloning of full-length NS2 constructs for expression in the cell-free BY-2 tobacco cell system

For the expression of difficult target proteins, often of eukaryotic origin, cell-free *in vitro* translation systems based on cell extracts from pro- and eukaryotes have been developed alongside the conventional expression systems like *E. coli* or insect and mammalian cell lines. With cell-free systems, proteins can be produced, that would otherwise be toxic for conventional expression hosts. In addition, the conditions and the composition of the *in vitro* translation reaction can be manipulated, to supply for instance artificial amino acids, essential cofactors or nanodiscs for membrane proteins. Besides established eukaryotic systems like wheat germ extract (WGE), rabbit reticulocyte lysate (RLL) and insect cell extract (ICE), and more recently developed systems based on Chinese hamster ovary cell lines (CHO) [36] or *Saccharomyces cerevisiae* [105], a plant-based system applying BY-2 tobacco cell extract has been developed [42]. While WGE, RLL, ICE as well as extracts from CHO cells and *S. cerevisiae* are either laborious and expensive to produce or have only very low yields – in the worst case both –, the BY-2 extract is supposed to be quickly prepared from inexpensive shaking BY-2 cell cultures and deliver high protein yields [42].

2.3. Design and cloning of full-length NS2 constructs for expression in the cell-free BY-2 tobacco cell system

For the heterologous production of full-length NS2, a cost-effective, cell-free system that provides a eukaryotic expression and folding machinery, and the option to add protein- or polymer-enclosed lipid bilayer discs – termed nanodiscs – for the co-translational incorporation of the membrane protein, could be of great value. The BY-2 tobacco cell-based cell-free expression system is currently established in our laboratory under the supervision of Dr. Sebastian Krossa. To take advantage of this system with respect to the NS2 expression, 6 constructs were designed and cloned for future usage in cell-free BY-2 extracts. The constructs are depicted in figure 2.18. They all contain the same *N*-terminally *StrepII*-tagged full-length NS2 or NS2_{Jiv90} sequence with the 7 aa of NS3 necessary for autoproteolytic cleavage, that was already used in the pFastBac constructs. An analysis comparing this sequence with *Nicotiana tabacum* codon usage revealed no especially critical codons or sequence areas (see section A.1 figure A.5). *C*-terminally, NS2-NS3_{7aa} is fused via a tobacco etch virus (TEV) protease cleavage site to HA-tagged EGFP as a fluorescent reporter and target for immunological detection. The coding sequences of the expression constructs are flanked by either SP6 promoter and T7 terminator (pEUMod-derived constructs, **A** and **B**) or T7 promoter and terminator (pUC57-derived constructs, **C**) to enable translation initiation and termination. In addition, 4 constructs contain a 3' enhancer sequence derived from tobacco mosaic virus (TMV), which increases translation efficiency (**B** and **C**). Furthermore, the pUC57-derived constructs (**C**) feature another 5' translation enhancer sequence from TMV, also known as Ω .

The constructs described here were cloned as explained in detail in section 7.8.5. Their sequence identity and integrity was verified by sequencing (data not shown). Unfortunately, the constructs could not yet be tested for expression of full-length NS2, since the establishment of the BY-2 tobacco cell-based cell-free expression system is still an ongoing process in our laboratory.

3. Discussion

The BVDV NS2 autoprotease is a multi-pass membrane protein, which is localized in the ER membrane of the respective eukaryotic host cell. It represents the key regulatory element in the temporal modulation of BVDV infections and causes the unique viral biotype switch from non-cytopathic to cytopathic behavior upon alteration of its coding sequence in the highly variable viral ssRNA(+) genome (see section 1.3). It could therefore be a promising drug target. However, for the rational design of specific inhibitory substances, a profound understanding of the molecular mechanism of autoprotease activation, substrate recognition and cleavage reaction based on the three-dimensional structure of NS2 is an essential prerequisite. To date, no structural information has been reported on BVDV NS2. In addition, due to the lack of structurally characterized, homologous proteins, homology modeling of the NS2 autoprotease based on the amino acid sequence does not provide any valuable information. The *ab initio* generation of structural data on NS2 is therefore inevitable.

In most cases, high resolution structure determination still relies on the classical, random based crystallization of heterologously expressed and subsequently purified protein. In this thesis, non-cytopathic (ncp) and cytopathic (cp) variants of either full-length or *N*-terminally truncated BVDV NS2 autoprotease were expressed in a range of *E. coli* strains and two insect cell lines to determine the best conditions for the production of protein suitable for future biophysical characterization and structure determination experiments.

3.1. *E. coli* expression of *N*-terminally truncated NS2 fusion constructs

With its 7 predicted transmembrane α -helices in the *N*-terminal portion, the full-length NS2 autoprotease is a challenging target for heterologous expression in *E. coli* (see section 1.3 for further details on NS2). In addition to the toxic effects that heterologously expressed membrane proteins themselves can have on the prokaryotic expression host, the latter also lacks organellar membranes and can only allocate its inner and outer membrane for the integration of membrane proteins. The heterologous expression of membrane proteins therefore often overwhelms *E. coli*'s protein folding and membrane integration capacity, resulting in the deposition of un- or partly folded protein into inclusion bodies. However, as experimental evidence confirmed, inclusion bodies can also consist of large portions of natively folded, co-aggregated protein [296, 273, 223]. While membrane proteins containing β -barrels can often be recovered successfully from inclusion bodies for subsequent characterization, transmembrane proteins featuring several α -helices – like NS2 autoprotease – need to be correctly integrated into a lipid bilayer upon expression to yield functional protein [15].

3.1.1. Codon-optimized, N-terminally truncated NS2 fusion constructs yield inclusion bodies upon expression in *E. coli*

In the case of the NS2 autoprotease, all putative catalytically relevant residues are located in its cytosolic C-terminal portion, which does not contain any transmembrane helices (see section 1.3). Based on this information, several N-terminally truncated constructs omitting the first 5, 6 or all 7 critical transmembrane helices were designed for the expression in *E. coli* in the course of this thesis.

The first expression constructs featured the authentic coding sequence for three different N-terminal truncated variants of non-cytopathic (ncp) BVDV strain NCP7. They were all N-terminally fused to His₆-tagged SUMO (small ubiquitin-like modifier) as affinity tag and solubility enhancer [44, 163]. A second set of constructs represented the situation in cytopathic (cp) BVDV biotypes: They additionally contained the coding sequence for the so-called Jiv90 insert, which was originally found in the cp strain NADL and is derived from the bovine DnaJ C14 chaperone (see figure 2.1 for further construct details).

After introduction into the standard expression strain BL21 (DE3) and its derivative BL21 (DE3)pLysS, a specialized strain with reduced basal expression for potentially toxic proteins, neither of these 6 constructs yielded any detectable expression product upon IPTG induction. In contrast, strains BL21 (DE3)RIL, RosettaTM 2(DE3) and RosettaTM 2(DE3)pLysS, which all provide additional rare tRNAs for less frequently used codons in *E. coli*, were able to produce significant amounts of both ncp and cp variants of the shortest NS2 truncation construct (pET-His₆-SUMO-[1393]NS2* and -[1393]NS2_{Jiv90}*; see section 2.1.2). In conjunction with an *in silico* codon usage analysis, which revealed several critical codons (see section A.1 figure A.2), these findings supported the hypothesis that the initially observed complete lack of overexpression in BL21 (DE3) and BL21 (DE3)pLysS was due to an unfavorable codon usage. However, even in expression strains supplying rare tRNAs, neither of the two longer NS2 truncation constructs in their ncp or cp variant could be expressed in detectable amounts.

In addition, first purification trials with protein expressed in BL21 (DE3)RIL and RosettaTM 2(DE3) indicated that the NS2 fusion protein was solely localized in inclusion bodies. As already mentioned above, this behavior is observed regularly with *E. coli*, when the prokaryotic cell is not able to properly handle the fast overexpression and correct folding of high amounts of heterologous protein under the particular conditions. This is especially true for eukaryotic proteins – or viral proteins originally produced in eukaryotic host cells: They often have a more complex fold and might therefore even require the addition of posttranslational modifications like specific glycosylations, which *E. coli* cannot provide. To identify potentially critical glycosylation sites, an *in silico* analyses of the NCP7 NS2 amino acid sequence for characteristic N- or O-glycosylation motifs with the artificial neural network-based prediction servers NetNGlyc (version 1.0, prediction of N-glycosylation sites in human proteins [96]), NetOGlyc (version 4.0, prediction of N-acetylgalactosamine (GalNAc) O-glycosylation sites in mammalian proteins [264]) and YinOYang (version 1.2, prediction of N-acetylglucosamine (GlcNAc) O-glycosylation sites in eukaryotic proteins [95]) of the Department of Bio and Health Informatics at the Technical University of Denmark was performed at medium stringency level. The servers predicted 3 potential GlcNAc O-glycosylation sites – which are all located in putative transmembrane helices and can therefore be disre-

garded –, and 3 *N*-glycosylation sites. Since the *N*-glycosylation reaction is performed in the ER lumen, only one of these positions was qualified as actual *N*-glycosylation site. The respective asparagine 1357 is located in the ER residing loop directly preceding the seventh putative transmembrane helix (see figure 1.3). However, one has to keep in mind that the *N*-glycosylation prediction server was trained on human proteins and that no experimental evidence for the glycosylation of this site in the bovine host cell has been reported to date. Therefore, it seems unlikely that a missing glycosylation is the reason for inefficient folding of the NS2 fusion constructs.

The variation of parameters influencing expression speed and produced protein amount – like the expression temperature or the induction intensity – giving the bacterial cell more time for the protein folding, did not improve the solubility of the expressible NS2 truncation variants in BL21 (DE3)RIL and RosettaTM 2(DE3). Therefore, several specialized *E. coli* strains, which provide additional skill sets for the expression and folding of challenging target proteins, were tried. Most of the tested strains provided accessory chaperones, like BL21 (DE3) and OrigamiTM 2(DE3) derivatives, which co-express different combinations of the chaperone systems DnaK/DnaJ/GrpE, GroEL/GroES or trigger factor, or ArcticExpressTM (DE3), which supplies the cold-adapted chaperones Cpn10 and Cpn60 (GroEL/GroES homologs from *Oleispira antarctica*) for a significantly enhanced folding efficiency at very low expression temperatures. OrigamiTM 2(DE3) strains also support disulfide bond formation in the cytoplasm of *E. coli* due to mutations in thioredoxin and glutathione reductase genes, although no experimental evidence has been reported for disulfide bond formation in BVDV NS2 to date.

To enable the expression in these strains independent of rare tRNA supplementation, the shortest NS2 truncation constructs were codon-optimized with respect to *E. coli*'s codon usage. In addition, 8 new NS2 truncation constructs with codon-optimized coding sequences were created, featuring either an *N*-terminal or a *C*-terminal fusion to a His₆-tagged GFP as solubility enhancer and folding state and localization reporter. These constructs contained one of two *N*-terminal truncation variants of NCP7 NS2 – the larger variant starting directly after the sixth putative transmembrane helix and including the seventh, and the shorter variant only encompassing the complete cytosolic part directly following the seventh putative transmembrane helix. Both variants were constructed with and without the Jiv90 insert (see figure 2.3 for further construct details).

As expected, expression trials with these codon-optimized new GFP fusion constructs and the two successful SUMO fusion constructs in the different above-mentioned *E. coli* strains confirmed that the codon-optimization rendered the supplementation of rare tRNAs unnecessary for efficient expression.

In summary, both His₆-SUMO constructs could be overexpressed in BL21 (DE3), OrigamiTM 2(DE3) and their chaperone-supplying derivatives as well as in ArcticExpressTM (DE3) (see section 2.1.3). However, strains providing accessory chaperones yielded in general less fusion protein. This is probably due to a competitive situation at the ribosomes between the overexpression of the chaperones on one side and the fusion protein on the other, which is also reflected by the prominent chaperone overexpression bands that most lysates of the corresponding strains exhibit on SDS-PAGE. The GFP fusion constructs did only yield detectable overexpression for the shorter NS2 truncation variant. While ArcticExpressTM (DE3)

supported the expression of both N- and C-terminal GFP fusion proteins, Origami™ 2(DE3) only exhibited overexpression of the C-terminal GFP fusion protein. In contrast, for the other Origami™ 2(DE3) derivatives and all BL21 (DE3) strains, overexpression could be detected neither for the shorter nor for the longer NS2 truncation variants. The complete lack of overexpression for the latter variants correlates with the results for the two longer N-terminal NS2 truncation variants of the SUMO fusion proteins, which also exhibited no noticeable overexpression.

3.1.2. N-terminally truncated NS2 fusion proteins from inclusion bodies can be affinity-purified after solubilization with N-lauroylsarcosine surfactant

With fusion protein from successful expression experiments, the parameters for cell lysis and protein solubilization were optimized. Since the best conditions for each individual fusion construct cannot be predicted, but have to be determined empirically, a range of buffer substances and pH values, salt concentrations, detergents and additives as well as different lysis methods were evaluated (see section 2.1.4). Unfortunately, all overexpressed NS2 truncation variants with either SUMO or GFP fusion from the different expression strains exhibited similar behavior: Despite all attempts to enable proper folding of the fusion proteins in *E. coli* by slowing down the expression with low-level induction, reduced expression temperatures and the supplementation of additional folding chaperones, the NS2 fusion proteins still remained in the pellet fractions in the absence of a detergent – indicating their localization in inclusion bodies. From all tested surfactants, the anionic N-lauroylsarcosine (NLS) was the only one able to transfer SUMO and GFP fusion proteins from most expression experiments into solution. The initially determined necessary NLS concentration was about 0.3 % (w/v; 10.2 mM), which is below the critical micelle concentration of 14.6 mM (Merck/Sigma-Aldrich®) and correlates with effective concentrations reported for the successful recovery of active protein from inclusion bodies in literature [223, 81].

Fortunately, the addition of 0.3 % NLS to the lysis and purification buffers did not interfere with the subsequent His₆-tag-based affinity chromatography of the NS2 fusion proteins (see section 2.1.5). While the gradual, but finally complete removal of NLS from the purification buffers during the protein's stay on the affinity column matrix – to allow for on-column-refolding of putatively not correctly folded protein – resulted in its aggregation on the column, the NLS concentration could at least be reduced to about 0.08 % without impairing solubility.

The expressible SUMO and GFP fusion proteins behaved quite similar with respect to their poor solubility in the absence of NLS. Nevertheless, although the major fraction of overexpressed fusion protein was obviously deposited in inclusion bodies, the presence of minor amounts of correctly folded, soluble protein could not be ruled out with certainty from SDS-PAGE results alone. Affinity chromatography results from the soluble fraction of GFP fusion proteins expressed in ArcticExpress™ (DE3) – which provide best conditions for slow expression and efficient protein folding at low temperatures – indeed confirmed that fusion proteins were also present in the lysis supernatants. However, as further western blot analyses revealed, these small amounts of GFP fusion proteins were accompanied by large amounts of co-purified Cpn60 chaperone. The purified GFP fusion proteins are therefore very likely not soluble on their own, but only bound to the chaperones, which mediate their solubility.

This is also supported by the fact that the addition of Mg^{2+} and ATP did not release soluble fusion proteins from the GroEL-like chaperones [142].

All previous results summarized, it was not possible to generate inherently soluble SUMO or GFP fusion proteins of N-terminally truncated NS2 autoprotease despite all efforts to optimize expression and solubilization conditions. Only the application of at least 0.08 % of the anionic surfactant NLS in lysis and purification buffers enabled the affinity purification of the fusion proteins. However, under these conditions the integrity and authentic fold of the purified C-terminal NS2 portion is questionable. Unfortunately, for obvious reasons the native tertiary structure is the one crucial prerequisite for all subsequent biophysical characterization and structure determination experiments.

3.1.3. Evaluation of SUMO protease cleavage as means to assess NS2 folding state is impaired by presence of NLS surfactant

Since NS2 is a cysteine autoprotease, the detection of its autoproteolytic activity with a reliable cleavage assay would be the best way of confirming its native tertiary structure. However, several requirements would have to be met to enable autoproteolytic cleavage: 1) The authentic cleavage site at the C-terminus of NS2 would have to be followed by at least 7 amino acids of NS3 to allow cleavage site recognition (see section 1.3, [140]). 2) The constructs would need a sufficiently large, easily detectable C-terminal protein portion to cleave off. 3) While Jiv90-containing NS2 truncation variants are supposed to represent a permanently active NS2 autoprotease, authentic ncp NS2 autoprotease would have to be supplied with stoichiometric amounts of its cellular cofactor Jiv to exhibit cleavage activity (see section 1.3.2). 4) N-terminal truncations would have to be less than 73 aa to ensure cleavage activity at all (see section 1.3, [140]). While the first three points could be addressed by further construct optimization, the fourth poses the actual problem: since all current NS2 truncation constructs feature an N-terminal truncation of at least 149 amino acids, it can be assumed that neither of these constructs is inherently capable of autoproteolytic cleavage.

To gain at least some insights into the folding state of the purified fusion proteins, the fusion partners SUMO and GFP could be consulted. GFP fused to the C-terminus of a protein is supposed to report its folding state. A misfolded fusion partner would impair proper folding of GFP and thereby prevent its fluorescence. On this basis, GFP fluorescences could be correlated to fusion protein folding and solubility, which subsequently led to the development of further GFP-based folding reporter systems, such as circularly permuted split-GFP variants [285, 45]. In this thesis, four constructs with GFP fused to the C-terminus of N-terminal NS2 truncation variants were generated. Of these constructs, the ones encompassing the shorter NS2 truncation variant with and without Jiv90 insert could be expressed in *E. coli*, exhibiting green fluorescence from the initial expression culture to the final purification product (data not shown). Unfortunately, the presence of GFP fluorescence alone does not guarantee the correctness of the tertiary structure of the corresponding NS2 truncation variant, since – based on literature – fluorescent GFP could also be detected in inclusion bodies and was even identified as aggregation mediator for the deposition of otherwise soluble protein into inclusion bodies [108].

For constructs containing an *N*-terminal SUMO fusion as solubility mediator, another strategy for the assessment of the folding status could be applied. Correctly folded SUMO is the substrate for the highly specific SUMO protease. Successful cleavage of the SUMO fusion constructs by SUMO protease after the expression and purification procedure would thereby confirm the correct tertiary structure of SUMO in the fusion proteins, indicating that the chosen conditions do not impair the integrity of a stable protein. However, cleavage experiments with SUMO fusions of NS2 truncation variants and control substrates revealed that the NLS concentration necessary for fusion protein solubility strongly interferes with protease activity (see section 2.1.6). While at least 0.08 % NLS are required for NS2 fusion protein solubility, SUMO protease only tolerates up to 0.02 % NLS without significant impairment of cleavage efficiency. Subsequent dilution to tolerated NLS concentrations did not restore cleavage in samples treated with NLS concentrations above 0.02 % (data not shown). However, based on these results alone it is not possible to determine whether NLS is involved in substrate or protease denaturation or only prevents protease-substrate interaction by tightly binding to relevant interacting surfaces. Regardless of the underlying mechanism, it is undeniable that SUMO cleavage cannot serve as indicator for the integrity of the tertiary structure of SUMO fusion constructs under these circumstances.

In conclusion, during this thesis codon-optimized *N*-terminal truncated NS2 constructs encompassing solely the cytosolic *C*-terminal portion of NS2 could be successfully overexpressed in various *E. coli* strains. Unfortunately, despite all efforts to support proper protein folding during expression, the fusion proteins were always deposited in inclusion bodies. After the screening of a variety of buffer conditions and lysis methods, buffers containing at least 0.08 % NLS finally enabled fusion protein solubilization and subsequent His₆-tag-based affinity chromatography. Unfortunately, based on literature data it has to be assumed that the expressible NS2 truncation variants are inherently inactive [140], rendering a direct autoprotease activity assay for the assessment of NS2 folding integrity pointless. While the detected GFP fluorescence could be interpreted as indicative for proper folding of the adjacent NS2 truncation variant, it could not be assigned to the intact fusion protein with certainty. The attempt to ascertain the folding status of SUMO – and thereby indirectly of the NS2 truncation variant – by cleavage with the folding-sensitive SUMO protease failed due to the low tolerance of the cleavage reaction to NLS concentrations above 0.02 %. Nevertheless, further biophysical characterization techniques like circular dichroism (CD) or nuclear magnetic resonance (NMR) spectroscopy could still be tried to gain more data on the tertiary structure of the NS2 truncation variants. However, since it is already clear that the purified truncation variants do not represent the authentic, active autoprotease, the scientific value of these results for subsequent characterization and structure determination experiments would be very limited.

3.2. Expression of full-length NS2 fusion proteins in insect cells

The previous experiments clearly showed that *E. coli* is a suboptimal choice as expression host for functional NS2 autoprotease. Only the short, inherently inactive cytosolic *C*-terminal portion of NS2 could be expressed, but not confirmed as suitable for further structural characterization experiments. To ensure full autoprotease functionality, the expression of full-length

NS2 is an important prerequisite. Since *E. coli* does not support the expression of NS2 variants that encompass one or more of its 7 transmembrane helices, another expression host had to be chosen. Although mammalian expression systems would naturally provide optimal expression conditions, they are cost-intensive, delicate and require special laboratory equipment. In addition, protein yield is often low. Therefore, the baculovirus-mediated expression in insect cell lines presented a reasonable compromise between the need for a sophisticated eukaryotic protein expression machinery and the monetary and technical investment. Insect cell expression systems provide more complex posttranslational modifications and a better folding machinery for proteins of mammalian host cell origin than simpler eukaryotic expression hosts like yeast. Yet, they are still easily sustained in adherent or suspension culture. Therefore, two insect cell lines, Sf21 and High FiveTM, were chosen for the expression of full-length NS2.

3.2.1. Design of functional full-length ncp and cp NS2 fusion constructs and choice of bacmid shuttle vectors

In this thesis, 8 constructs encompassing either full-length ncp NS2 autoprotease or its cp variant with the NADL-like Jiv90 insertion were designed and tested for expression (see figure 2.10 for further construct details). In these constructs, *N*-terminally *StreptII*-tagged ncp or cp NS2 was *C*-terminally followed by the first 7 amino acids of NS3 – enabling autoprotease cleave site recognition (see section 1.3) – and the red fluorescing protein mCherry as expression and localization reporter. For both ncp and cp NS2 variants, 4 different constructs were generated. Three of these constructs contained additional signal sequences to direct the expression product to the ER (ER; ER-KDEL for ER retention) and peroxisomes (SKL), while the fourth would express its fusion construct into the cytosol. All constructs were successfully transposed into a bacmid shuttle vector (see section 2.2.1 figure 2.11).

In this thesis, two different bacmids were tested in parallel: bMON14272 from the commercial Bac-to-Bac[®] baculovirus system, and the advanced EMBAcY (EMBL Grenoble), which features additional protease and chitinase deficiencies, extended cell viability and reduced proteolysis compared to bMON14272, as well as separately expressed EYFP under the same polyhedrin promoter as the NS2 fusion construct for the easy estimation of target protein expression rates. While the EYFP reporter is in general a useful feature, this strategy ensures that subsequent experiments can also be performed without EYFP presence, if necessary.

3.2.2. Activable, full-length ncp and cp NS2 fusion protein is successfully expressed in Sf21 and High FiveTM insect cells

After the establishment of a reliable insect cell culture system in our laboratory, first small scale expression experiments were performed in 6 well plates with Sf21 and High FiveTM cells transfected with the different bacmids or infected with their higher titer virus stocks. Upon visual inspection, the adherent cell layer already exhibited a promising reddish tint in addition to the common signs of baculovirus infection. The presence of mCherry – as well as EYFP for the corresponding constructs – could subsequently be confirmed by fluorescence excitation and detection on a Blue/Green LED Transilluminator (see section 2.2.2 figure 2.12). Interestingly,

while EYFP fluorescence is prominent in for all NS2 constructs, mCherry fluorescence is more apparent for all insect cell cultures expressing ER and ER-KDEL constructs and only faintly visible for cytosolic and SKL constructs. However, this preference seems reasonable, considering the authentic cellular localization of NS2 in the ER membrane.

Cell pellet and medium supernatant samples from expression cultures of all constructs were subsequently analyzed on SDS-PAGE under only mildly denaturing conditions to preserve mCherry (and EYFP) fluorescence (see section 2.2.3 figure 2.13). Indeed, fluorescence images of these SDS-PAGEs confirmed the presence of red fluorescing, full-length fusion protein bands for all 8 constructs at the expected molecular weight (taking into account the incomplete denaturation status). In addition, a band corresponding to separate mCherry could only be detected for cell pellet and medium supernatant samples from cells expressing Jiv90-containing NS2 fusion proteins. These promising results strongly suggest that the Jiv90 insert in these proteins was able to activate NS2 autoproteolytic activity - proving its native, functional tertiary structure. However, in contrast to the bands of the full-length fusion proteins, these bands were rather faint. This might indicate that only a minor fraction of the fusion proteins is cleaved.

If the NS2 fusion protein is cleaved, the separate *StreptII*-tagged full-length NS2 should also be detectable besides mCherry. A western blot detecting the *StreptII*-tag was therefore performed on SDS-PAGEs obtained under mildly denaturing conditions and overlaid with the corresponding fluorescence image (see figure 2.13). As expected, samples from Jiv90-containing fusion protein exhibit two bands in western blot. While the band with the apparently lower molecular weight shows additional mCherry fluorescence and is assumed to be the complete fusion protein, the other is not fluorescent and must therefore represent the separate NS2 after cleavage. The contradicting running behavior is probably due to the stable and more compact fold of mCherry in the complete fusion protein, simplifying its migration through the pores of the PA gel.

3.2.3. Full-length ncp NS2 autoprotease might be activated for cleavage by insect cell-derived Jiv homolog

Unexpectedly, two analogous bands are also visible for the samples from the ncp NS2 fusion variant, indicating that these samples likewise contain *StreptII*-tagged NS2 without mCherry. Curiously, separate mCherry could not be detected in the fluorescence images of these samples. There are two possible explanations: either separate NS2 can also occur as fragment of an aborted translation or a degradation process (and in addition to the actual cleavage product in cp NS2 variants), or the NS2 autoprotease is somehow activated by an unknown insect cell-derived factor. The first option might also explain the absence of separate mCherry, if translation is aborted prior to mCherry production. However, the cells are clearly able to produce the complete fusion proteins, and codon usage analyses with the NS2 coding sequence (see appendix figure A.3) and the complete fusion construct (data not shown) did not reveal any critical sections. It therefore seems unlikely that the specific separate NS2 fragment arises from abortive translation.

In the second case, the insect cells might provide a homolog of Jiv (DnaJ C14), which is able to induce cleavage with low efficiency. This is not unlikely, since the Jiv90 region

can be found in many eukaryotes. Database research with Jiv or Jiv90 sequences did not reveal any relevant homologous proteins for *Spodoptera frugiperda* (Sf21) or *Trichoplusia ni* (High Five™). In fact, only a single DnaJ protein could be found for *S. frugiperda* (GenBank Acc. No.: AHA86298.1). However, for the closely related tobacco cutworm *Spodoptera litura*, a great number of DnaJ homologs are annotated. One of them, auto-annotated dnaj-5 (NCBI Ref. Seq.: XP_022816320.1), showed high sequence similarity around the DnaJ domain and Jiv90 (see section A.2 figure A.6). In addition, for other related *Lepidoptera* species – namely the silk moth *Bombyx mori* and the butterflies *Papilio machaon* and *Papilio xuthus* – Jiv90-containing proteins are listed. An alignment of these four proteins from *Lepidoptera* with four Jiv90-containing proteins from *Artiodactyla* – bovine (*Bos taurus*), sheep (*Ovis aries*), pig (*Sus scrofa*) and Bactrian camel (*Camelus ferus*) – also confirmed the high sequence similarity around the Jiv90 (and DnaJ) region (see section A.2 figure A.7). As outlined in section 1.3.2, an alanine scan identified W571 in the Jiv90 sequence as the only crucial residue for NS2 cleavage induction [140]. This residue is highly conserved in all eukaryotic Jiv90-containing proteins, which is also reflected by the alignment in figure A.7. Although no direct statements concerning the presence of a Jiv homolog in Sf21 or High Five™ can be made, based on these data from related species it seems highly likely that the insect expression cell lines also possess a Jiv homolog.

Based on this theory, the low cleavage efficiency of ncp NS2 fusion protein is easily explained by the inherent nature of NS2 cleavage activation: As outlined in detail in section 1.3.2, stoichiometric amounts of cellular Jiv are necessary for optimal NS2 cleavage efficiency. After NS2 cleavage, Jiv stays bound to NS2, so that each cleavage reaction is induced by its own Jiv. Since the AcMNPV polyhedrin promoter controlling the expression of the NS2 fusion protein is very strong, high amounts of heterologous protein are produced in the infected insect cells. In contrast, the cellular Jiv-homolog pool in the insect cells is limited – a situation that resembles that of BVDV-infected bovine cells at a late stage of infection. Thus, the cellular Jiv pool is depleted – resulting in minor amounts of cleaved NS2 – and the major fraction of NS2 fusion protein remains uncleaved.

The absence of detectable mCherry is therefore most likely due to the very low cleavage efficiency. In this case, since separate mCherry was already hardly detectable in samples from Jiv90-containing fusion protein with the applied makeshift imaging system, a more sensitive setup with a professional fluorescence imaging system might also reveal fluorescence from separate mCherry in samples from ncp NS2 fusion proteins.

To clarify the situation in the future, a mass spectrometric analysis of the NS2 fragments could be performed to confirm that the fragments contain the authentic cleaved C-terminus. The identification of a Jiv homolog via a PCR with degenerated primers on cDNA from the insect cell lines could be an additional option [144].

3.2.4. Insect cells expressing ER-targeted NS2 fusion proteins exhibit *in vivo* crystal growth

Interestingly, in contrast to Sf21 cells, High Five™ cells expressing ER or ER-KDEL (and in fewer cases SKL) constructs did not only exhibit mCherry fluorescence, indicating the presence of full-length fusion proteins, but additionally often contained prominent crystal-like

structures (see figure 2.14). Spontaneous *in vivo* crystallization of proteins in cells has been reported for a range of cell types and for different reasons: from protein crystals as a means of storage in seeds [58], secretory cells like pancreatic β -cells or eosinophiles [68, 152], protists [281] or bacteria [2], over peroxisomal catalytic crystals in yeast [284, 113], to the protective encapsulation of virus particles observed for several insect virus genera (cytoplasmic and nuclear polyhedrosis viruses, granulosis viruses, entomopoxviruses [8, 115]). The baculovirus-mediated insect cell expression system employed in this thesis actually exploits the strength of the polyhedrin promoter, which naturally drives the expression of viral polyhedrin. Polyhedrin is expressed during later stages of baculovirus infection and auto-assembles into a highly stable, crystalline coating around the viral particle to protect it from environmental hazards. The crystallization of protein heterologously expressed in the baculovirus expression system therefore is not a rare event [79, 131].

With the recent development of highly brilliant, micro-focus synchrotron-generated X-ray beams and the computing power to enable the processing of diffraction images obtained not only from one single crystal, but from a series of crystals in arbitrary orientation (serial crystallography, [85]), as well as the implementation of X-ray free-electron lasers (XFELs), which provide the unprecedented opportunity to collect diffraction data from micro-crystals in suspension, it is now possible to use these *in vivo* grown crystals for three-dimensional structure determination on a molecular level [229, 76].

Upon infection of freshly seeded High FiveTM cells with P3 virus generation containing ER or ER-KDEL NS2 fusion constructs, the crystalline structures reliably appeared after 5 to 7 days, in some cases already at day 3. The crystals are colorless and translucent, and in most cases they have a thin, plate-like, triangular or even rectangular morphology. Their size ranges from hardly detectable to about the average diameter of an infected High FiveTM cell. The visible crystals stick out of the cell, with traceable parts inside the cell in a few cases. Since the thin plates are hardly visible when localized inside the cells, it cannot be ruled out that cells without prominent protruding crystals actually also have smaller crystals inside. Frequently, several crystals grow from one cell, often in several layers on top of each other. The crystal shapes and sizes do not correlate with the crystal origin, all shapes and sizes could be detected in all crystal-yielding conditions.

In cells infected with baculovirus carrying SKL constructs, crystals were much less frequent, and for cytosolic constructs no crystals could be observed. While the construct itself obviously influences the frequency of crystal occurrence, the employed bacmid did not: crystals could be detected for cells infected with constructs in both bacmids (bMON14272 and EMBacY) in comparable amounts and shapes. In contrast, individual mCherry expressed as a control from cytosolic, ER, ER-KDEL and SKL constructs from both bacmids and in Sf21 as well as High FiveTM cells did not exhibit any sign of *in vivo* crystal growth. These results support the assumption that the crystals actually contain the NS2 fusion protein and are not an artifact of baculovirus-mediated expression.

Provided that the crystals are of NS2 fusion protein origin, the construct-dependent differences in crystal yield could be due to the varying amounts of produced fusion protein. As mentioned above, Sf21 and High FiveTM cells expressing ER and ER-KDEL constructs exhibited a considerably higher mCherry fluorescence than those expressing cytosolic or SKL

constructs. However, for other proteins that exhibit spontaneous *in vivo* crystallization in insect cells, a preference for a specific compartment for crystallization has been reported: While *Trypanosoma brucei* cathepsin B (*TbCatB*) exclusively crystallizes in the ER, crystals of GFP- μ NS – a GFP-tagged portion of the avian reovirus (ARV) nonstructural protein μ NS – grow in the cytosol, and firefly luciferase crystals assemble in peroxisomes [246, 130]. Therefore, it is more likely that the ER localization just provides the best conditions for NS2 fusion protein crystal growth. Considering the 7 transmembrane helices of NS2, it seems reasonable that the expression into membranous compartments like the ER (or peroxisomes) is advantageous for NS2 crystal formation.

3.2.5. Full-length ncp and cp NS2 fusion protein is localized in the insect cell ER, but cannot be detected in *in vivo* crystals

To determine the cellular localization of the NS2 fusion protein in crystal-containing insect cells, living High FiveTM cells expressing ER or ER-KDEL constructs from both bacmids were analyzed under an inverted fluorescence microscope (see section 2.2.6 figure 2.15). Interestingly, for all constructs and bacmids, the mCherry fluorescence is localized in a well-defined shell-like area around the nucleus, which is assumed to be the ER. In contrast, in cells which additionally express EYFP without any targeting sequence, the EYFP fluorescence is distributed uniformly throughout the cell, representing cytosolic localization. However, while the mCherry fluorescence confirms that the NS2 fusion protein is actually localized in the ER membrane, the crystalline structures do not exhibit mCherry fluorescence at all. In cells expressing additional EYFP, they exhibit the same EYFP fluorescence as the rest of the cell. This result can arise from two possible scenarios: either the crystallizing protein is not the NS2 fusion protein, or the crystals exclusively consist of already cleaved NS2.

The first option seems rather unlikely, considering the fact that from two different bacmids only ER and ER-KDEL NS2 fusion constructs show significant crystal growth, but neither the cytosolic or the SKL NS2 fusion constructs nor all mCherry control constructs. In other words: The only feature all crystal-producing cells have in common is an ER-targeted NS2 fusion protein, which makes it the most likely candidate for the crystallizing protein.

Based on this line of argument, it appears more plausible to assume that the crystals consist of cleaved NS2. Since crystals were observed for both cp and ncp NS2 fusion protein variants, this assumption would agree with the earlier finding that *StrepII*-tagged NS2 without mCherry is also present in cells expressing the ncp variant. Interestingly, while the autoproteolytic cleavage of the NS2 fusion protein seemed to be of low efficiency, judged by SDS-PAGE and western blot results (see above), the occurrence of crystals is rather frequent. Nevertheless, since samples for SDS-PAGE only represent the average of a cell population, it is possible that some cells contain high amounts of activated, crystallizing NS2, while the majority of the cell population contains mostly uncleaved NS2. However, this argumentation might disagree with the earlier assumption that all insect cells provide the same, limited Jiv-homolog pool. As a possible solution for this contradiction, it is also conceivable that some time-dependent protein maturing has to take place prior to NS2 activation and cleavage, so that the SDS-PAGE samples capture the average of different maturation stages – which might arise from asynchronous infection of the insect cells.

3.2.6. Delicate *in vivo* crystals disintegrate upon fixation for immunofluorescence and TEM

Despite all these indirect evidence, the direct detection of *StreptII*-tagged NS2 in the *in vivo* crystals would clarify most of the above-mentioned issues. The *StreptII*-tag is ideally suited to serve as epitope for immunofluorescence experiments. Therefore, crystal-containing insect cells, which expressed ER and ER-KDEL NS2 fusion constructs from bMON14272 bacmid (to exclude interference from EYFP fluorescence), were fixated, permeabilized and subsequently stained with a green fluorescing anti-*StreptII*-tag Chromeo™ 488 antibody conjugate and the DNA-binding dye DAPI to counterstain the nucleus (see section 2.2.7 figure 2.16). In immunostained cells, mCherry fluorescence from NS2 fusion protein could be detected in the shell-like area of the ER around the nucleus, which correlates with earlier fluorescence microscopy results, indicating that the original cellular and protein structures are not affected by the chosen conditions for immunofluorescence staining. As expected, the Chromeo™ 488 fluorescence of the anti-*StreptII*-tag antibody completely co-localizes with the mCherry fluorescence, which confirms that *StreptII*-tag and mCherry are both located at the ER membrane as part of the NS2 fusion protein. Unfortunately, while the immunofluorescence staining was successful in principle, upon fluorescence microscopic examination of a great number of cells it became clear that the crystalline structures had not survived the cell preparation procedure. Therefore, the identity of the crystallizing matter remains an object of speculation.

A microscopic examination of the first cell preparation steps revealed that the crystalline structures already disappeared during fixation. This effect was independent of the chosen fixation protocol (either based on glutaraldehyde or paraformaldehyde). Immunostaining of crystals in fixed cells is therefore not an option to identify the nature of the crystallizing matter. While antibodies unfortunately cannot pass the membrane of a living cell on their own, a potential alternative to the immunostaining of fixed cells could be the delivery of antibodies into the cytosol of living cells with the aid of a special delivery reagent like Ab-DeliverIN™ Transfection Reagent (Ozbiosciences), PULSin® (Polyplus transfection®) or Chariot™ Protein Delivery Reagent (Active Motif®). However, since these reagents are lipid-based, their effects on the putatively ER membrane-embedded protein crystals cannot be predicted.

NS2 with its 7 transmembrane helices is embedded in the ER membrane of the original bovine host cell and – based on earlier findings in this thesis – likewise in the ER membrane of its insect cell expression host. If the *in vivo* crystals indeed consist of NS2, the encompassing membrane can be expected to be a distinct feature of the crystal structure. For *in vivo* grown protein crystals from *TbCatB*, high resolution transmission and scanning electron microscopy (TEM/SEM) revealed a highly ordered crystal composition enveloped by membranous structures [131].

To gain some insight into the structural composition of the crystals grown in High Five™ cells expressing ER-targeted NS2 fusion proteins, sections of fixed, immobilized cells were imaged with a transmission electron microscope and searched for crystal-like structures protruding from the cells (see section 2.2.8 figure 2.17). While the infected cells exhibited clear symptoms of a baculovirus infection (increased size, enlarged nucleus, granular appearance) and even baculovirus cross-sections were visible, no crystal-like structures were present. In-

stead, the infected cells contained "furry" structures in their nucleus and cytosol, which could not be detected in uninfected control cells. However, although no infected control cells were imaged, it is not likely that these structures are remnants of the missing crystals, since comparable structures have been imaged by TEM in insect cells infected with Baculovirus heliothis elsewhere (see section A.3 figure A.8). They can therefore most likely be attributed to the baculovirus infection itself.

As further experiments revealed, the insect cell preparation for TEM suffered from the same problem as the immunofluorescence preparation procedure (see above): upon fixation, the protruding crystals disintegrated and disappeared. As an alternative for further TEM experiments, cryogenic transmission electron microscopy of vitreous sections (CEMOVIS) should be considered [5].

3.2.7. *In vivo* crystals do not diffract micro-focused synchrotron X-radiation *in cellulo*

In vivo-grown crystals are already an intriguing biological feature. But their existence, origin and formation mechanism are not the only issues of scientific interest. Since *in vivo* crystallization is not a rare event, but on the contrary is observed for a growing number of overexpressed proteins in insect cells [74], it can also be exploited for the structural characterization of heterologously expressed protein. *In vivo* crystallization circumvents conventional, laborious protein purification and crystallization and enables X-ray diffraction experiments for protein structure determination either on crystal suspensions prepared from crystal-containing cells or directly *in cellulo* by serial crystallography [76, 34].

Since the ultimate goal of the heterologous expression of full-length NS2 in this thesis and beyond is the determination of its three-dimensional structure on a molecular scale, the occurrence of *in vivo*-grown crystal-like structures from insect cells infected with NS2 fusion protein constructs could be of great value. However, initial serial crystallography experiments with suspended, cryo-preserved cells did not yield any diffraction from the putative NS2 crystals. The results from immunofluorescence and TEM cell preparations discussed above already confirmed that the crystals are very delicate. Although crystals were in this case still visible after cryo-preservation prior to flash-freezing, it is possible that already the handling of the cells for the cryo-preservation and subsequent mounting onto a mesh loop as well as the flash-cooling event disrupts the inner crystal order, resulting in a lack of diffraction. To minimize cell manipulation, the cells could be directly grown in X-ray-transparent 96-well plates for further diffraction experiments. This technique has already successfully been implemented in pilot experiments with *TbCatB* crystals (Prof. Dr. Lars Redecke, oral communication).

Nevertheless, considering the membrane-embedded nature of the NS2 autoprotease, it is also conceivable, that the crystal-like structures do in fact only consist of a few layers of membrane, merely creating two-dimensional crystals instead of three-dimensional ones. This could explain the crystal fragility during the preparation for immunofluorescence and TEM experiments as well as the lack of detectable X-ray diffraction.

3.3. Conclusion and future prospects

For the production of protein suitable for further biophysical and structure determination experiments, full-length or *N*-terminally truncated NS2 fusion proteins representing either the ncp or cp BVDV biotype were expressed in *E. coli* and insect cells in the course of this thesis.

Since the expression of membrane proteins in *E. coli* can be a grueling task due to the bacteria's limited membrane resources and the toxic effects that some membrane proteins can have on the bacterial host cell, only *N*-terminally truncated variants of NS2 – omitting the first 5, 6 or all of its putative transmembrane helices – were tested as His₆-tagged SUMO or GFP fusion protein for expression in *E. coli*. These constructs were expressed in 14 different *E. coli* expression strains, in several growth media and under varying induction and expression conditions. Overall, only codon-optimized constructs without any hydrophobic helices could be overexpressed in most of the tested strains. However, the expressed fusion protein was nearly exclusively deposited in inclusion bodies. Minor amounts of seemingly soluble fusion protein observed after cell lysis and affinity chromatography was actually bound to solubility-mediating chaperones, as western blot experiments revealed. From a range of different solubilization buffer formulations, only buffers containing at least 0.08 % of the surfactant *N*-lauroylsarcosine (NLS) enabled fusion protein solubilization and subsequent affinity chromatography. Since most *N*-terminal truncations of NS2 render the autoprotease inactive, the authentic tertiary structure of the NS2 truncation variants could not be evaluated directly by detecting NS2 autoproteolytic activity. The assessment of folding-sensitive SUMO protease cleavage efficiency as indirect indicator for the folding state of NS2 in SUMO fusion proteins was unfortunately prevented by the presence of NLS, which significantly inhibited SUMO protease cleavage at concentrations above 0.02 %. Another indirect approach – solely relying on the presence of GFP fluorescence as indicative for proper folding of the adjacent fusion protein – was abandoned due to its questionable reliability.

In summary, although SUMO and GFP fusion proteins of *N*-terminally truncated NS2 variants could be successfully overexpressed in *E. coli* as well as solubilized and affinity purified from inclusion bodies with the help of the anionic surfactant NLS, the folding state of the truncated NS2 variants in the corresponding fusion proteins could not be determined.

The expression of activable full-length ncp or cp NS2 autoprotease as mCherry fusion into different compartments of two insect cell lines, however, yielded the complete fusion protein localized in the insect cell ER membrane, as fluorescence microscopy demonstrated. The presence of separate mCherry as well as separate *StreptII*-tagged full-length NS2 detected in in-gel-fluorescence SDS-PAGE and western blot experiments surprisingly confirmed autoprotease cleavage activity not only for Jiv90-activated cp NS2 variants, but also for the ncp variants. This indicates the presence of an activating factor – most likely a Jiv-homolog – in the insect cells. Interestingly, High FiveTM expressing ER-targeted NS2 fusion proteins also exhibited the growth of prominent *in vivo* crystals. However, all attempts to identify the crystallizing matter and characterize its crystal structure by immunofluorescence, TEM and X-ray diffraction experiments failed so far, due to the delicate nature of the crystals.

Taking these results together, insect cell expression of full-length NS2 is clearly superior to the expression of *N*-terminal truncation variants in *E. coli* and represents the promising basis

for subsequent biophysical characterizations – and ultimately the structure determination – of the NS2 autoprotease.

Here, different experimental routes can be conceived for future experiments: The classical approach of membrane protein structure determination would lead from the detergent-based extraction of full-length NS2 to the crystallization of protein–detergent complexes or alternatively *in meso* (also known as lipid cubic phase, LCP) crystallization [46]. A more recently developed technique would directly transfer the full-length NS2 from the insect cell membranes to discoidal lipid bilayers (termed nanodiscs) enclosed by either membrane scaffold proteins (MSPs) or a polymer (e. g. styrene-maleic-acid copolymer, SMA) without the need for accessory detergent [37, 185]. While the detergent-based approach completely extracts the membrane protein from its native, stabilizing lipid environment and the best detergent choice therefore has to be determined empirically, the transfer into nanodiscs ensures an almost native-like lipid environment for all subsequent experiments. In addition, membrane proteins embedded in nanodiscs are even suitable for LCP crystallization. They can also be structurally characterized with cryo-electron microscopy (CryoEM), which recently became increasingly popular for the determination of three-dimensional protein structures due to important technical advancements [215, 271]. In addition to these advantages for the handling of NS2 expressed in insect cells, the implementation of nanodiscs would furthermore enable the direct, cell-free expression of NS2 into their lipid bilayer, even circumventing the transfer of the membrane proteins from cellular membranes into nanodiscs. First steps on this path have already been undergone in the course of this thesis by creating full-length NS2 fusion constructs – based on the successful insect cell constructs – for the expression in the cell-free BY-2 tobacco cell extract system, which is currently established in our laboratory (see section 2.3 for further details) [42].

PART II

**Expression, Purification
and Biophysical Characterization
of *Sorghum bicolor* Deoxyhypusine/
Homospermidine Synthase**

4. *Sorghum bicolor* deoxyhypusine synthase

In 1971, Tetsuo Shiba and his coworkers reported the isolation of a new, basic amino acid from bovine brain: N^ϵ -(4-amino-2-hydroxybutyl)lysine, which they termed hypusine – a composite name derived from **hydroxy**putrescine and **lysine** [256, 257]. The presence of hypusine was subsequently reported for a broad range of mammalian tissue [180, 111]. Further studies revealed that hypusine was exclusively formed as a unique posttranslational modification on a single, acidic 18 kDa protein, that was ubiquitous and highly conserved in various cell lines, and from mammalian cells to archaea (see section A.5 figure A.14 for sequence alignments) [59, 52, 16].

This protein was soon identified as eukaryotic translation initiation factor 5A (eIF-5A, until 1991 known as eIF-4D), due to the close association of the hypusine modification with an increase in protein synthesis during lymphocyte growth stimulation experiments [60]. Recent studies showed that eIF-5A is not only well conserved, but also present in abundance. In *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and human HeLa cells, it is among the top 50 genes with the highest expression rates [136].

The hypusine modification was demonstrated to be the product of a two-step process, in which the aminobutyl moiety of the triamine spermidine (N -(3-aminopropyl)1,4-diaminobutane) is first transferred in a highly specific reaction onto the ϵ -amino group of a single, conserved lysine residue of eIF-5A to yield deoxyhypusine (N^ϵ -(4-aminobutyl)lysine) and 1,3-diaminopropane. In a second reaction, hypusine is generated by hydroxylation of the aminobutyl moiety of deoxyhypusine [205, 206, 214, 218, 208].

The triamine spermidine, as well as putrescine (1,4-diaminobutane) and the tetraamine spermine (N,N' -bis(3-aminopropyl)1,4-diaminobutane), are ubiquitous polyamines found in all eukaryotes. At physiological pH, their amino groups are protonated and thus positively charged, enabling them to interact with all kinds of negatively charged biomacromolecules, like DNA, RNA, phospholipids and proteins. Polyamines are involved in a variety of essential cellular functions, from proliferation to apoptosis [219, 98, 64]. However, one of their most distinct functions is in the generation of the unique (deoxy)hypusine residue on eIF-5A from spermidine.

The generation of deoxyhypusine and the subsequent hydroxylation to hypusine takes place during or shortly after translation of the eIF-5A precursor mRNA and is mostly regulated by intracellular polyamine levels. Thereafter, the hypusination of eIF-5A is neither reversed nor further modified, and the protein stays stable for a long period [75, 89, 202, 275, 72].

The formation of the hypusine residue at a specific lysine of eIF-5A is catalyzed by two different enzymes: The first step, the transfer of the aminobutyl moiety from spermidine onto the lysine residue, is performed by the NAD^+ -dependent deoxyhypusine synthase (DHS, also: DHPS; EC 2.5.1.46) [53, 293], while the subsequent hydroxylation reaction is realized by deoxyhypusine hydroxylase (DOHH; EC 1.14.99.29) [1]. In this thesis, the emphasis is placed on the context of the DHS reaction.

The enzyme DHS is ubiquitous in eukaryotic cells and highly conserved in eukaryotes and archaea (see section A.5 figure A.15 for sequence alignments). Its biological significance is based on a single, unique and crucial post-translational modification on the eIF-5A precursor. To fully understand the importance of DHS in the cellular context, it is necessary to gain deeper insight into the physiological role of its modified reaction product, (deoxy)hypusinated eIF-5A. The following section will therefore provide further information on eIF-5A. Subsequent sections will explore structural and mechanistic details of the deoxyhypusination reaction and the DHS enzyme (see section 4.2), as well as the plant-specific HSS, which evolved from DHS (see section 4.3).

4.1. Eukaryotic translation initiation factor 5A

As already mentioned above, hypusinated (or deoxyhypusinated) eIF-5A is ubiquitous in eukaryotes – as its archaeal homolog aIF-5A is in archaea [213, 89, 240, 300, 51, 16]. To date, about 1260 protein sequences belonging to the eIF-5A family have been reported for more than 600 different eukaryotic species – from metazoa over viridiplantae to fungi and unicellular organisms – as well as 113 aIF-5A sequences from archaeal species (entry PF01287 of the protein family database Pfam, <http://pfam.xfam.org/>). Especially the amino acid sequence around the hypusinated lysine residue is strictly conserved, but the complete eIF-5A still exhibits high sequence identity between distantly related species (see figure 4.1 and section A.5 figure A.14).

Human eIF-5A isoform 1 (eIF-5A-1), a representative of metazoan eIF-5A, comprises 154 amino acids, resulting in a molecular weight of about 16.7 kDa. Crystal structures from a variety of organisms – including human – have already been determined (PDB entries: 3CPF, *Homo sapiens* [274]; 5HY6, *Spodoptera frugiperda*; 3HKS, *Arabidopsis thaliana* [270]; 3ER0, *Saccharomyces cerevisiae*; 2EIF, *Methanococcus jannaschii* [127]; 1Z6, *Pyrococcus horikoshii* [301]). They all exhibit high similarity with a two-domain fold (see figure 4.1). The highly conserved *N*-terminal domain consists of 6 β -strands which form a partially open β -barrel, while the *C*-terminal domain resembles the oligonucleotide-binding fold (OB fold), comprising a slightly varying number of β -strands and α -helices [126]. The hypusination site (K50 in human eIF-5A) resides at the tip of a finger-like, long, unordered loop, located between two β -strands and protruding from the *N*-terminal domain. Like the hypusinated lysine, the surrounding amino acids in this loop are highly conserved [262].

Protein synthesis and cell viability experiments in eukaryotes or archaea confirmed that the unmodified eIF-5A precursor or a variant, in which lysine is substituted with arginine, cannot functionally replace (deoxy)hypusinated eIF-5A [245, 159]. In fact, overexpression of the precursor or the variant has no detectable effect on protein synthesis and cell viability – neither beneficial nor detrimental [261]. However, protein synthesis experiments *in vitro* or cell viability assays in *Saccharomyces cerevisiae* demonstrated that deoxyhypusinated eIF-5A is in principle capable of functionally substituting the fully modified eIF-5A [203, 204, 200]. In contrast, experiments in metazoan organisms, such as mouse [259], *Drosophila melanogaster* [216] and *Caenorhabditis elegans* [259], showed that the loss of deoxyhypusine hydroxylase (DOHH) activity – and therefore the lack of hypusinated eIF-5A – is lethal. Although the

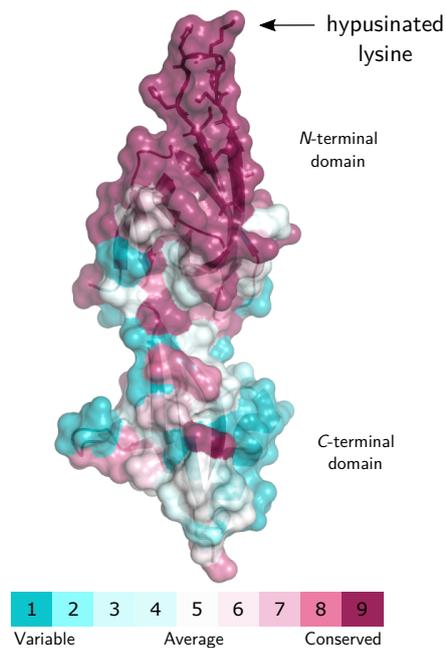


Figure 4.1.: Representation of *A. thaliana* eIF-5A tertiary structure and residue conservation

Cartoon representation of *A. thaliana* eIF-5A (PDB Acc. No.: 3HKS) overlaid with a surface representation, colored according to residue conservation. The conservation score was calculated with the ConSurf webserver (<http://consurf.tau.ac.il/2016/>, [9, 10, 50, 143, 87]). Individual side chains of the modified loop are depicted in stick representation. The graphic was generated with an open-source build of PyMOL (version 2.1.0) and Inkscape™ (version 0.92, <https://inkscape.org/>).

formation of deoxyhypusine itself is not impaired, deoxyhypusine levels are undetectable in DOHH knockdown mice, whereas in yeast deoxyhypusine levels are not altered in comparison to the wild type situation [259, 204]. Since the deoxyhypusination reaction is reversible [201], in metazoa hydroxylation might be necessary to stabilize and "cage" the hypusine modification without a further functional reason for the hydroxylation. This hypothesis is in accordance with findings from archaea, where no experimental evidence for the presence of a functional DOHH has been reported to date. Here, deoxyhypusinated aIF-5A is functionally sufficient and the dominant aIF-5A species. Although hypusine can actually be detected in some archaea species, the hydroxylation there is not attributed to DOHH activity [17].

Despite the quite specific substrate spectrum of DHS (see section 4.2.5 for further details), it also accepts *N*-(3-aminopropyl)cadaverine (*N*-(3-aminopropyl)1,5-diaminopentane) as substrate instead of spermidine and catalyzes the transfer of the aminopentyl moiety onto the lysine residue with low efficiency, resulting in the formation of homodeoxyhypusine (*N*^ε-(5-aminopentyl)lysine) [212]. Although homodeoxyhypusine and deoxyhypusine only differ in length by a single methyl group, homodeoxyhypusinated eIF-5A is not active in protein synthesis assays, indicating that the exact length of the modification is crucial for its function in eukaryotic ribosomes.

4.1.1. The role of eIF-5A and its homologs in protein synthesis

Early results from model translation initiation assays revealed that while eIF-5A stimulated methionine-puromycin synthesis – an assay which reports on translation initiation and elongation, based on the transfer of the initiator methionine onto the aminoacyl-tRNA analog puromycin –, it exhibited no influence on assays monitoring other steps of translation initiation [23, 24, 247]. This already indicated that eIF-5A – despite its eponymous assignment as initiation factor – was more likely involved in the late stages of translation initiation and the elongation phase.

Interestingly, amino acid sequence alignments and structural comparisons revealed that bacterial elongation factor P (EF-P) is a homolog of eukaryotic and archaeal a/eIF-5A. Although it does not feature a hypusinated lysine, however, the respective residue and some of the surrounding amino acids are still conserved, and a modification resembling the hypusination of eIF-5A was reported for EF-P from a number of bacterial strains: (*R*)- β -lysyl-lysine [199, 41], which can in addition also be hydroxylated (see figure 4.2). Intriguingly, this modification is more similar in length to the homodeoxyhypusine modification mentioned above than to hypusine. In other cases, the conserved lysine is replaced with an arginine, which is modified by rhamnosylation [146, 226]. Obviously, bacterial and eukaryotic/archaeal translation machineries have developed a comparable need for a modification of the specific lysine (or in some bacteria: arginine), but with different requirements for its exact length and chemical properties. However, while hypusinated eIF-5A in eukaryotes or deoxyhypusinated aIF-5A in archaea are essential for viability in all species, EF-P or its β -lysylated or rhamnosylated variants in bacteria are not.

Recent experiments revealed that a/eIF-5A and their bacterial counterpart, EF-P, play a crucial role in the translation of polyproline containing proteins, which are poor substrates for peptide bond formation [217, 291]. The synthesis of polyproline-containing proteins is stalled at these polyproline stretches. A/eIF-5A as well as EF-P have been demonstrated to stimulate the peptide bond formation in these cases [97, 69, 220, 278]. These polyproline-containing proteins are abundant in more complex organisms such as eukaryotes, but rare in prokaryotes, which correlates with the essential nature of eIF-5A in eukaryotes compared to the lesser importance of EF-P in bacteria [162].

Further investigations led to the assumption that eIF-5A might not only be necessary for the accelerated translation of polyproline containing proteins, but for a much broader range of proteins, that contain all kinds of ribosome-stalling motifs – including polyproline stretches. In addition, eIF-5A stimulated release factor 1 (RF1)-mediated peptide release. In the absence of eIF-5A, translation termination was therefore severely impaired on all tested peptides, indicating an essential role in the termination process. In contrast, the function of bacterial EF-P seems to be limited to the stimulation of the translation of polyproline-containing proteins [249, 221].

The determination of the molecular structure of eIF-5A bound to the eukaryotic ribosomal complex by X-ray diffraction and cryo-electron microscopy (CryoEM) revealed that eIF-5A, which structurally mimics the acceptor stem of tRNA, is localized at the ribosomal E-site and probably stabilizes the peptidyl-carrying 3'-CCA-end of the tRNA in the P-site with its hypusine residue [97, 244, 166]. Experiments with bacterial EF-P indicated that EF-P, which

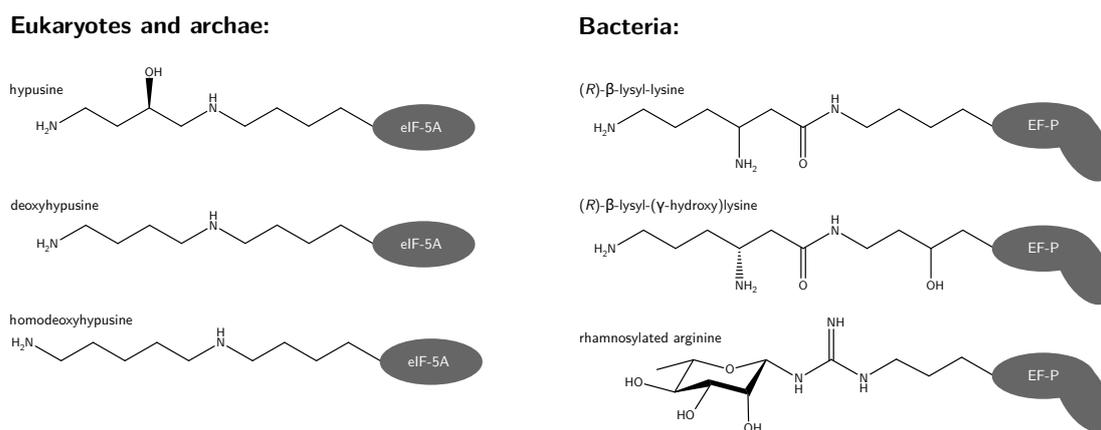


Figure 4.2.: Unique posttranslational modifications on eukaryotic and archaeal eIF-5A and its bacterial counterpart EF-P

structurally resembles the acceptor and – in contrast to eIF-5A – also the anti-codon stem of a tRNA, reduces reaction entropy by stabilizing an optimal geometry of the peptidyl-tRNA through interaction with its 3'-CCA-end [70, 110]. This orienting and stabilizing interference might be especially necessary if structurally inflexible polyproline peptides are present at the P-site.

Based on these results and their own findings, Schuller and coworkers proposed that if the ribosome is engaged in peptide bond formation on stalling motifs, such as polyproline stretches, or in the terminating RF1-mediated hydrolysis for a prolonged period, the discharged tRNA leaves the E-site without being directly replaced by the former P-site tRNA. The empty E-site can then be occupied by eIF-5A, stimulating the translation/termination process by accelerating the peptidyl transfer at the ribosomal peptidyl transferase center (PTC) through optimization of the peptidyl-tRNA orientation for peptidyl transfer or hydrolysis [249, 221]. In addition to this function as hypusine-carrier, further studies implicate the eIF-5A protein itself – independent of its hypusine modification – as general stimulating agent in protein synthesis. In these studies, eIF-5A was able to substitute for polyamines such as spermine, which have been reported to be required for high translational activity *in vitro* [258].

4.1.2. Involvement of eIF-5A in pathologic processes

In addition to its crucial function in protein synthesis, eIF-5A has been implicated in a number of other cellular processes, which are also connected to a variety of pathologic conditions. A selected overview highlighting some further roles for eIF-5A and its unique modification is given in this section.

As already mentioned in previous sections, the inhibition of eIF-5A expression or (deoxy)hypsination severely affected cell viability and cell cycle progression in a number of experiments. Studies specified cell cycle arrest to occur at the G(1)-S phase transition upon depletion of hypusinated eIF-5A. Since no direct influence of eIF-5A on the cell cycle regulation has been found to date, it seems most likely that eIF-5A exerts its control over the cell

cycle progression by promoting the translation of relevant genes [209, 100, 122, 116, 304]. Due to its influence on cell cycle and proliferation, elevated levels of eIF-5A have been connected to several proliferative diseases [164]. In this context, eIF-5A was, for instance, found to act as nucleocytoplasmic shuttle protein, which might translocate cell cycle-relevant mRNAs [237, 4, 172]. Targeting the eIF-5A/DHS/DOHH network could therefore be promising in the treatment of cancer [182, 250].

In diabetes, hypusinated eIF-5A has also been shown to promote the β -cell inflammation, and the inhibition of DHS supported β -cell survival and insulin production [236, 161, 160].

However, eIF-5A is not only involved in the pathologic conditions related to cellular malfunctions, but also in the progression of certain viral diseases. For the expression of the structural proteins of human immunodeficiency virus type 1 (HIV-1), an ssRNA(+) virus of the *Retroviridae* family, the presence of the viral trans-activator protein Rev is essential. Rev specifically binds to the Rev response element (RRE) and mediates the nucleocytoplasmic transport and translation of incompletely spliced and unspliced viral transcripts [224, 92]. Rev activation domain was shown to specifically interact with host cell eIF-5A. Rev activation is therefore assumed to depend on the presence of eIF-5A, and its functionality *in vivo* can be blocked by inhibition of eIF-5A expression or (deoxy)hypusination [238, 28]. The components of the eIF-5A expression and modification pathway are thus highly interesting drug targets for the treatment of HIV-1 infections [106, 248].

Two other lethal human pathogens, Ebola virus (EBOV) and Marburg virus (MARV), members of the *Filoviridae* family of ssRNA(-) viruses, strongly rely on eIF-5A modification for their replication. Inhibition of deoxyhypusination could be correlated to a specific reduction of VP30, a phosphoprotein essential for the initiation of virus transcription. Since the expression of other viral proteins was not affected, eIF-5A is assumed to be necessary for efficient VP30 mRNA translation. [194]

In *Aedes albopictus*, elevated eIF-5A levels upon dengue virus type 2 (DENV-2) infection are discussed as protective measure of the infected mosquito cells. While mammalian host cells regularly undergo apoptosis due to protein synthesis shutdown upon DENV-2 infection, the insect host cells do not suffer from detrimental effects [62].

These selected scientific findings clearly show that the biological role of eIF-5A is still not fully understood to date and further efforts have to be made to exploit the eIF-5A/DHS/DOHH system for therapeutic purposes. Since the deoxyhypusination of eIF-5A is crucial for most of its functions, the controlled and specific inhibition of DHS is a central aspect of therapeutic endeavors. It is therefore essential to understand the structural and mechanistic context of the deoxyhypusination reaction, which will be described in detail in the following sections.

4.2. Deoxyhypusine synthase

4.2.1. The proposed reaction mechanism of deoxyhypusine formation

Deoxyhypusine synthase (DHS) specifically catalyzes the deoxyhypusination of eIF-5A as the first, reversible step in the post-translational formation of hypusine (N^ϵ -(4-amino-2-hydroxybutyl)lysine), which is essential for eIF-5A activity in eukaryotes. Figure 4.3 provides

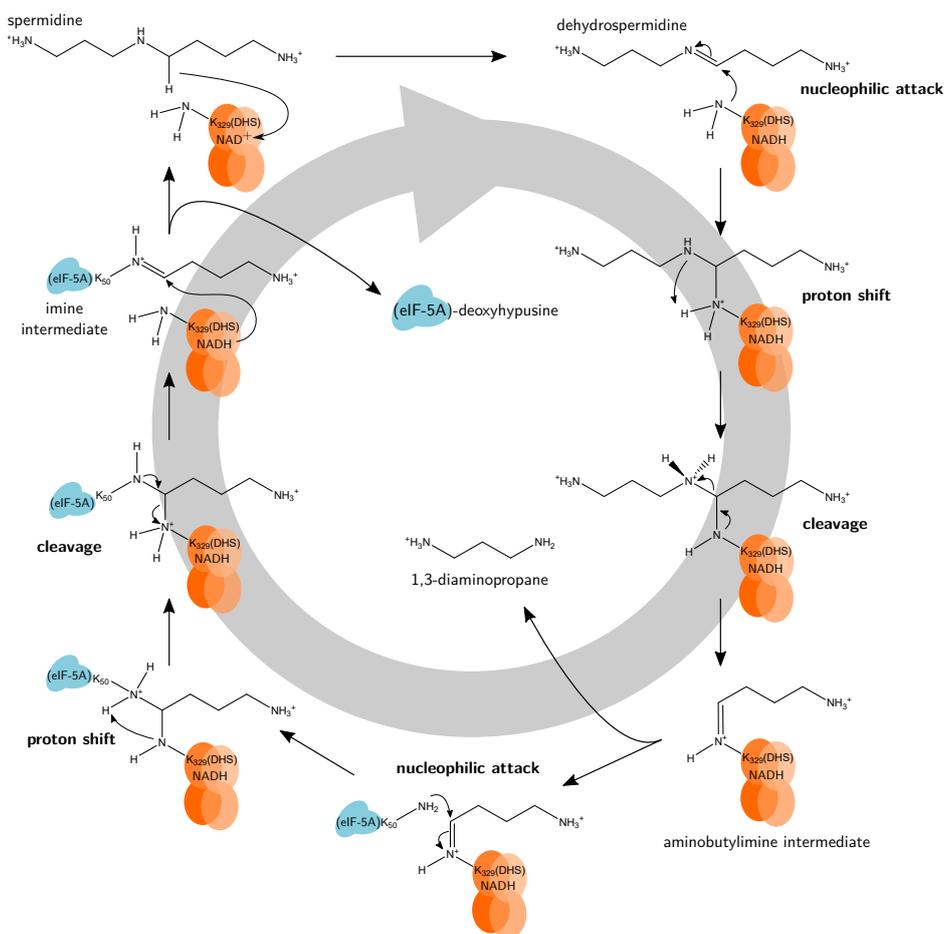


Figure 4.3.: Schematic representation of the putative DHS reaction mechanism

The scheme of the catalytic cycle of the DHS reaction was created based on the information and resources summarized in this section. It is intended as visual support and makes no claims of being complete.

an overview of the putative DHS reaction steps. The second reaction, the irreversible hydroxylation of deoxyhypusine (N^ϵ -(4-aminobutyl)lysine) to hypusine at the C2 of the 4-aminobutyl chain, is subsequently performed by deoxyhypusine hydroxylase (DOHH) [198].

The complex deoxyhypusine synthesis reaction can be subdivided into two distinct phases: in the first part, which can take place in the absence of eIF-5A, the spermidine substrate (N -(3-aminopropyl)1,4-diaminobutane) is cleaved in an NAD^+ -dependent reaction, generating an enzyme-bound 4-aminobutylimine intermediate at the ϵ -amino group of the catalytic lysine, NADH and 1,3-diaminopropane as remnant of the spermidine substrate.

In the second part of the reaction, the 4-aminobutyl moiety is transferred from the catalytic lysine onto the eIF-5A precursor, to yield an eIF-5A-aminobutylimine intermediate. This

intermediate is finally reduced to deoxyhypusine by the enzyme-bound NADH, which closes the catalytic cycle [118].

In the first step during the first phase, spermidine is dehydrogenated by NAD⁺ at C5 (or C1 of the aminobutyl moiety) [207], yielding the imine dehydrospermidine and 4*R*-NADH – NADH stereospecifically hydrated at its *re* face [293]. In watery solution, NADH exhibits strong fluorescence at 457–465 nm upon excitation with light of 340 nm wavelength. In contrast, in the presence of non-polar solvents or bound to a protein, the emission peak is blue-shifted by about 20 nm, allowing the distinction between tightly bound NADH and only loosely associated NADH, which might be exchanged during the reaction. Exploiting this biophysical characteristic, in 2000 Wolff and coworkers could demonstrate that NADH does not leave the active site, but stays bound to DHS to deliver the hydride to the last step of the reaction cycle, a mechanism termed "transient hydride transfer" [82]. In the absence of eIF-5A precursor, NADH was mostly maintained in the protein for more than 4 h at room temperature and over 24 h at 4 °C [294]. However, small amounts of NADH are slowly released over time. In more recent studies, these amounts were utilized to quantify DHS activity via the luciferase-dependent NAD⁺/NADH-GloTM Assay (Promega) [210].

In the second step of the first phase, dehydrospermidine is nucleophilically attacked at C5 by the ϵ -amino group of the catalytic lysine (K329 in human DHS [292]). In a transimination reaction, the resulting intermediate undergoes a proton shift and 1,3-diaminopropane is subsequently released, resulting in the 4-aminobutyl moiety of spermidine linked to K329 via an imine linkage (Schiff base). The presence of a covalent intermediate was demonstrated by Wolff and coworkers in 1997 by trapping the intermediate with sodium cyanoborohydride (NaBH₃CN) [292]. In the absence of eIF-5A, the intermediate is slowly converted to 1-pyrroline, leaving the DHS with bound NADH unavailable for further reaction cycles until NADH also slowly dissociates (see figure 4.6) [293, 292]. Since the appearance of 1-pyrroline is delayed in relation to the release of 1,3-diaminopropane, it can be assumed that 1-pyrroline indeed results from the intramolecular nucleophilic attack of the primary amine on the imine at the secondary nitrogen of the enzyme intermediate, rather than from an intramolecular reaction of dehydrospermidine [292].

In the presence of eIF-5A, the third and fourth reaction step (phase two) can be completed. Therefore, the aminobutyl moiety is transferred in a second transimination reaction from the ϵ -amino group of the catalytic lysine of DHS onto the ϵ -amino group of the acceptor lysine residue of the eIF-5A precursor to form yet another imine intermediate. This imine intermediate is subsequently reduced to deoxyhypusine by the enzyme-bound NADH.

4.2.2. Tertiary and quaternary structure of DHS and relevance of specific amino acid residues

In 1995, the purification of DHS from rat testis delivered first insights into the structural nature of the enzyme [295]. Amino acid sequences derived from tryptic digests of rat testis DHS subsequently enabled the identification and cloning of DHS from *S. cerevisiae* and human [118, 124]. In these studies, DHS was demonstrated to be a tetramer of four identical subunits with a molecular weight of about 42 kDa.

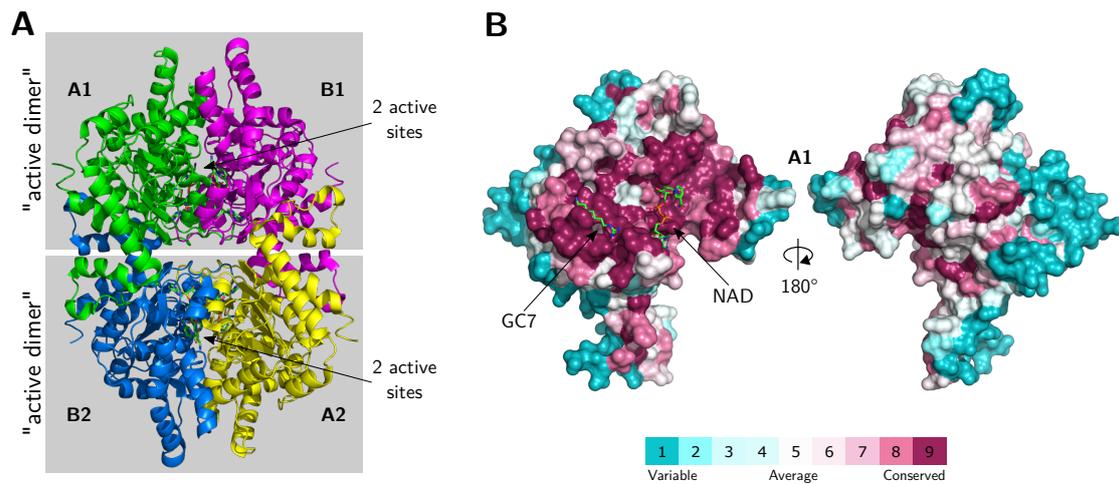


Figure 4.4.: Representation of human DHS tertiary structure and residue conservation

A: Cartoon representation of human DHS homotetramer (PDB Acc. No.: 1RQD) with NAD^+ and the inhibitor GC7 in the active sites. **B:** Surface of a single subunit of the DHS tetramer, colored according to residue conservation. Left: "active dimer" interface with NAD^+ and GC7 in the two different active sites; right: solvent-exposed backside. The conservation score was calculated with the ConSurf webserver (<http://consurf.tau.ac.il/2016/>, [9, 10, 50, 143, 87]). The graphic was generated with an open-source build of PyMOL (version 2.1.0) and Inkscape™ (version 0.92, <https://inkscape.org/>).

The first crystal structure of a DHS – human DHS in complex with its cofactor NAD^+ – was solved in 1998 by Liao and coworkers [153]. This structure confirmed the homotetrameric nature of the DHS and revealed that the individual subunits were arranged in two tightly associated dimers, which in turn dimerize to form the active tetramer ("dimer of dimers", see figure 4.4 **A**).

Four active sites were identified in the tetramer: At the highly conserved interface of each tightly associated dimer (see figure 4.4 **B**), two antiparallel narrow tunnels are formed by the combined residues of both subunit interfaces. The tunnels each start at a broad funnel on the enzyme surface (see section 4.3 figure 4.10) and intrude about 17 Å-deep into the enzyme core. Each tunnel contains an active site, in which an NAD^+ molecule bound to a Rossmann dinucleotide-binding motif could be identified. The NAD^+ molecule is in contact with both enclosing subunits – leaving only the nicotinamide ring accessible from the active site tunnel [153]. Aspartic acid 313 of one subunit, which interacts with the adenine moiety of NAD^+ , and aspartic acid 238 of the other subunit, which forms a hydrogen-bond with the nicotinamide ribose, are completely conserved in eukaryotes, and aspartic acid 342, which interacts with the adenine ring of NAD^+ , is in some cases replaced by glutamic acid. The mutation of either of these residues to alanine prevents binding of NAD^+ – and subsequently spermidine (see below) –, thereby confirming the importance of these residues in correctly orienting and anchoring the nicotinamide ribose and the adenine ring for the following binding and reaction steps (see section A.5 figure A.15 for a sequence alignment of eukaryotic and archaeal DHS sequences and highlighting of specific residues mentioned in this section) [149].

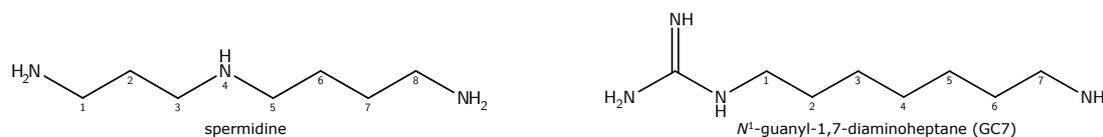


Figure 4.5.: Structural representation of DHS substrate spermidine and its inhibitor GC7

Interestingly, in the first structure of human DHS each of the four entrances to the active sites was obstructed by a so-called "ball-and-chain" motif. This motif was composed of a terminal two-turn α -helix, which consisted of amino acids alanine9 to valine17 of a third subunit from the other tightly associated dimer, attached to an extended loop structure (leucine18 to proline25). The helix was held in place directly in front of the entrance into the active site tunnel, mainly by hydrophobic interactions. Since crystallization was performed at an acidic pH of about 4.5 and high ionic strength – conditions, which intensify hydrophobic interactions and under which DHS is not active *in vitro* – the structure was assumed to represent an inactive conformation. This was in accordance with the finding that although NAD^+ was readily identified in the active site, the most potent DHS inhibitor and spermidine analog *N*¹-guanyl-1,7-diaminoheptane (GC7, see figure 4.5) [114, 211], which was also present in the mother liquor, was not [153].

Indeed, in the DHS structure from crystals obtained at pH8.0 and low ionic strength – conditions near the pH optimum of DHS around pH9.0 [151, 148] –, the ball-and-chain motif could not be identified in the electron density, indicating that it did not adopt a distinct conformation, but moves freely in an unordered fashion from a hinge region around glutamic acid27 and serine28 [280]. This conformational change was therefore independent of the binding of eIF-5A or other ligands. The ball-and-chain motif was thus suspected to have a regulatory role. However, earlier studies already showed that a deletion of the ball-and-chain motif did not increase DHS activity. In fact, a deletion of more than the first 10 *N*-terminal amino acids significantly reduced DHS activity, and with a deletion of the first 20 amino acids activity is abolished completely [123]. This indicates a crucial role of most of the ball-and-chain motif for DHS activity.

Besides the visual absence of the ball-and-chain motif in the electron density, the accessibility of the active sites in the new structure was in addition demonstrated by the presence of the inhibitor GC7 in the active sites of soaked crystals [280]. In the structure the inhibitor is oriented with its amino group near the entrance of the active site tunnel, where it is stabilized by interactions with asparagine292 of one subunit and aspartic acid243 of the other, and is in close contact with the nicotinamide ring of NAD^+ . The guanidinium group of GC7 is localized towards the inner end of the tunnel, making contact with aspartic acid316 and glutamic acid323. These residues as well as aspartic acid243 are highly conserved in eukaryotic DHS and a mutation of either of these residues to alanine nearly abolished spermidine binding [149]. Therefore, they are likely crucial for the binding of the terminal amino groups of spermidine. The active site tunnel is additionally lined with further negatively charged residues, namely glutamic acids136 and 137, which are also essential for spermidine binding [149, 280].

At the narrowest region of the active site tunnel, the ϵ -amino group of the catalytic K329, the ϵ^2 -nitrogen of histidine288 and C4, the hydrogen-accepting carbon of the NAD⁺ nicotinamide ring, are positioned closely to C4 of GC7, which correlates to the position of the secondary amine in spermidine [280]. Tryptophan 327, which is completely conserved in eukaryotes and archaea, is positioned opposite of the imidazole ring of histidine288 and the nicotinamide ring of NAD⁺, close to the putative position of the spermidine secondary nitrogen. A mutation of this residue to alanine significantly impairs NAD⁺ and spermidine binding [149]. Tryptophan 327 is assumed to improve spermidine binding by hydrophobic interactions and might contribute to the deoxyhypusination reaction by stabilizing an intermediate cation by cation- π interaction – a strong, non-covalent interaction between a positively charged moiety and the π -face of an aromatic system [73].

This configuration is well suited to perform the deoxyhypusination reaction, according to the proposed mechanism outlined above. Histidine288 would act as proton acceptor during the dehydrogenation of spermidine to the imine dehydrospermidine by C4 of NAD⁺. The imine is then attacked nucleophilically by the ϵ -amino group of the catalytic K329, resulting in the formation of the enzyme–imine intermediate. Protonated histidine288 could also act as proton donor in the following reaction steps. In addition, the carboxyl group of the highly conserved glutamic acid 137, a negatively charged residue of the active site tunnel, is also in the immediate vicinity of C4 of GC7 and C4 of the nicotinamide ring [280]. While enhancing NAD⁺ binding, a mutation of this residue to alanine completely abolishes spermidine binding, emphasizing its importance for the deoxyhypusination reaction [149].

A side pocket is localized at the bottom of the active site tunnel. It is lined with highly conserved polar and charged side chains. The already mentioned aspartic acid 316 and glutamic acid 323, which interact with the guanidinium group of GC7, are positioned at the junction of tunnel and side pocket. Tyrosine305 forms the back wall of the pocket. A mutation of this residue to alanine impaired binding of both spermidine and NAD⁺ [149]. Although tyrosine 305 is not in direct contact with either NAD⁺ or spermidine, it might influence the binding of NAD⁺ via a neighboring strand, which is connected to the loop that binds the adenine ring of NAD⁺. The side chain of lysine 287 protrudes into the void of the side pocket, forming a salt bridge with aspartic acid 316. For lysine287, Umland and coworkers proposed a role as modulator of the concentration of negative charges in the active-site tunnel and for the definition of the volume of the side pocket, which might serve to accommodate a reaction intermediate [280].

4.2.3. Binding sequence of NAD⁺, spermidine and eIF-5A

As sequential binding experiments revealed, the binding of the two substrates spermidine and eIF-5A and the NAD⁺ cofactor to DHS in the course of the deoxyhypusination reaction seems to be interdependent to a certain extent (see figure 4.6 for a schematic overview).

The interaction of eIF-5A with DHS does not require NAD⁺ or spermidine and exhibits high affinity [151]. Actually, the deoxyhypusination reaction proceeds regularly after association of eIF-5A with DHS upon addition of NAD⁺ and spermidine, which indicates that the presence of eIF-5A at the entrance to the active site does not interfere with the subsequent binding of the missing reaction partners. However, the binding of spermidine or GC7 to the active site is

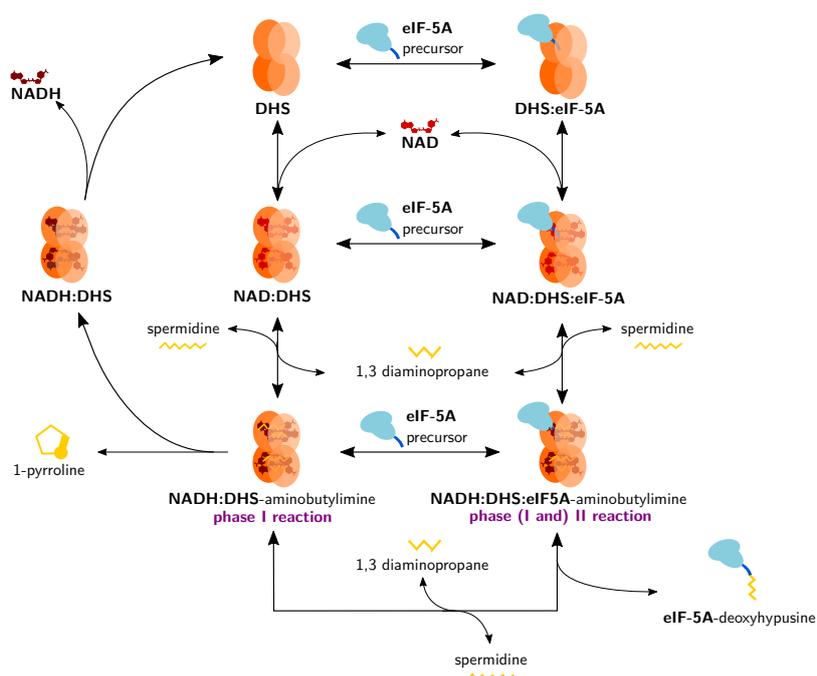


Figure 4.6.: Schematic representation of DHS substrate binding sequence and its catalytic cycle

This scheme was created based on the information summarized in this section and figures from Wolff et al. (2000) and Lee et al. (2000) [294, 148] with Inkscape™ (version 0.92, <https://inkscape.org/>).

crucially dependent on the prior binding of NAD^+ – both in the presence and absence of eIF-5A [148, 149]. In contrast, NAD^+ does not require the presence of spermidine or eIF-5A for DHS binding, although if both present, NAD^+ and spermidine mutually enhance each other's binding affinity. Since this is also true for the inactive component pairs NADH/spermidine and NAD^+ /GC7, this effect might be due to some conformational adaptations, which improve binding of both components. This is supported by results from fluorescence measurements, which revealed quenching of the intrinsic tryptophan fluorescence upon binding of NAD^+ , and further quenching at the subsequent addition of spermidine – indicating conformational changes during complex formation [148, 294, 149]. Although eIF-5A does not depend on the presence of NAD^+ and spermidine for binding, the detailed effects that these conformational changes might also have on the affinity between eIF-5A and DHS have not been studied yet.

Taken together, the dependence of spermidine binding on the prior association of NAD^+ to DHS enforces a certain sequence of events at this point. Nevertheless, in the presence of all necessary components, the other binding events probably occur randomly, only influenced by the respective concentrations and association/dissociation constants of the individual component species.

4.2.4. DHS cooperativity effects upon binding of eIF-5A

As already described in section 4.1, the hypusinated lysine residue is located at the tip of a flexible loop protruding from the *N*-terminal domain of eukaryotic and archaeal eIF-5A alike. The surrounding residues from phenylalanine 30 to aspartic acid 80 – encompassing most of the *N*-terminal domain – are crucial for the deoxyhypusination reaction, which explains their high conservation among species (see figure 4.1 and section A.5 figure A.14 B) [262, 117]. The mostly basic and hydrophilic residues directly enclosing the modified lysine residue in the conserved loop (–STSKTGK₅₀HGHAK– in eukaryotes, see figure 4.7 and the alignment in section A.5 figure A.14) are especially critical: Point mutation of most of these residues directly impairs deoxyhypusination [67, 48]. This correlates with many positively charged residues on the funnel surface leading to the active site tunnel (see figure 4.10), and with results from a mutation of the aspartic acid 243 in the vicinity of the tunnel entrance to alanine. While the mutation abolishes spermidine binding (see section 4.2.2), it in contrast enhances eIF-5A binding – probably due to the elimination of a potentially repulsive negative charge. Interestingly, the mutation of the catalytic DHS residue lysine 329 to alanine also seems to improve the association of eIF-5A, while reducing spermidine binding and naturally abolishing deoxyhypusination. Both mutations have another intriguing feature in common, which also distinguishes them from other mutants: while the stoichiometry of wild-type DHS(tetramer)–eIF-5A(monomer) complexes was estimated to be 1:1 as derived from experimental results [151], these two mutations exhibit 1:4 stoichiometry [149].

Although the active DHS tetramer provides four seemingly equal active sites and nearly complete NAD⁺ occupancy [153, 294, 280], the wild-type DHS(tetramer)–eIF-5A(monomer) complex stoichiometry led to the assumption that only one active site at a time performs the deoxyhypusination reaction. Meanwhile, the other sites are not active, and it was speculated that this was either because of cooperative conformational changes alone or in combination with additional sterical hindrance for two eIF-5A molecules on the same side of the DHS tetramer. The observed 1:4 stoichiometry for the two mutants D243A and K329A obviously refutes the latter assumption.

In addition, results from different experiments are in support of negative cooperative behavior [151, 294, 280]. As already mentioned in section 4.2.2, the crystal structures revealed that each DHS subunit maintains considerable contact with two other subunits: either by forming the active dimer (A1–B1 and A2–B2) or the tetramer, the "dimer of dimers" (A1–B2 and A2–B1). The contact to the corresponding fourth subunit (A1–A2 and B1–B2) is not extensive, but distinct. Besides the ball-and-chain motif, which is capable of blocking the active site on the opposite active dimer (see section 4.2.2 [153]), phenylalanine 54 – e. g. of subunit A1 –, which is highly conserved in eukaryotes and archaea, interacts with a loop of the corresponding fourth subunit (A2) and another one from the third subunit (B2), which are both in direct contact with the same NAD⁺ via asparagine 106, aspartic acid 313 and glycine 314. The first two residues interact with the adenine moiety of NAD⁺ at the bottom of the active site tunnel while glycine is in contact with the nicotinamide [280].

Intriguingly, the substitution of aspartic acid 313 to alanine results not only in a severe decrease of NAD⁺ binding, but also in a significant reduction of eIF-5A binding [149]. Since eIF-5A binding does not depend on NAD⁺ and the residue is positioned deep inside the

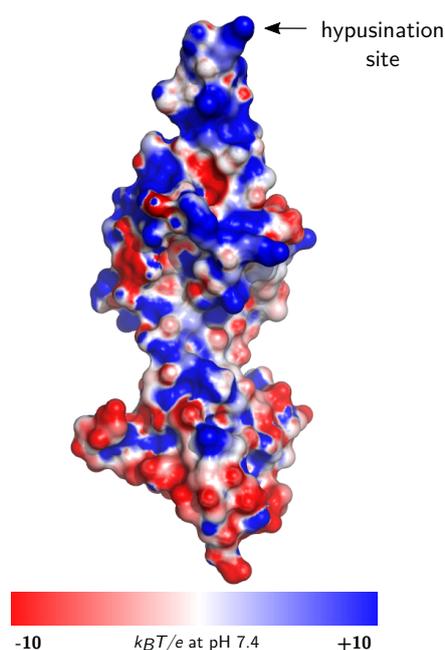


Figure 4.7.: Representation of *A. thaliana* eIF-5A surface electrostatic potential at pH 7.4

Surface representation of *A. thaliana* eIF-5A (PDB Acc. No.: 3HKS), colored according to the residues' electrostatic potential at pH 7.4. The electrostatic potential was calculated with the PDB2PQR webserver (version 2.1.1, http://nbcx-222.ucsd.edu/pdb2pqr_2.1.1/, [71]) and APBS (version 1.4.1, [119]). The graphic representation was generated with an open-source build of PyMOL (version 2.1.0) with the APBS plugin and Inkscape™ (version 0.92, <https://inkscape.org/>).

active site tunnel, this effect can most likely be attributed to the impaired interaction with phenylalanine 54 of the other subunit. For the crystal structures of DHS with and without the spermidine analog GC7, no significant conformational changes were described. Taking these findings together, it is therefore most likely that the binding of eIF-5A is the critical event that triggers the conformational changes which prevent the binding of further eIF-5A molecules [280]. This regulatory effect might also explain why DHS is an obligate homotetramer, although a homodimer is the active subunit. However, since no structures of DHS in complex with eIF-5A are available to date, these assumptions must remain speculative.

4.2.5. DHS substrate specificity, inhibitors and side reactions

Already early experimental results revealed an unusual basic pH optimum of the deoxyhypusination reaction: in crude cell lysates deoxyhypusine was generated at a pH of 9.5, while the subsequent hydroxylation to spermidine took place at about pH 7.1 [179]. These findings were later also confirmed for purified and reassembled reaction mixtures. Although deoxyhypusine formation could also be demonstrated at neutral pH, it was much less efficient than at basic pH for low spermidine concentrations [72, 118]. In contrast, high spermidine concentrations were shown to compensate for the low reaction efficiency [151].

The basic pH optimum is most likely not an inherent feature of the DHS enzyme itself, but is owed to the deoxyhypusination reaction mechanism, which requires the secondary amino group of spermidine to be uncharged, as inhibitor studies and the crystal structures consistently imply [114, 153, 280]. The three amino groups of spermidine are considered to have slightly different pK_a values, ranging from pH 8.4 for the secondary amino group to about 9.9 and 10.8 for the two terminal primary amino groups [11]. At neutral pH, the majority of spermidine molecules would be triprotonated, and therefore only a minor portion would be available for the deoxyhypusination reaction. To increase the available fraction, higher spermidine concentrations would have to be applied. In contrast, at pH 9.5 the proton at the secondary amino group will be abstracted in the majority of spermidine molecules, resulting in sufficient concentrations of diprotonated spermidine even at low overall spermidine concentrations [114, 151].

The special requirements of the active site of DHS strongly limit the enzyme's substrate spectrum. [208]. Early studies about DHS substrate requirements, specificity and inhibitory effects revealed a number of compounds that were structurally related to spermidine and inhibited deoxyhypusine formation on eIF-5A. Some of these compounds were the diamines 1,3-diaminopropane, 1,7-diaminoheptane, 1,8-diaminooctane and 1,9-diaminononane, the triamines homospermidine (*bis*(4-aminobutyl)amine), caldine (also: norspermidine; *bis*(3-aminopropyl)amine) and *N*4-benzylspermidine, and the tetraamine norspermine (also: thermine; *N,N'*-*bis*(3-aminopropyl)-1,3-diaminopropane) [208]. From these and later experiments it became clear that effective inhibition could be achieved with diamino compounds consisting of two primary amino groups – which can optionally be guanylated – and a connecting carbon chain, that was similar in length to spermidine (e. g. 1,7-diaminoheptane or 1,8-diaminooctane; *N*¹-guanyl-1,7-diaminoheptane or -octane; *N*¹,*N*⁷-*bis*guanyl-1,7-diaminoheptane or *N*¹,*N*⁸-*bis*guanyl-1,8-diaminoheptane) [114, 211]. However, a secondary amino group – as present in spermidine – was not a necessary prerequisite. In fact, *N*¹-guanyl-1,7-diaminoheptane (also termed GC7, see figure 4.5), is the most potent DHS inhibitor *in vitro* and *in vivo* known to date.

These findings are concurrent with the structural circumstances at the active site of DHS, which possesses an extended, narrow tunnel with two binding sites for the positively charged, terminal amino groups of spermidine (see section 4.2.2). Monoamines or compounds with derivations of one or both primary amino groups (except for guanyl groups) as well as compounds with space-demanding or flexibility-limiting derivations at the carbon chain therefore were naturally inefficient inhibitors (or substrates, see below).

Interestingly, the short diamines 1,3-diaminopropane – a product of the deoxyhypusination reaction (see section 4.2.1) – as well as 1,4-diaminobutane (putrescine) were also identified as surprisingly strong DHS inhibitors [208, 114]. Since they were apparently too short to substitute for spermidine in the active site tunnel, their inhibitory effect was assumed to be based on some kind of product inhibition. However, in 2003 Park and coworkers could demonstrate that the deoxyhypusine reaction was in fact completely reversible in the presence of 1,3-diaminopropane and NAD^+ [201]. The inhibitory effect of 1,3-diaminopropane was therefore due to the promotion of the reversed reaction – from deoxyhypusinated eIF-5A back to spermidine. Likewise, putrescine could replace 1,3-diaminopropane in this reversed

DHS substrates and reaction products:

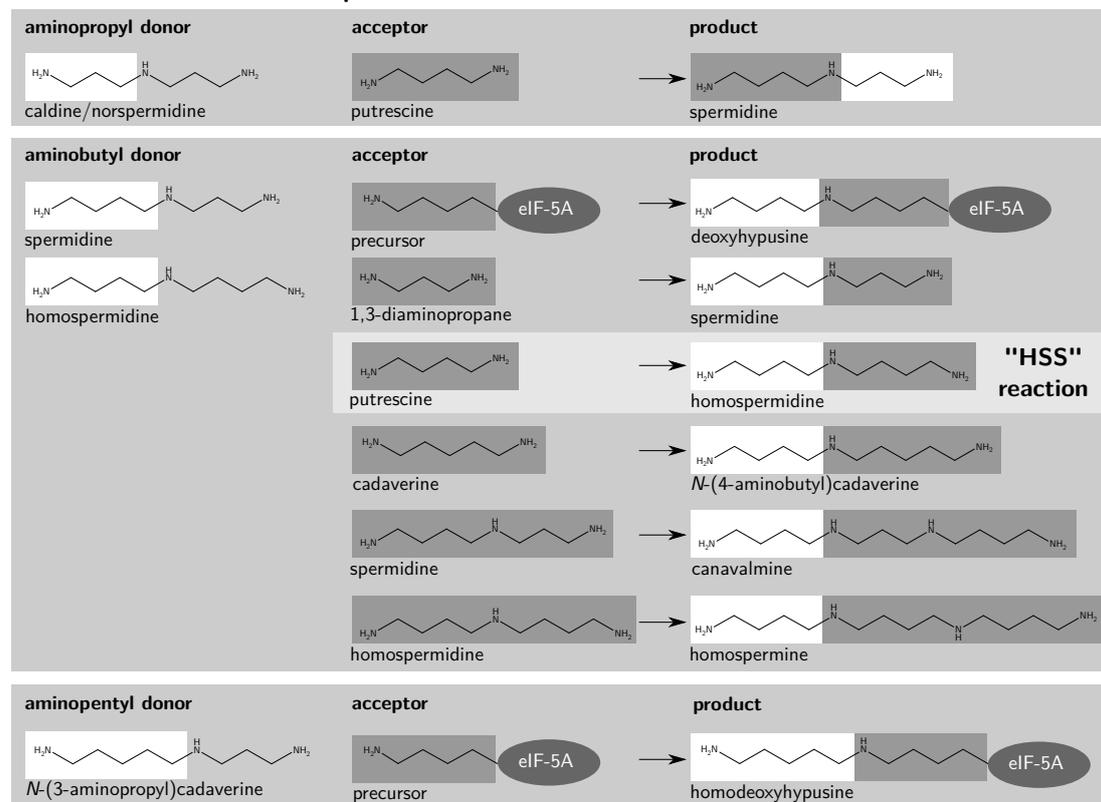


Figure 4.8.: Overview of DHS amino donor/acceptor substrates and their reaction products

reaction and accept the aminobutyl moiety from the enzyme–imine intermediate to form homospermidine. This homospermidine synthase activity had already been described for DHS from tobacco and *Senecio vernalis* and will be described in more detail in section 4.3 [189, 190, 192].

Further experiments revealed that DHS accepts a small number of spermidine related compounds as amine donors for the forward reaction, namely spermidine, homospermidine, caldine, *N*-(3-aminopropyl)cadaverine, *cis*- and *trans*-spermidine, 1- or 8-methylspermidine and *N*¹- or *N*⁸-ethylspermidine – although the methyl- and ethyl-derivatives were very inefficient donors (see figure 4.8 for an overview of substrates and reaction products). As acceptors can act either the eIF-5A precursor, or the short diamines 1,3-diaminopropane, putrescine or cadaverine (1,5-diaminopentane) – but also spermidine and homospermidine themselves. While spermidine and homospermidine as donors both generate either deoxyhypusine on the eIF-5A precursor, or spermidine, homospermidine or *N*-(4-aminobutyl)cadaverine from the corresponding short diamines – or even canavalmine or homospermine from spermidine and homospermidine, respectively –, the aminopropyl moiety of caldine can only be transferred onto putrescine to yield spermidine, and in contrast, the aminopentyl moiety of

N-(3-aminopropyl)cadaverine can only be transferred onto eIF-5A precursor, resulting in the formation of homodeoxyhypusine (*N*^ε-(5-aminopentyl)lysine) [201, 192].

The above-mentioned inhibiting effect of homospermidine on deoxyhypusine formation in a mixture of homospermidine and spermidine can therefore be attributed to the fact that homospermidine can act as aminobutyl donor – releasing putrescine, which in turn can act as aminobutyl acceptor, yielding homospermidine again. In contrast, caldine has a different mode of action: While it can act as donor for the acceptor putrescine – confirming its ability to form the enzyme–imine intermediate –, eIF-5A is not able to accept its aminopropyl moiety.

4.3. DHS paralog homospermidine synthase in angiosperm species: trailhead of the pyrrolizidine alkaloid biosynthesis pathway

As mentioned in the previous section, DHS can transfer its enzyme-bound aminobutyl moiety derived from spermidine (or homospermidine) not only onto its main acceptor eIF-5A, but also in a less efficient side reaction onto putrescine, resulting in the formation of homospermidine (see figure 4.8 "HSS" reaction). This side reaction was first described for DHS from tobacco, but was not restricted to plant DHS, as subsequent experiments revealed (see previous section) [189, 190, 192, 201].

Indeed, this discovery was able to unravel the origin of the enzyme homospermidine synthase (HSS, EC 2.5.1.45) in a variety of plant species. HSS was identified as the first pathway-specific enzyme of the pyrrolizidine alkaloid (PA) biosynthesis in more than 6000 species belonging to several angiosperm families – mainly the Apocynaceae, Asteraceae, Boraginaceae, Fabaceae and Orchidaceae. PAs are a diverse group of constitutive plant secondary metabolites, which are mono- or diesters of mostly aliphatic mono- or dicarboxylic acids (necine acids) with a necine base – a derivate of the eponymous pyrrolizidine, like the widespread retronecine ((1*R*,7*aR*)-7-(hydroxymethyl)-2,3,5,7*a*-tetrahydro-1*H*-pyrrolizin-1-ol, see figure 4.9). In many species, PAs are synthesized as defense compounds against herbivores. However, many insects co-evolutionarily developed strategies to detoxify and store these compounds, and utilize them for their own defense or the production of pheromones [157].

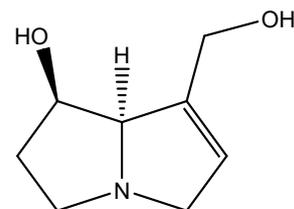


Figure 4.9.: Retronecine

Basic compound for the biosynthesis of the PAs' necine base is homospermidine, which is generated from the polyamines spermidine and putrescine by HSS. The discovery of the homospermidine producing side reaction in DHS, sequence comparisons and the distribution of HSS and DHS coding sequences in plants confirmed that HSS evolved from DHS through gene duplication and subsequent specialization of one duplicate on the HSS side reactivity (see section A.5 figure A.16 for an alignment of DHS and HSS amino acid sequences from representatives of the above-mentioned angiosperm families) [189, 190]. In the course of this specialization, the duplicate DHS lost its ability to bind and modify the eIF-5A precursor due to non-synonymous mutations in one of the duplicated genes, resulting in a separate HSS – a so-called paralog [192]. As analyses of the properties of DHS and HSS from *Senecio ver-*

4.3. DHS paralog homospermidine synthase in angiosperm species: trailhead of the pyrrolizidine alkaloid biosynthesis pathway

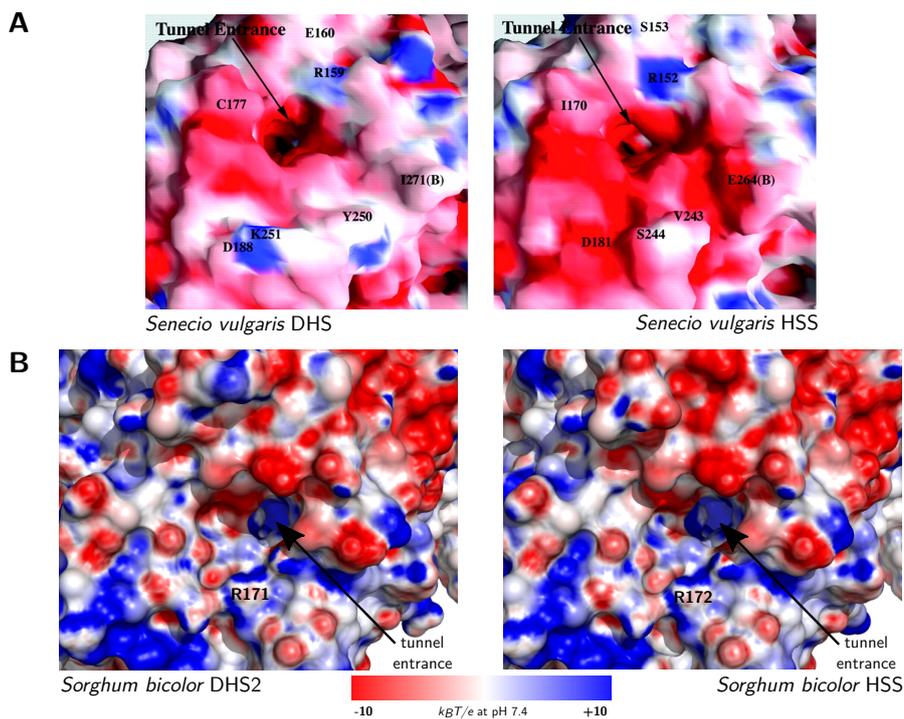


Figure 4.10.: Representation of the surface electrostatic potential of the active site tunnel entrance of *S. vulgaris* and *S. bicolor* DHS and HSS models

A: Surface representation of *S. vulgaris* DHS and HSS, colored according to the residues' electrostatic potential; from Umland et al. (2004) [280]. **B:** Surface representation of *S. bicolor* DHS2 and HSS (DHS3), colored according to the residues' electrostatic potential at pH 7.4. The *Sb*DHS2 and *Sb*HSS models were built based on the structure of human DHS (PDB Acc. No.: 1DHS) with the SWISS-MODEL webserver (<https://swissmodel.expasy.org/>). The electrostatic potential was calculated with the PDB2PQR webserver (version 2.1.1, http://nbcrc-222.ucsd.edu/pdb2pqr_2.1.1/, [71]) and APBS (version 1.4.1, [119]). The graphic representation was generated with an open-source build of PyMOL (version 2.1.0) with the APBS plugin and Inkscape™ (version 0.92, <https://inkscape.org/>).

nal demonstrated, this specialization does not necessarily alter reaction kinetics or substrate specificities apart from the loss of eIF-5A binding [192]. Recalling the mutation experiments performed on human DHS (see section 4.2.2), a mutation of D313, which abolished eIF-5A binding to human DHS, seems intriguing as reason for the loss of eIF-5A binding in HSS. However, this residue is also conserved in HSS [280]. Comparisons of the area around the active site funnel in a *Senecio vulgaris* HSS model, based on the crystal structure of human DHS, with that of DHS only revealed a number of non-conserved substitutions, which increase the overall negative electrostatic potential in the funnel area for HSS – including the disappearance of distinct, positively charged patches, which might be involved in eIF-5A binding specificity (see figure 4.10 A) [280]. However, models of a DHS and its HSS paralog from *Sorghum bicolor*, a model plant from the Poaceae family, whose DHS/HSS isoforms were studied in this thesis, did not exhibit similar distinct changes in their electrostatic surface po-

4.3. DHS paralog homospermidine synthase in angiosperm species: trailhead of the pyrrolizidine alkaloid biosynthesis pathway

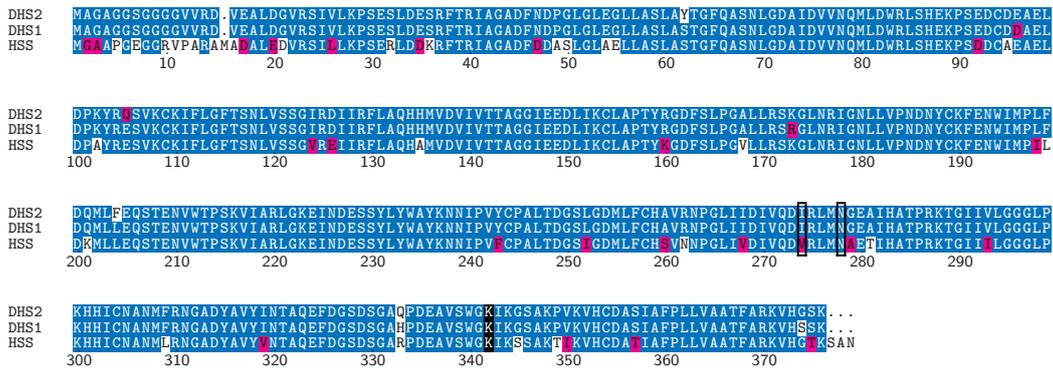


Figure 4.11.: Alignment of the amino acid sequences of two DHS homologs and their HSS paralog from *S. bicolor*

For the alignment of DHS homologs, the amino acid sequences from isolates¹ of DHS1 and DHS2 as well as HSS (DHS3) from *Sorghum bicolor* were used. □: non-matching residues; ■: similar residues in >50% of the aligned sequences; ■: conserved residues, ■: catalytic lysine. Boxed residues might represent telltale residues for the distinction between DHS and HSS [120, 155]. Residue numbering according to *SbDHS2*.

tential (see figure 4.10 B), although the HSS clearly exhibits HSS functionality (A. Wesseling and E. Kaltenecker¹, internal communication).

Since the presence of HSS is not uniform throughout angiosperm species, but scattered and incoherent, it was hypothesized that HSS was either developed several times independently during angiosperm evolution, or arose once in an early ancestor and was subsequently lost again several times in a variety of species. Phylogenetic studies showed that HSS – and the subsequent pyrrolizidine alkaloid biosynthesis pathway – developed at least eight times independently: twice in monocotyledons (Orchidaceae, Poaceae), twice in the Asteraceae family (Senecioneae and Eupatorieae), once in the Boraginaceae family, once in the Convolvulaceae family, once in the Apocynaceae family and once in the genus *Crotalaria* in the Fabaceae family. The evolution of specialized HSS from ubiquitous DHS is an excellent example for a rare evolutionary strategy, termed "evolution by change of function" [192, 230, 6, 188, 120, 289, 155].

Efforts have been made to distinguish between DHS and HSS based on the amino acid sequence. To date, two or three sequence positions have been implicated in defining the DHS or HSS characteristic of the paralogous enzymes from amino acid sequence analyses: 252, 263 and 267 (numbering based on human DHS) – of which the latter two seem to be more predictive [120, 155]. In all functionally characterized plant DHS, the residue at position 263 is an isoleucine, while HSS contains valine (or alanine in one exception) [155]. At position 267, plant DHS contains an asparagine, while HSS features an aspartic acid. Figure 4.11 exemplarily shows the amino acid sequence alignment of two DHS homologs and their HSS

¹Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

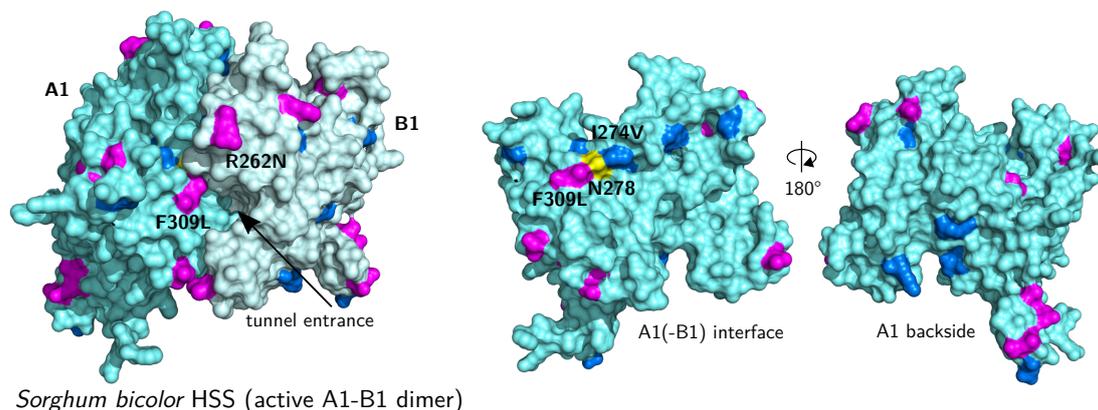


Figure 4.12.: Surface representation of the *S. bicolor* HSS model and its differences compared to the DHS model

Surface representation of the "active dimer" of the *S. bicolor* HSS (DHS3) and the separate A1 subunit. The amino acid changes compared to *SbDHS1* and *SbDHS2* are colored as follows: ■: unchanged residues; ■: similar residues; ■: different residues; ■: "telltale" N278. Residue numbering according to *SbDHS2*. The *SbHSS* model was built based on the structure of human DHS (PDB Acc. No.: 1DHS) with the SWISS-MODEL webserver (<https://swissmodel.expasy.org/>). The graphic representation was generated with an open-source build of PyMOL (version 2.1.0) and Inkscape™ (version 0.92, <https://inkscape.org/>).

paralog from *Sorghum bicolor*, the plant whose DHS and HSS were studied in this thesis. From these two telltale residues (274 and 278 in *SbDHS2*), only the isoleucine is changed to valine in the *S. bicolor* HSS. In figure 4.12, all amino acid changes between *SbDHS* and *SbHSS* are emphasized in a surface representation of the active dimer of an *SbHSS* model. It becomes clear that only very few residues might be responsible for the lack of eIF-5A binding in HSS enzymes. In this case, R262N and F309L are the only changes on the surface in the vicinity of the active site funnel. The telltale residues are located side by side, buried between the interfaces above the entrance of the active site tunnel. Since they are not accessible on the surface – at least in these models – they probably influence the interaction of the interfaces and thereby HSS reactivity.

Mutagenesis experiments with *Ipomoea neei* DHS indicated that the individual I263V substitution slightly improves HSS activity without affecting DHS activity, while N267D significantly reduced DHS activity, influencing the HSS activity only a little [120]. However, if these findings can be generalized remains to be proved. Interestingly, beside the independent appearance of HSS in the different plant families, recent analyses also indicate that HSS characteristics – in these studies pinpointed to the two variants described before – did also get lost again several times independently during evolution of the Aponaceae [155].

The different origins of independently evolved HSS are also reflected in their differing expression patterns in plant tissues. While DHS is ubiquitous – although at low level – in all eukaryotic cells, HSS is expressed only in distinct plant cell types or tissues, which vary for different species. In addition, HSS expression – and thereby PA production – often depends on the developmental stage of the respective tissue [175, 6, 7, 184, 134].

The exclusive expression in selected tissue might also be a possibility to prevent detrimental effects of the duplicated new DHS paralog – HSS – on DHS functionality in the majority of plant tissue. Since DHS is an obligate homotetramer, it depends on its distinct protein-protein interactions for functionality (see section 4.2.2). The interaction of paralogous HSS and DHS subunits in the tetramer could impair the modification of the eIF-5A precursor and thereby possibly a number of downstream cellular processes (see section 4.1). This cross-interaction between paralogous subunits and the formation of so-called paralogous heteromers influences the evolutionary fate of the duplicate gene, a process termed "paralog interference" [12, 121].

Directly after gene duplication, the paralogs will be under the same selective pressure as their ancestor. Subsequently, both paralogs can co-evolve by positive selection – maintaining and optimizing the original function of the complex – or a purifying selection prevents changes in both paralogs, since the deleterious mutation of one paralog might functionally impair not only its own homomeric complexes, but also the paralogous heteromers. The development of HSS from DHS is most likely based on the latter scenario, since DHS function relies on the assembly of the active site between subunits, their regulatory interactions and additionally on the specific binding of the substrate eIF-5A on the enzyme surface – probably also involving two subunits at a time [121]. To escape the purifying selection under these circumstances and to allow for the loss of DHS functionality and the specialization on the HSS side activity, the paralogs need to be separated from each other. This can be either by spatial separation due to differential expression – something that has already been described for HSS and DHS (see above) – or by a change in the interacting interfaces, preventing the paralogs from forming heteromers [121]. However, in the case of DHS and HSS, the latter route would be significantly complicated by the fact that the interfaces make up the active site, which is essential for DHS and HSS activity alike. Nevertheless, to date no proof for the interaction of HSS and DHS paralogs has been described.

4.4. Bacterial HSS and plant HSS – similar, but not the same

The evolution of an HSS from DHS is a unique feature of a variety of angiosperm species. However, in many proteobacteria – which include human pathogens such as *Legionella pneumophila* and *Pseudomonas aeruginosa* – a highly conserved HSS is the nexus of bacterial polyamine metabolism. Since this enzyme also catalyzes the formation of homospermidine in an NAD⁺-dependent manner and with a similar pH optimum [191, 299], it was not clear until recently, if these enzymes were evolutionarily related. Sequence comparisons and finally the determination of the crystal structure of the *Blastochloris viridis* HSS in our laboratory revealed that – although both enzymes share many features – they are not evolutionarily related [133]. While plant HSS – like DHS – is a homotetramer, which creates its active sites at the interface of two subunits, the bacterial HSS is a homodimer with a single, independent active site in each subunit. Plant HSS creates its product via a Schiff base intermediate at its catalytic lysine and cannot synthesize homospermidine from putrescine alone – in contrast to bacterial HSS, which has no imine intermediate, as confirmed by the absence of a lysine residue in the active site, and readily accepts putrescine as sole substrate [191].

Nevertheless, in view of the rising numbers of multidrug-resistant strains, the bacterial HSS and its homospermidine synthesis product, which is essential for bacterial growth and might play a crucial role in pathogenic mechanisms such as biofilm formation and the evasion or repression of the host immune system, could represent a valuable target for new therapeutic approaches [64, 254, 150, 183]. The recently solved crystal structure of a bacterial HSS now enables the targeted design of specific inhibitors of the bacterial HSS[133]. Further knowledge on the binding of eIF-5A to DHS from biophysical experiments and future crystal structures of the eIF-5A complex will probably support these efforts and provide a basis to minimize cross-interaction of putative therapeutic inhibitors.

4.5. Aim of this thesis

The (deoxy)hypusination of a specific lysine residue of the eIF-5A precursor is crucial for most of its function in cellular processes of eukaryotes and archaea. Without this modification, the cells can ultimately not survive. Consequently, the modifying deoxyhypusine synthase (DHS) – which catalyzes the first of two modification steps – is very well conserved in all eukaryotic cells and archaea. In several angiosperm plant lineages, gene duplication of this highly conserved enzyme gave rise to a paralogous enzyme, the homospermidine synthase (HSS), which subsequently specialized on a side reactivity of the original DHS enzyme, the synthesis of homospermidine from spermidine and putrescine. In the course of this specialization, the HSS enzyme lost its inherited DHS activity, and in many plant species developed into the trail head enzyme of the pyrrolizidine alkaloid biosynthesis pathway that produces toxic substances as defense against herbivores.

The DHS enzymes from several eukaryotic species have already been studied for some time. Many details about the mechanism and dynamics of the deoxyhypusination reaction, the involved substrates and cofactors and their interdependence have been elucidated. For the angiosperm HSS, however, information is comparatively scarce concerning the defining differences that separate the HSS from the parental DHS, and that give rise to the first enzyme of a complete biosynthesis pathway. Although the structure of human DHS and the eIF-5A precursor from several species have been determined to date, structural data on an HSS enzyme – or a complex between a DHS and the eIF-5A precursor – is still not available.

In this thesis, several different biophysical techniques were applied to gain a deeper understanding of the differences and common features of two DHS homologs and their HSS paralog from the model plant *Sorghum bicolor*. *S. bicolor* is an example for a species containing a functional HSS besides two DHS isoforms, and for which no PA production has been shown to date. The occurrence of two functional DHS isoforms is somewhat special, since DHS is mostly maintained as a single copy in eukaryotes. It is therefore likely that the second duplication is a quite recent event – in evolutionary terms – and that *S. bicolor* enables the study of two different stages on the evolutionary pathway from DHS to HSS functionality. In addition, the biophysical data produced in the course of this thesis is expected to facilitate future efforts to obtain structural data for the individual DHS and HSS enzymes and the DHS–eIF-5A complex.

5. Results

5.1. Optimization of *SbDHS* isoform purification

5.1.1. Scale-up of the *E. coli* expression and purification of His₆-tagged *SbDHS* isoforms

The heterologous expression in *E. coli* and subsequent purification of plant DHS and HSS for the determination of enzyme activity and kinetic characteristics has been performed in the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober (Botanical Institute of Kiel University) for several years and an established protocol is applied. For the experiments of this thesis, three expression constructs containing the three *Sorghum bicolor* DHS isoforms as C-terminally His₆-tagged fusion constructs were kindly provided by the research group of Prof. Dr. Dietrich Ober (see figure 5.1 **A**).

Initial small scale expression and purification trials with the established expression and purification protocols were successful (data not shown). However, since the purified protein was intended for biophysical characterization methods and crystallization for the structure determination by X-ray diffraction, the expression had to be scaled up to supply the necessary protein amounts.

Unfortunately, the scale-up from the original protocol with 500 mL LB culture volume in unbaffled 1 L shaking flasks (induction at an OD₆₀₀ of 0.4 with 0.4 mM IPTG and expression for 4 h at 37 °C) to 2 L culture volume in plain 5 L shaking flasks not only resulted in a higher protein yield, but also in a huge loss of expressed protein at an established 1.5 M NaCl wash step during His₆-tag-based immobilized metal affinity chromatography (IMAC). Figure 5.2 **A** shows an exemplary IMAC profile of the purification of *SbDHS2*-His₆ from an up-scaled *E. coli* expression culture. After the initial loading of the cleared lysate (see figure 5.2 **B** sample S) onto a 5 mL HisTrap HP at an ÄKTA chromatography system (**A**: volume 0–63 mL), the column is washed with 3 CV of a buffer containing 1.5 M NaCl to reduce unspecific binding

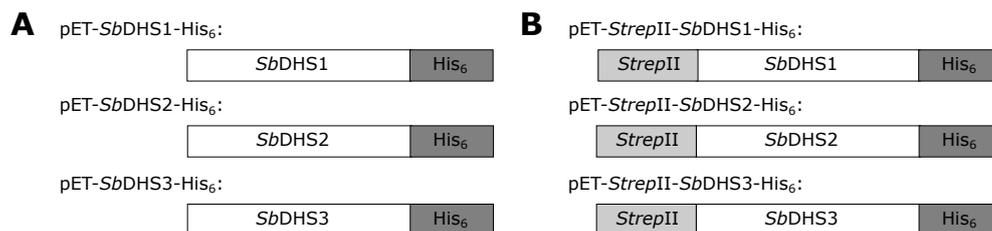


Figure 5.1.: Constructs for expression of *SbDHS* isoforms in *E. coli*

A: Full-length *SbDHS* isoform 1, 2 and 3 fused to a C-terminal His₆-tag. **B:** Full-length *SbDHS* isoform 1, 2 or 3 preceded by an N-terminal *StrepII*-tag and C-terminally fused to a His₆-tag. Constructs not to scale.

5.1. Optimization of *SbDHS* isoform purification

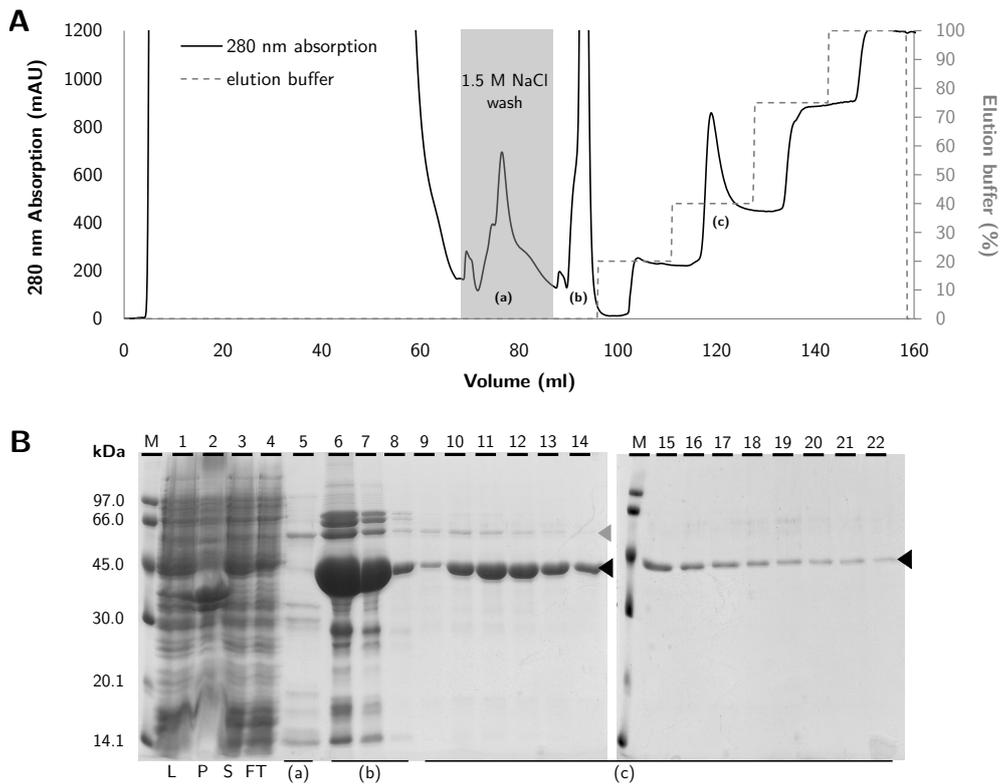


Figure 5.2.: IMAC profile and SDS-PAGE result of *SbDHS2-His₆* purification according to the original protocol from up-scaled *E. coli* expression culture

A: IMAC profile of the purification of *SbDHS2-His₆* from up-scaled *E. coli* expression culture based on the established protocol. Purification was performed on cleared lysate with a 5 mL HisTrap HP (GE Healthcare) at an ÄKTA chromatography system (GE Healthcare). **B:** SDS-PAGE results from selected samples of the purification from **A**. L: lysate, P: pellet; S: supernatant; FT: flow-through; (a): elution fraction of the 1.5 M NaCl wash step; (b): elution fraction after wash step; (c): stepwise elution from 20–500 mM imidazole. Black arrowhead indicates the band of the *SbDHS2-His₆* subunit with a calculated molecular weight of 41.8 kDa; gray arrowhead indicates the putative band of the bacterial chaperone GroEL. M: molecular weight marker (MWM; Low Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare).

(**A** (a): volume 63–82 mL). While only minor impurities were eluted during this step (see **B** lane 5), a huge amount of protein is eluted from the column in about 1 CV upon buffer exchange back to the original binding buffer with only 300 mM NaCl (see **A** (b) and **B** lanes 6–8). Based on the amount and the apparent molecular weight of about 45 kDa of the major protein fraction in the eluate, which correlates with the expected molecular weight of 41.8 kDa of the expressed fusion protein, it could be assumed that a significant portion of the bound *SbDHS2-His₆* was lost at this step. A smaller portion of the His₆-tagged protein stayed stably bound to the matrix of the IMAC column and was not eluted until an imidazole concentration between 120 and 220 mM was reached.

In the established protocol, no protein was lost at this wash step. The overall protein yield from the elution step was only approx. 1 mg in total from 4 L expression culture – which is about the same amount that could be purified from 1 L expression culture of the original scale (data not shown). Since all protein initially bound to the IMAC column – which is supposed to retain about 40 mg His₆-tagged protein per mL resin, according to manufacturer's information – and the finally eluted protein amount was so low, an overloading of the 5 mL HisTrap HP column could be ruled out, even if steric hindrance of the tetramers would be taken into account.

Although the omission of the 1.5 M NaCl wash step prevented the loss of protein at this step and enabled a regular affinity purification (data not shown), the distinct difference in binding behavior indicated an inhomogeneous situation in the His₆-tagged protein pool. Since large amounts of homogeneous protein are a prerequisite especially for the protein crystallization for structure determination purposes, this quality discrepancy needed to be addressed prior to further biophysical characterization experiments.

Several parameters during the expression and purification of *SbDHS* isoforms were therefore examined and optimized in the course of this thesis, as summarized in the following section.

5.1.2. Optimization of expression and purification conditions

The inhomogeneous situation in the pool of the expressed *SbDHS2*-His₆ must be due to differences between the tetramers or even single subunits, that influence the binding behavior of the His₆-tag. Since the His₆-tag is attached to the C-terminus of *SbDHS2*, the formation of truncation fragments missing the affinity tag from abortive translation might result in tetramers with one or more subunits without a His₆-tag, which could exhibit a different binding behavior than tetramers which contain a His₆-tag at each subunit.

In addition, the SDS-PAGE results from several purification experiments often contained a band at about 60 kDa even in the final purified protein pool (also visible in figure 5.2 B, especially in lanes 6 and 7), which might represent the bacterial chaperone GroEL and could be indicative for difficulties with the proper folding of the overexpressed protein.

A codon usage analysis of the *SbDHS* sequences in correlation to the codon usage of *E. coli* B (see section A.4 figures A.11, A.10 and A.9) revealed several critical codons. Therefore, instead of BL21 (DE3), the strain RosettaTM 2(DE3), which provides extra tRNAs for the most critical codons in *E. coli* on a separate plasmid, was tested for expression. Additionally, the expression temperature was lowered to 25 °C with an elongated expression period of about 22 h. However, these modifications did not significantly change the binding behavior of the His₆-tagged *SbDHS2* in the subsequent purification procedure.

To ensure optimal conditions during *E. coli* cell lysis and the following purification, different buffer conditions were tested with purified *SbDHS2*-His₆ and evaluated by differential scanning fluorimetry (DSF, data not shown). Therefore, protein samples in a variety of buffer conditions were heated in the presence of the fluorescent dye SYPRO Orange[®]. Upon protein unfolding, the dye binds to exposed hydrophobic patches and starts to fluoresce. The fluorescence signal can then be used to generate a melting curve and calculate a melting temperature for the corresponding protein. Substances or conditions that enhance protein stability result in an increase of the protein's melting temperature (see section 7.14.5 for measurement details).

figure 5.12 in section 5.4.1 exemplarily shows the melting curves of *StrepII-SbDHS1-His₆* in 11 different buffers at three pH values per buffer. Especially for the more extreme pH values, a distinct shift in the melting curves to lower temperatures is visible, which results in a lower melting temperature and indicates impaired stability under the corresponding conditions. However, the results of these preliminary experiments only confirmed that the three *SbDHS* isoforms were stable over a broad pH range around the physiological pH (from about pH 5.5 to 9.0), and without a pronounced buffer preference. Comparable results were generated later in the course of this thesis with *StrepII*- and *His₆*-tagged *SbDHS* samples. They are depicted in figure 5.13 **A** and described in the corresponding section.

The original purification protocol was based on a phosphate buffer (50 mM NaH₂PO₄, 300 mM NaCl, 5 mM β -mercaptoethanol, 20 mM imidazole, pH 8.0) for lysis and IMAC, and a glycine-NaOH buffer (20 mM glycine, 200 mM KCl, 1 mM DTT, pH 9.0) for the subsequent size exclusion chromatography (SEC). Based on the thermal stability results from DSF experiments, the purification protocol was initially simplified by implementation of 50 mM Tris-HCl buffer at pH 7.5 for IMAC as well as SEC instead of the phosphate and glycine-NaOH buffer, although no significant positive effect on sample homogeneity was achieved. Further experiments with buffers containing Mg²⁺ and ATP to help GroEL-like chaperones release their substrates did also not improve the situation (data not shown).

In a parallel attempt to try a different affinity purification approach, the original C-terminally *His₆*-tagged *SbDHS* constructs were modified with an additional N-terminal *StrepII*-tag (see figure 5.1 for a schematic representation of the constructs and section 7.9.3 for cloning details). This enabled highly specific affinity purification on *Strep*-Tactin[®] resin. Interestingly, the N-terminal *StrepII*-tag significantly enhanced the expression of the three *SbDHS* in *E. coli* and thereby increased the overall protein yield even after *His₆*-tag based IMAC (data not shown).

The quality of a purified protein charge is routinely evaluated prior to flash freezing and storage and after thawing by dynamic light scattering (DLS) measurements. DLS is a biophysical method that determines the apparent hydrodynamic diameter of particles in solution or suspension – which move randomly due to Brownian motion – by autocorrelating the intensity fluctuation of their scattered light measured at a fixed angle (see section 7.14.6 for further experimental and theoretical details). From these autocorrelation functions, a distribution of particle sizes in the given solution or suspension is derived.

DLS measurements of freshly prepared *SbDHS2-His₆*, which was purified by IMAC and SEC, exhibited a nearly monodisperse particle size distribution (21.4 % polydispersity) around an average diameter of about 11.1 nm (see figure 5.3 **A**, gray line). This corresponds to an estimated protein molecular weight of approx. 184 kDa, which is in accordance with the expected molecular weight of the corresponding active tetramer of 167.2 kDa. In contrast, after flash freezing and storage at –80 °C, the DLS measurements of a thawed *SbDHS2-His₆* sample show a distinct shift in the size distributions towards a larger average diameter of about 18.6 nm, which corresponds to an estimated protein molecular weight of approx. 624 kDa. Interestingly, the polydispersity (21.6 %) did not change significantly, indicating that the size distribution of the aggregated particles is still quite narrow. This could be due to a well defined higher-order oligomeric state. In this case, a tetramer of the active tetramer with an

5.1. Optimization of *SbDHS* isoform purification

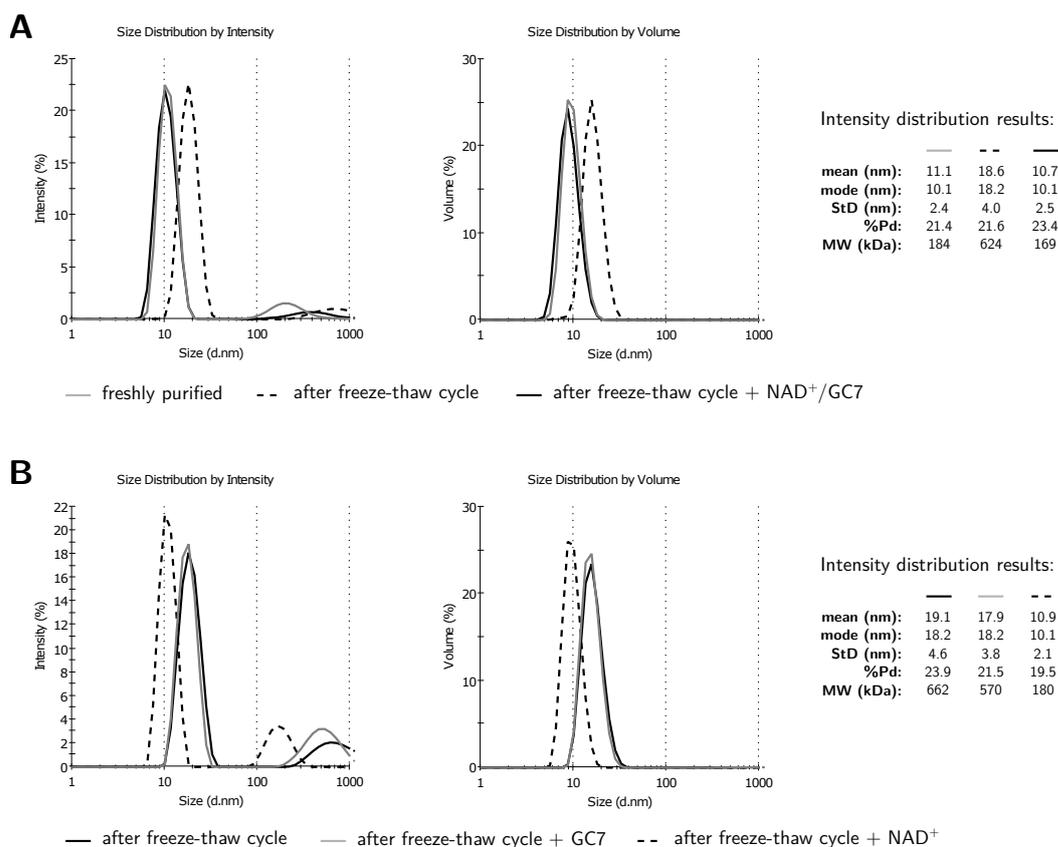


Figure 5.3.: DLS results of *SbDHS2-His₆* prior to and after flash freezing/thawing, and the effect of NAD⁺ and GC7

Intensity- and volume-based size distributions of *SbDHS2-His₆* samples at a concentration of 2 g L⁻¹. Intensity distribution results are only reported for the main peaks. **A:** Gray line: fresh *SbDHS2-His₆* after IMAC and SEC purification; average of 50 DLS measurements (13 runs of 10 s each). Black dotted line: *SbDHS2-His₆* after IMAC, SEC and a flash freeze/thaw cycle; average of 30 DLS measurements (10 runs of 10 s each). Black line: *SbDHS2-His₆* after IMAC, SEC, a flash freeze/thaw cycle and the addition of 2 mM NAD⁺ and GC7; average of 50 DLS measurements (20 runs of 10 s each). **B:** Black line: *SbDHS2-His₆* after IMAC, SEC and a flash freeze/thaw cycle. Gray line: *SbDHS2-His₆* after IMAC, SEC, a flash freeze/thaw cycle and the addition of 2 mM GC7. Black dotted line: *SbDHS2-His₆* after IMAC, SEC, a flash freeze/thaw cycle and the addition of 2 mM NAD⁺. Average of 6 DLS measurements (15 runs of 10 s each) for each sample. StD: standard deviation; Pd: polydispersity; MW: molecular weight.

estimated molecular weight of approx. 668 kDa would fit the reported average particle size best.

DHS needs NAD⁺ as cofactor and spermidine as substrate for the catalyzed deoxyhypusination reaction on its protein substrate eIF-5A. The binding of NAD⁺ and spermidine to DHS is independent of the binding of eIF-5A. The binding of spermidine (or the potent inhibitor GC7), however, depends on the prior binding of NAD⁺ (see section 4.2.3). Since at least the binding of NAD⁺ is assumed to initiate a conformational adaptation of the active tetramer,

it seemed reasonable to evaluate the effects of NAD^+ and GC7 – supplied separately or in combination – on the oligomeric situation in the thawed *SbDHS2*-His₆ samples with further DLS measurements.

In figure 5.3 **A**, the black line represents the particle size distribution of a thawed *SbDHS2*-His₆ sample supplied with 2 mM NAD^+ and an equal amount of GC7. Interestingly, directly upon addition of cofactor and inhibitor, the distribution is shifted towards a lower average diameter. It becomes nearly indistinguishable from the particle size distribution of freshly purified *SbDHS2*-His₆ (**A**, gray line) with an average particle diameter of 10.7 nm and a polydispersity of 23.4%. The average diameter corresponds to an estimated protein molecular weight of approx. 169 kDa, which is again in the range of the expected molecular weight of the active DHS tetramer.

The individual effect of NAD^+ and GC7 on the thawed *SbDHS2*-His₆ sample was tested in a second experiment and is depicted in figure 5.3 **B**. The black line represents the situation after a freeze/thaw cycle. Similar to what was already shown in **A**, the particle sizes are distributed around an average diameter of 19.1 nm with a polydispersity of 23.9%, which corresponds to an estimated protein molecular weight of 662 kDa. Upon addition of 2 mM NAD^+ (**B** black dotted line), the distribution is immediately shifted to the average particle diameter of the freshly purified tetramer (10.9 nm, 19.5% polydispersity). In contrast, the addition of 2 mM GC7 to the thawed *SbDHS2*-His₆ sample did not result in such a distinct effect: the average particle diameter is slightly lower (17.9 nm, 21.5% polydispersity), leading only to a small decrease in the calculated protein molecular weight (570 kDa). Taking into account the fact that GC7 requires prior binding of NAD^+ for its own binding to DHS, this minor effect of separately supplied GC7 compared to that of separately added NAD^+ is not unexpected.

5.1.3. Optimized purification strategy

As the DLS measurement results described above clearly demonstrate, the presence of NAD^+ can have a significant effect on the aggregation state of *SbDHS2*. With these findings in mind, the inhomogeneous binding behavior of *SbDHS2*-His₆ upon an established high salt wash step during IMAC (see section 5.1.1) was reviewed. Since the heterologous expression of the DHS constructs in *E. coli* was aimed at and optimized for a high protein yield, it might be possible that *E. coli* was not able to supply the DHS tetramers with a sufficient amount of NAD^+ to prevent aggregation upon lysis and purification. This could lead to the formation of higher order oligomers – as demonstrated for thawed *SbDHS2*-His₆ samples. These aggregates might impair the accessibility of the His₆-tag, resulting in a less efficient binding of the aggregates to the IMAC column resin.

Based on this hypothesis, the original purification protocol was modified as follows: The cells were lysed in a buffer supplied with 2 to 5 mM NAD^+ (final buffer composition: 100 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 2–5 mM NAD^+ , 0.5–1 mM PMSF, pH 8.5 at 22 °C). The addition of the reducing agent was omitted in all used buffers to maintain the oxidized state of NAD^+ .

The soluble His₆-tagged *SbDHS* was subsequently affinity purified from the cleared lysate in the same buffer, with a slightly lower NAD^+ concentration of 1 to 2 mM (and without

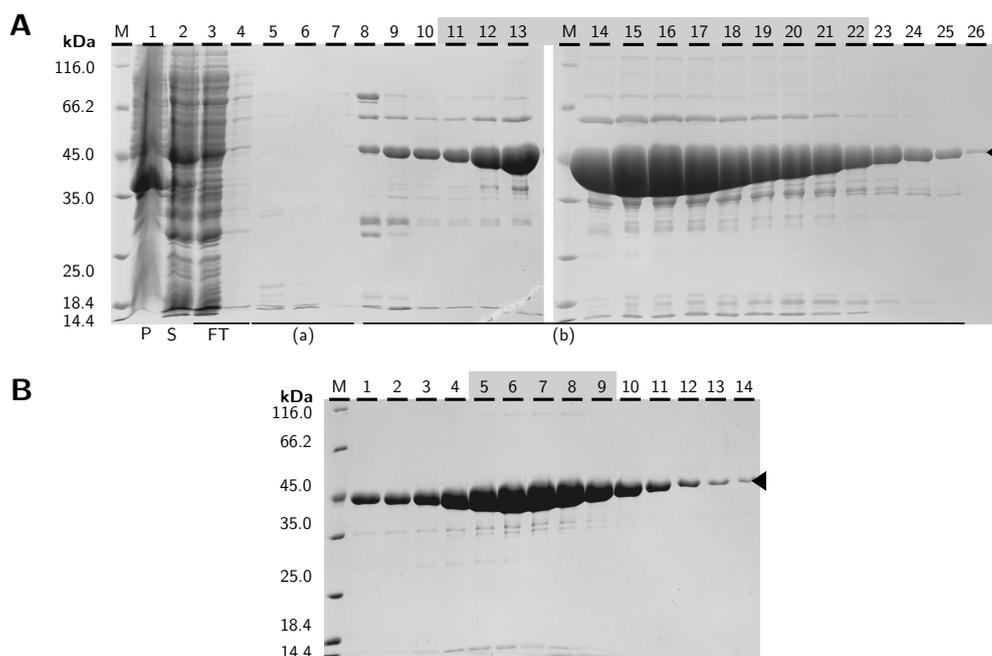


Figure 5.4.: SDS-PAGE results of *StrepII-SbDHS2-His₆* purification according to the optimized protocol from up-scaled *E. coli* expression culture

A: SDS-PAGE result from selected samples of the optimized IMAC purification of *StrepII-SbDHS2-His₆*. P: pellet; S: supernatant; FT: flow-through; (a): elution fraction of 1.0 M NaCl wash step; (b): gradient elution from 68–350 mM imidazole. **B:** SDS-PAGE result from selected samples of the optimized SEC purification of *StrepII-SbDHS2-His₆* from retention volume 58–71 mL (Superdex 200 HiLoad 16/60, GE Healthcare). Black arrowhead indicates the band of the *StrepII-SbDHS2-His₆* subunit with a calculated molecular weight of 43.6 kDa. Gray boxes: pooled fractions of IMAC and SEC, respectively. M: molecular weight marker (MWM; Pierce™ Unstained Protein MW Marker, Thermo Scientific).

protease inhibitor PMSF). Included in the purification procedure was the high salt wash step from the original protocol. The protein from the pooled and concentrated elution fractions was then polished by SEC in a glycine-NaOH buffer at pH 9.0 to maintain conditions that were similar to those of the smaller scale expressions and purifications, and ensure comparability of subsequent activity results (final buffer composition: 100 mM glycine-NaOH, 150 mM NaCl, pH 9.0 at 22 °C). While the SEC running buffer at first also contained 1 to 2 mM NAD⁺ as a preventive measure, purification trials finally revealed that NAD⁺ supplementation was not necessary at this point (data not shown). In sections 7.12.2 and 7.12.2 the complete optimized purification protocol is described in more detail.

Due to the presence of NAD⁺ in the buffers, the 280 nm absorption signal during the IMAC purification at an ÄKTA chromatography system (GE Healthcare) is usually very noisy. The depiction of the profile of an IMAC purification with optimized buffers is therefore omitted. In figure 5.4, however, the SDS-PAGE results of the IMAC (**A**) and SEC (**B**) purification of *StrepII-SbDHS2-His₆* expressed in Rosetta™ 2(DE3) and purified according to the optimized

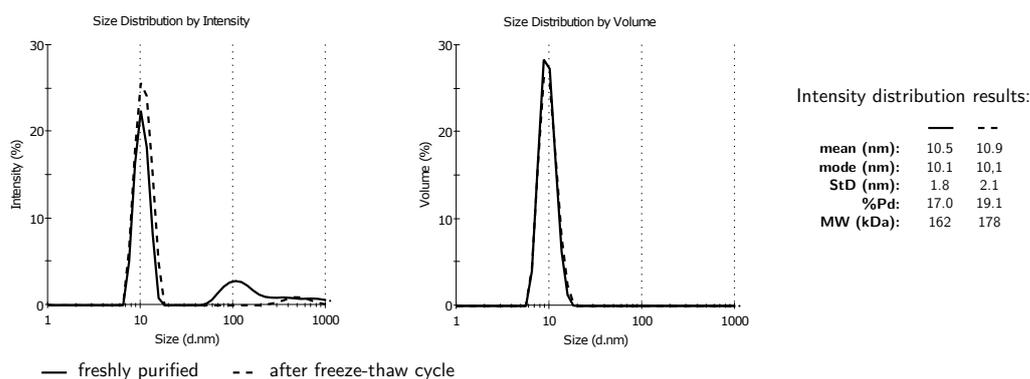


Figure 5.5.: DLS results of *StreptII-SbDHS2-His₆* purified according to optimized protocol prior to and after flash freezing/thawing

Intensity- and volume-based size distributions of *StreptII-SbDHS2-His₆* samples at a concentration of 1.5–2.0 g L⁻¹. Intensity distribution results are only reported for the main peaks. Black line: fresh *StreptII-SbDHS2-His₆* after optimized IMAC and SEC purification; average of 5 DLS measurements (20 runs of 10 s each). Black dotted line: *StreptII-SbDHS2-His₆* after optimized IMAC, SEC and a flash freeze/thaw cycle; average of 5 DLS measurements (20 runs of 10 s each). StD: standard deviation; Pd: polydispersity; MW: molecular weight.

purification protocol are shown exemplarily. As clearly visible, the high salt wash step (A (a)) did not cause protein loss and an ample amount of fusion protein could be eluted in the imidazole gradient, including several visible impurities. The subsequent SEC (see 5.4 B) – or the application of a second affinity purification step based on the *StreptII*-tag in combination with a *Strep-Tactin*[®] column (data not shown) – significantly reduced these impurities to a negligible level. Since the affinity purification via *StreptII*-tag is inconvenient for larger protein amounts due to the expensive desthiobiotin used for elution, the relatively low binding capacity of the *Strep-Tactin*[®] and the limited regeneration cycle number (a maximum of 5 cycles is recommended by the manufacturer), the combination of IMAC and SEC was maintained in the optimized protocol, even for additionally *StreptII*-tagged proteins.

In figure 5.5, the results of DLS measurements prior to and after flash freezing/thawing of the pooled *StreptII-SbDHS2-His₆* fractions from the SEC of figure 5.4 B (gray boxed lanes) are shown. Samples directly after SEC purification (black line) and after thawing (black dotted line) exhibit almost identical particle size distributions around and average diameter of 10.5 and 10.9 nm with a polydispersity of 17.0 and 19.1 %, respectively. The estimated protein molecular weights of 162 and 178 kDa are well in the range of that of the active tetramer (174.4 kDa). This demonstrates that the supplementation of NAD⁺ during purification not only homogenizes the protein binding behavior, but also safeguards the protein from aggregation upon freezing/thawing.

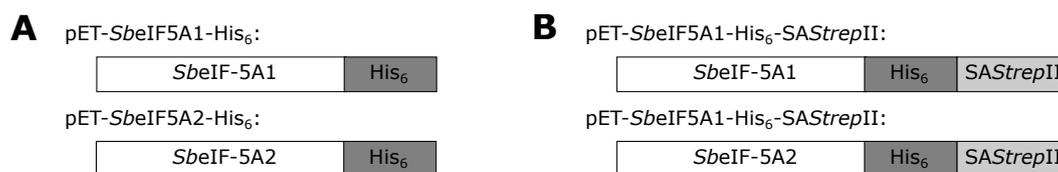


Figure 5.6.: Constructs for expression of *SbeIF*-5A isoforms in *E. coli*

SbeIF-5A isoform 1 or 2 fused to a C-terminal His₆-tag. Constructs not to scale.

5.2. Heterologous production of *Sorghum bicolor* DHS substrate eIF-5A

5.2.1. Design and cloning of *SbeIF*-5A constructs

For the biophysical characterization of the interaction of the DHS with its substrate protein eIF-5A in the course of this thesis, heterologously expressed, authentic *Sorghum bicolor* eIF-5A was chosen over *Zea mays* eIF-5A, which is routinely used in plant DHS activity assays in the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober (Botanical Institute of Kiel University), to represent the native situation. In *S. bicolor*, two distinct eIF-5A-like coding sequences had been identified in 2015, when the expression constructs were designed and cloned (GenBank Acc. No.: KXG34270.1, termed "*SbeIF*-5A1"; GenBank Acc. No.: KXG37504.1, termed "*SbeIF*-5A2"). Recently in 2017, a third eIF-5A-like coding sequence was annotated (GenBank Acc. No.: KXG36613.1, termed "*SbeIF*-5A3"). In section A.5 figure A.13, these three protein sequences are aligned. As expected, they exhibit a very high sequence identity, especially in the N-terminal domain.

For the DHS activity assays in the research group of Prof. Dr. Dietrich Ober, a C-terminally His₆-tagged *Z. mays* eIF-5A construct is used. Since the position of the His₆-tag at the C-terminus demonstrably did not interfere with the deoxyhypusination reaction on lysine 50 in the N-terminal domain of eIF-5A, the already existing *Z. mays* construct was used as backbone. The coding sequence for *ZmelF*-5A was replaced with the coding sequence for either of the two *SbeIF*-5A by an RF-Cloning strategy. The necessary primers were designed based on the database sequences indicated above (see section 7.9.1 for cloning details). The raw material for the PCR-based amplification of *S. bicolor* eIF-5A coding sequences was derived from fresh *S. bicolor* leaves (Botanical Garden, Kiel University) by RNA extraction and subsequent reverse transcription into cDNA.

Besides the C-terminally His₆-tagged eIF-5A isoforms, constructs with an additional C-terminal serine-alanine-*StrepII*-tag for the specific and stable binding to special antibodies for binding kinetic experiments (see section 5.5) were generated for both *SbeIF*-5A homologs (see section 7.9.2 for cloning details). In figure 5.6, the constructs are depicted schematically. The sequence integrity of the final constructs was verified by sequencing (data not shown).

5.2.2. *E. coli* expression and purification of *SbeIF*-5A constructs

The expression and purification of the *SbeIF*-5A constructs described in the previous section was performed based on established protocols for the purification of the *ZmelF*-5A construct

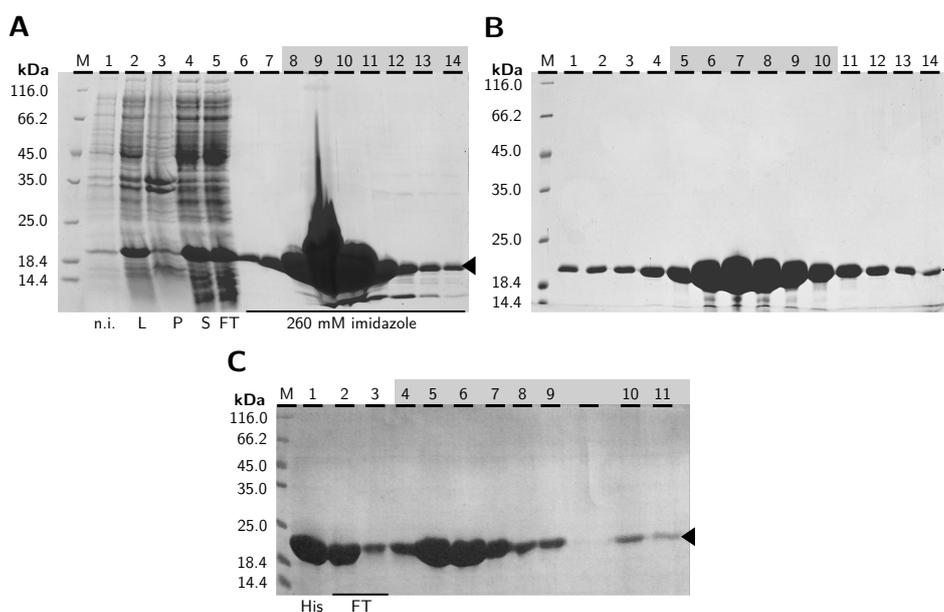


Figure 5.7.: SDS-PAGE results of *SbelF*-5A1 IMAC, SEC and *Strep*-Tactin[®] purifications

A: SDS-PAGE result from selected samples of the IMAC purification of *SbelF*-5A1- His_6 . n.i.: sample prior to induction; P: pellet; S: supernatant; FT: flow-through. **B:** SDS-PAGE result from selected samples of the SEC purification of *SbelF*-5A1- His_6 from retention volume 56–82 mL (Superdex 75 HiLoad 16/60, GE Healthcare). **C:** SDS-PAGE result from selected samples of the *Strep*-Tactin[®] purification of *SbelF*-5A1- His_6 -SA*Strep*II. His: pooled fraction from prior IMAC (data not shown); FT: flowthrough. Black arrowheads indicate the band of *SbelF*-5A1- His_6 or *SbelF*-5A1- His_6 -SA*Strep*II with a calculated molecular weight of 18.5 and 19.7 kDa, respectively. Gray boxes: pooled fractions. M: molecular weight marker (MWM; Pierce[™] Unstained Protein MW Marker, Thermo Scientific).

with some minor modifications. The detailed expression and purification strategies are described in sections 7.11.1 and 7.11.2.

The exemplary results of the His_6 -tag-based IMAC with a 5 mL HisTrap HP (GE Healthcare) and subsequent SEC purification of *SbelF*-5A1- His_6 are depicted in figure 5.7. The heterologous expression produced considerable amounts of fusion protein, as the heavily overloaded elution fractions and the sizable amount of fusion protein still present in the flow-through fraction demonstrates. After polishing of the pooled fractions of the IMAC purification (**A**, gray boxed lanes) by SEC, the pooled *SbelF*-5A1- His_6 fractions (**B**, gray boxed lanes) only contain minor impurities, which are negligible compared to the total protein amount. The IMAC and SEC purification of *SbelF*-5A2- His_6 gave comparable results (data not shown).

The *SbelF*-5A- His_6 -SA*Strep*II variants were polished by affinity chromatography with a *Strep*-Tactin[®] column after an IMAC purification step (see section 7.11.2 for further details). Figure 5.7 C shows the exemplary SDS-PAGE result of the *Strep*II-tag-based polishing step, following the IMAC purification of *SbelF*-5A1- His_6 -SA*Strep*II (data not shown). The final pooled fractions do not exhibit any visible impurities – however, in this case already the

IMAC eluate is of high purity (see lane 1, "His"). Comparable results were obtained for *SbelF-5A2-His₆-SAStrepII* (data not shown).

DLS measurements confirmed protein homogeneity with nearly monodisperse average particle size distributions around 5.9 nm (data not shown). Based on the assumption of a globular molecule, this diameter corresponds to a calculated molecular weight of approx. 42 kDa. Since the expected molecular weight of the different constructs and variants is in the range of 18.5 to 19.7 kDa, these values might indicate a dimer situation. However, DLS data analysis assumes all scattering molecules to be of globular shape, thereby only determining the size of a sphere that moves in the same manner as the scattering protein. Deviating from this assumption, the eIF-5A monomer does not feature a globular structure, but rather a quite elongated, maybe even dumb bell-like shape. This led in this case to an overestimation of the molecular weight calculated from DLS data, as later SEC-MALS measurements would reveal (see section 5.6).

Overall, the new *SbelF-5A* expression constructs provide high protein yields with an excellent purity upon purification with the strategy previously established for *ZmelF-5A* and a polishing SEC or *StrepII*-tag-based affinity chromatography step.

5.3. Crystallization of *SbDHS* and *SbelF-5A* isoforms

5.3.1. Classic crystallization approach

The three *S. bicolor* DHS isoforms characterized in this thesis only differ in a small number of amino acids (see section 4.3). However, these differences are sufficient to improve the homospermidine synthase (HSS) activity of *SbDHS3* and significantly impair the binding of the eIF-5A substrate protein compared to *SbDHS1* and *SbDHS2* (see section 5.5). To gain deeper insights into the molecular basis for this change of function, it is necessary to determine the detailed three-dimensional structure of the different isoforms – preferably in complex with the corresponding eIF-5A for the DHS paralogs. Although the structure of human DHS and several structures of eIF-5A from different species have already been solved, a structure of the complex has not been determined to date.

Despite recent advancements such as X-ray free-electron lasers (XFELs), X-ray diffraction of single, *in vitro*-grown protein crystals is still the gold-standard to obtain structural data of high-resolution. The crystallization process of proteins from solution is initiated by the careful supersaturation of the solution, which ideally results in the formation of a nucleation core and subsequently the growth of a single crystal. The slow supersaturation of the protein solution can be achieved by micro-dialysis, free-interface diffusion or vapor diffusion against a crystallization solution, which can contain a variety of chemical substances that might establish the necessary conditions for the crystallization of the protein in question. Unfortunately, each protein requires its very own, specific conditions, which encompass for example a certain pH value and buffer substance, the right precipitants and additives – and the optimal concentration of all components of the crystallization solution as well as the protein solution itself. Since the ideal conditions cannot be predicted beforehand, extensive screening of possible conditions has to be conducted.

In the course of this thesis, a number of crystallization screenings were performed to identify suitable conditions for the crystallization of the three *S. bicolor* DHS isoforms, the two eIF-5A and complexes thereof for subsequent X-ray crystallography. Most of these screenings were set up as sitting-drop vapor diffusion experiments in 96-well scale at the High Throughput Crystallization platform of The European Molecular Laboratory (EMBL) Hamburg and the High Throughput Crystallization Laboratory of EMBL Grenoble.

The separate IMAC- and SEC-purified, *N*-terminally *StreptII*-tagged *SbDHS* (see section 5.1) were all screened at concentrations of 5 or 8 g L⁻¹ in their storage buffer (100 mM glycine, 150 mM NaCl, pH 9.0) in the presence of 2 or 3 mM NAD⁺, or 2 or 3 mM NAD⁺ in combination with the same amount of inhibitor GC7. For the initial screening, the commercially available 96-well screens Morpheus[®] HT-96 I and II, Midas[®] HT-96, and Clear Strategy[™] HT-96 Screen I and II from Molecular Dimensions, as well as The PEGsII Suite, The JCSG+ Suite and The Protein Complex Suite from Qiagen[®] were applied. *StreptII-SbDHS2-His₆* was additionally screened with The Classics Suite, The MPD Suite, The PACT Suite and The PEGsI Suite from Qiagen[®], with Wizard Classic 1+2 from Rigaku Reagents and the Polymer Grid from Hampton Research. Several of these screens were also tested for crystallization of the two His₆-tagged eIF-5A, which were set up at a concentration of 10 g L⁻¹ without the addition of NAD⁺ or GC7.

In addition to the separate crystallization trials, *StreptII-SbDHS1-His₆* and *StreptII-SbDHS2-His₆* tetramers at a concentration of 8 g L⁻¹ were also screened in the presence of equimolar concentrations of either of the two *SbelF-5A-His₆* and 2 mM of GC7 and NAD⁺ with Morpheus[®] HT-96 I and II, Midas[®] HT-96 and Clear Strategy[™] Screen II HT-96 from Molecular Dimensions as well as The Protein Complex Suite (Qiagen[®]).

In figure 5.8 **A**, UV/vis images of some of the more promising initial hits from these screens are compiled. For both *SbelF-5A*, no real crystalline matter could be identified in the tested conditions from the different screens, only phase separation and spherulites (as depicted in **A**). Since they do not contain any tryptophan, their intrinsic fluorescence is very weak and no signal is visible in the UV images. For all three DHS, crystalline structures could be obtained for several conditions. The most promising conditions (e. g. depicted in **A**), were reproduced and tried to optimize in-house. However, while crystals could be grown reproducibly and quite reliably, these crystals were often strongly intergrown and/or consisted of a high number of twisted layers (see figure 5.8 **B**). Although these crystals showed diffraction upon illumination with a synchrotron-generated X-ray beam (beamline P14 of EMBL Hamburg at the PETRA III storage ring; DESY, Hamburg, Germany), the resulting diffraction data contained a high number of overlaid crystal lattices (see **C**), which made it impossible to index and process the data for subsequent structure determination. Therefore, the promising crystallization conditions determined by the initial screens still have to be optimized further to generate high quality, single crystals.

In crystallization trials with DHS1 or DHS2 in complex with either eIF-5A1 or eIF-5A2, mostly phase separation and spherulites was observed, as well as a few promising conditions with crystalline precipitate or small needles (see figure 5.8 **A**). The presence of a DHS in these spherulite or crystalline structures was proved by their UV fluorescence. However, the presence of a eIF-5A in these structures could not be confirmed. Unfortunately, all complex

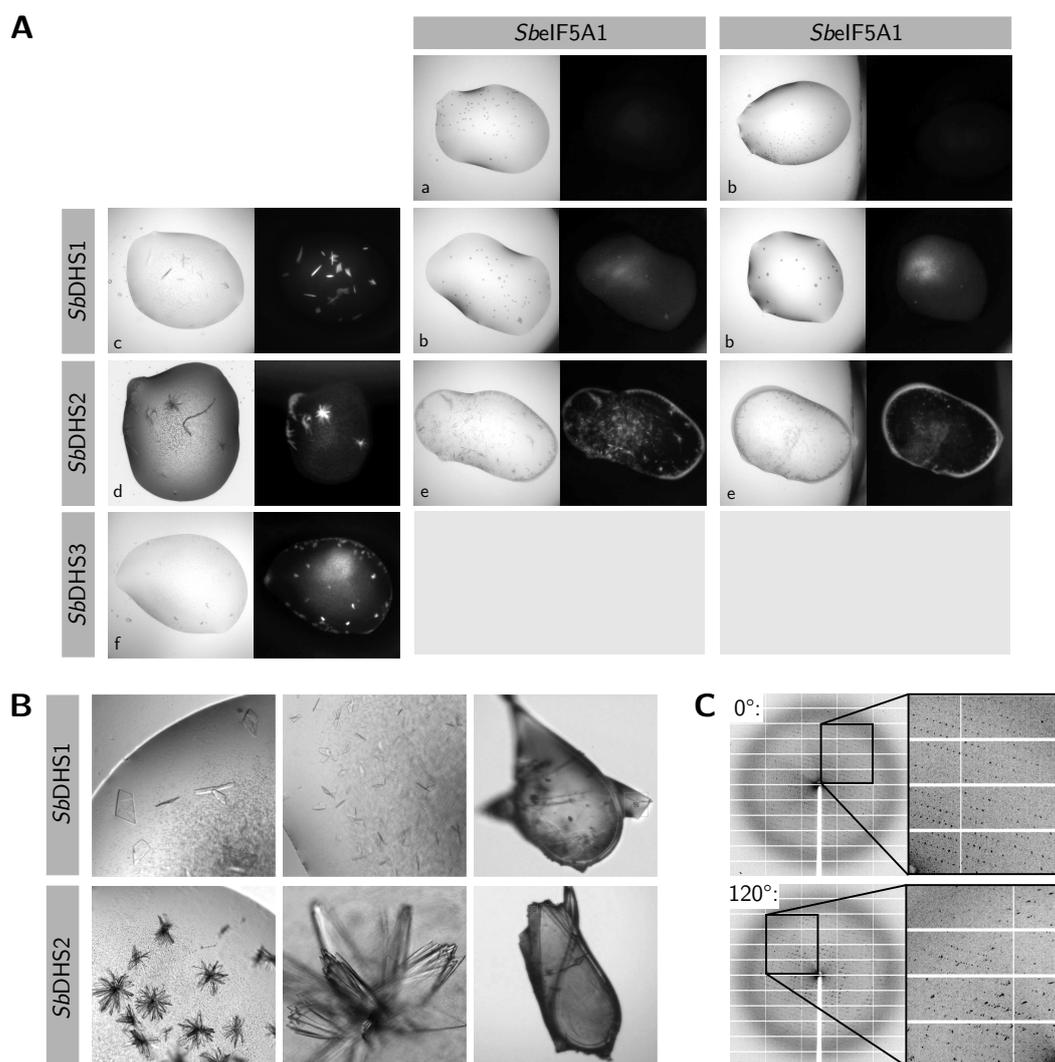


Figure 5.8.: Crystals and diffraction images of *S. bicolor* DHS and eIF-5A isoforms

A: Exemplary initial hits, depicted are visual and UV images. Protein solution was prepared in glycine buffer (0.1 M glycine, 150 mM NaCl, 3 mM NAD⁺, 2 mM GC7, pH 9.0) from *N*-terminally *StreptII*-tagged *SbDHS* at a concentration of 8 g L⁻¹ or *C*-terminally His₆-tagged *SbelF-5A* at 10 g L⁻¹ – or 3.4 g L⁻¹ in combination with either DHS1 or DHS2. 100 nL protein solution was mixed 1:1 with the crystallization solution. **a:** Clear StrategyTM Screen II HT-96 G-12 (Molecular Dimensions; 0.2 M calcium acetate, 0.1 M Tris pH 8.5, 8 % (w/v) PEG 20000, 8 % (v/v) PEG 500 MME); **b:** Clear StrategyTM Screen II HT-96 G-06 (0.2 M calcium acetate, 0.1 M Tris pH 8.5, 15 % (w/v) PEG 4000); **f:** Morpheus[®] HT-96 G-08 (Molecular Dimensions; 20 mM sodium formate, 20 mM ammonium acetate, 20 mM sodium citrate, 20 mM sodium potassium tartrate, 20 mM sodium oxamate, 100 mM HEPES/MOPS pH 7.5, 37.5 % MPD/PEG 1000/PEG 3350 (37.5 %)); **d:** Wizard Classic 1+2 A-06 (Rigaku Reagents; 0.1 M sodium citrate/citric acid pH 5.5, 20 % (w/v) PEG 3000); **e:** Morpheus[®] HT-96 I G-06 (Molecular Dimensions; 20 mM sodium formate, 20 mM ammonium acetate, 20 mM sodium citrate, 20 mM sodium potassium tartrate 20 mM sodium oxamate, 0.1 mM HEPES/MOPS, 20 % ethylene glycol, 10 % PEG 8000); **f:** Midas[®] HT-96 E-07 (Molecular Dimensions; 0.1 M sodium malonate, 0.1 M HEPES pH 7.0, 30 % (w/v) poly(acrylic acid sodium salt) 2100). **B:** Crystals grown in-house based on **d** (*SbDHS2*) and **f** (*SbDHS1*). **C:** Diffraction images of *SbDHS2* crystallized at condition **b** taken at 0° and 120°.

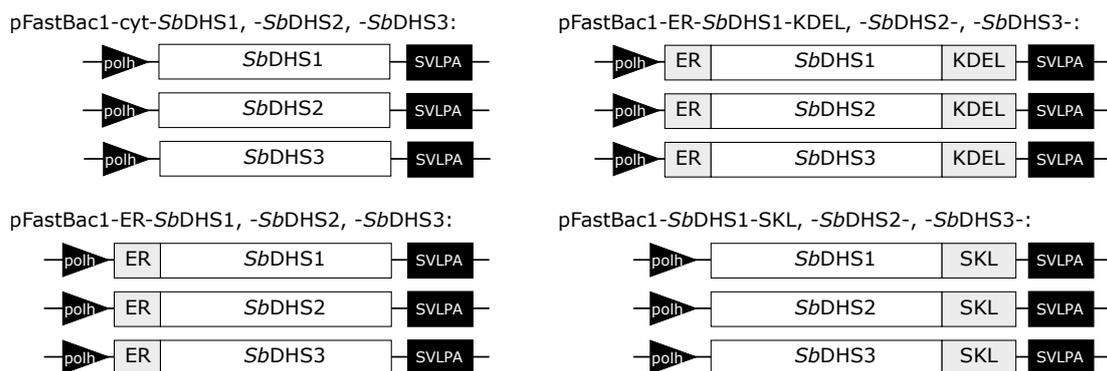


Figure 5.9.: Constructs for baculovirus-mediated expression of *SbDHS*1–3 in insect cells

Constructs for cytosolic, ER (ER, ER+KDEL) and peroxisomal (SKL) localization. *SbDHS*1, *SbDHS*2 and *SbDHS*3. Constructs not to scale. polh: Polyhedrin promoter. SVLPA: SV40 late polyA terminator.

crystallization set-ups tested so far contained 2 mM of the inhibitor GC7, since demonstrated by thermal stability experiments (see section 5.4.3) the inhibitor stabilized the DHS – an effect that was assumed to improve homogeneity and crystallizability in the complex situation. However, as later binding experiments with bio-layer interferometry (BLI) would reveal (see section 5.5.4), this concentration of inhibitor completely prevents any binding between DHS and eIF-5A. The fluorescing spherulites and crystalline matter therefore solely contained DHS protein, and conditions for the complex crystallization will have to be screened again from scratch in the absence of the inhibitor GC7.

5.3.2. *In vivo* crystallization in insect cells

The *in vivo* crystallization of proteins – a rather frequently occurring phenomenon especially in certain cell lines, such as insect cell lines expressing huge amounts of heterologous protein – opens up new possibilities for the structural characterization of proteins in combination with recent developments in data acquisition and analysis, like XFEL and serial crystallography [229, 85, 76, 34]. In the first part of this thesis, it could be demonstrated that the pestiviral autoprotease NS2, a transmembrane protein, yielded crystalline structures upon overexpression in High FiveTM insect cells (see section 2.2.5). However, most likely due to their membrane association, these crystals were very fragile and not suitable for further characterizations (see sections 2.2.7 to 2.2.9).

In an attempt to evaluate the insect cell crystallization method with another heterologously expressed protein and to obtain crystals suitable for X-ray-based structure determination by serial crystallography, 12 constructs of the two *Sorghum bicolor* DHS homologs and their HSS paralog for the expression in Sf21 and High FiveTM insect cells were designed, cloned and tested in the course of this thesis. In figure 5.9, these constructs are depicted schematically. They are based on four modified pFastBacTM 1 vectors from the Bac-to-Bac[®] baculovirus system (Thermo Scientific), which were already applied in the first part of this thesis (see sec-

tion 2.2.1). They are suitable for the transposition of the expression construct into an artificial baculovirus genome (bacmid) and put the expression of the heterologous proteins under the control of the strong polyhedrin promoter from the baculovirus *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV). Additionally, these vectors contain signal sequences for the targeting of the expressed protein to different cellular compartments: pFastBac1-cyt for cytosolic localization, pFastBac1-ER with an *N*-terminal signal sequence for ER targeting, pFastBac1-ER-KDEL with the *N*-terminal ER targeting sequence and a *C*-terminal retention signal for permanent ER localization and pFastBac1-SKL with a *C*-terminal signal sequence for peroxisomal targeting. The empty vectors were obtained from the work group Structural Infection Biology applying new Radiation Sources of Prof. Dr. Lars Redecke from Hamburg and Lübeck University.

The coding sequences for *SbDHS1* (HSS), *SbDHS2* and *SbDHS3* were cloned into these four vectors to yield the final 12 constructs as described in detail in section 7.9.6. Since the expression in insect cells was in this case solely meant for the production of *in vivo*-grown crystals, the addition of an affinity tag was omitted. A codon usage analysis performed on the authentic *SbDHS* sequences in correlation to the codon usage in *S. frugiperda* did not reveal any critical sequence parts, so that sequence optimizations were not necessary (data not shown).

The integrity of the expression cassettes was verified by sequencing (data not shown) prior to transformation into *E. coli* DH10BacTM or DH10EMBacY. As already described in more detail in section 2.2.1, these *E. coli* strains harbor the baculovirus shuttle-vector bMON14272 or EMBacY, respectively, and a helper plasmid, pMON7124, which mediate the transfer of the expression cassette and adjacent regions into the respective bacmid by the site-specific mini-Tn7 transposon. The substitution of the lacZ α fragment in the original bacmid mini-Tn7 transposon in one of the bacidms by the desired expression cassette enables subsequent blue/white screening of the resulting clones on selective agar plates to identify successful transposition events. To ensure transposition homogeneity of the clone of choice for all 12 constructs, the clones were re-streaked twice onto fresh agar plates supplemented with the necessary antibiotics, X-gal and IPTG.

For each construct, three white clones were chosen for a colony PCR with primers M13for and M13rev to verify the completeness of the expression cassette. Figure 5.10 shows the result of the agarose gel electrophoresis of the colony PCR products. Clones without a successfully transposed expression cassette would yield a PCR product of about 350 bp (see figure 2.11 B, lane 12). Positive clones would feature a PCR product in the range from 3.433 bp for the shortest cytosolic constructs to 3.508 bp for the longest ER-KDEL constructs. Mixed clones – which can occur occasionally and contain a mixture of bacteria harboring the original and the already transposed bacmid – would contain both PCR products (data not shown). However, in neither of the clones the PCR product of the empty bacmid could be detected. From 72 tested clones, 57 show a distinct band for a successfully transposed expression cassette. The other colonies yield either no band at all or only a faintly visible band at the size of the correct transposition product.

For bacmid preparation, a single clone was chosen for each expression construct in either of the bacidms (indicated by an asterisk in figure 5.10). In addition to some advancements over

5.3. Crystallization of *SbDHS* and *SbE1F-5A* isoforms

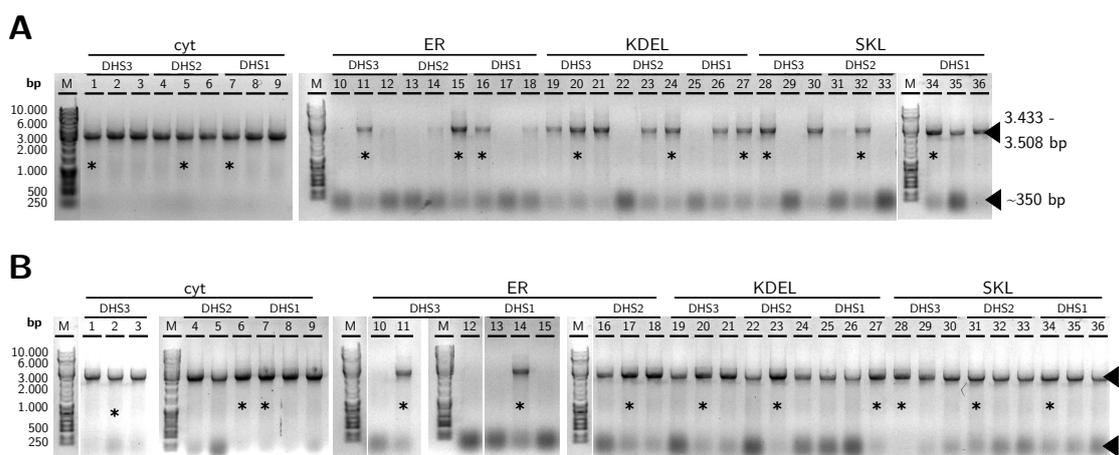


Figure 5.10.: Colony PCR results of *E. coli* DH10EMBacY and DH10BacTM clones transformed with *SbDHS* variant pFastBac1 constructs

A: DH10BacTM clones. **B:** DH10EMBacY clones. The constructs all contain one of the three *SbDHS* homo/paralogs. The type of localization signal in the respective clone is denoted above the lanes. The expected PCR product sizes for successfully transposed expression cassettes range from 3.433 bp for the cytosolic constructs to 3.508 bp for the ER-KDEL constructs, as indicated above. Empty bacmid produces a PCR product with a size of about 350 bp. M: Molecular weight marker (GeneRulerTM 1 kb DNA Ladder, Thermo Scientific). Asterisks indicate clones chosen for bacmid preparation and subsequent insect cell transfection.

the older bMON14272 bacmid – like protease and chitinase deficiencies, extended cell viability and reduced proteolysis – EMBacY also allows the easy monitoring of the heterologous protein expression rates via the simultaneous expression of EYFP, driven by the same polyhedrin promoter as the desired heterologous protein. Since the *SbDHS* expression constructs do not possess any easily detectable features themselves, the EMBacY bacmids were initially chosen for subsequent experiments.

Therefore, these 12 bacmids were transfected into Sf21 insect cells, as described in section 7.13.2. Transfected insect cells were incubated for 7 d in 6-well plates, and virus of the first generation (P1) was obtained between 48 to 72 h after transfection by removing the cell culture supernatant and replacing it with fresh medium. Fresh Sf21 insect cells were infected with P1 supernatant to generate a second virus generation (P2), and from this, a third virus generation (P3) was derived (see section 7.13.3 for details).

The infected cells were visually checked for the fluorescence of EYFP with a Blue/Green LED Transilluminator (Nippon Genetics) and the corresponding Amber filter to monitor *SbDHS* expression rates. Figure 5.11 **A** and **C** show images of the 12 constructs expressed in Sf21 insect cells from P1 (**A**) and P3 (**C**) virus generation. For ER-*SbDHS3*-KDEL and *SbDHS3*-SKL, the transfection – or infection, respectively – failed. For all other constructs, a clearly visible EYFP fluorescence could be detected at about day 4, while first signs of infection such as a disruption of the confluent cell layer and an enlargement of infected cells was noticeable already at day 2 to 3, although they were not as prominent for Sf21 cells as for High FiveTM.

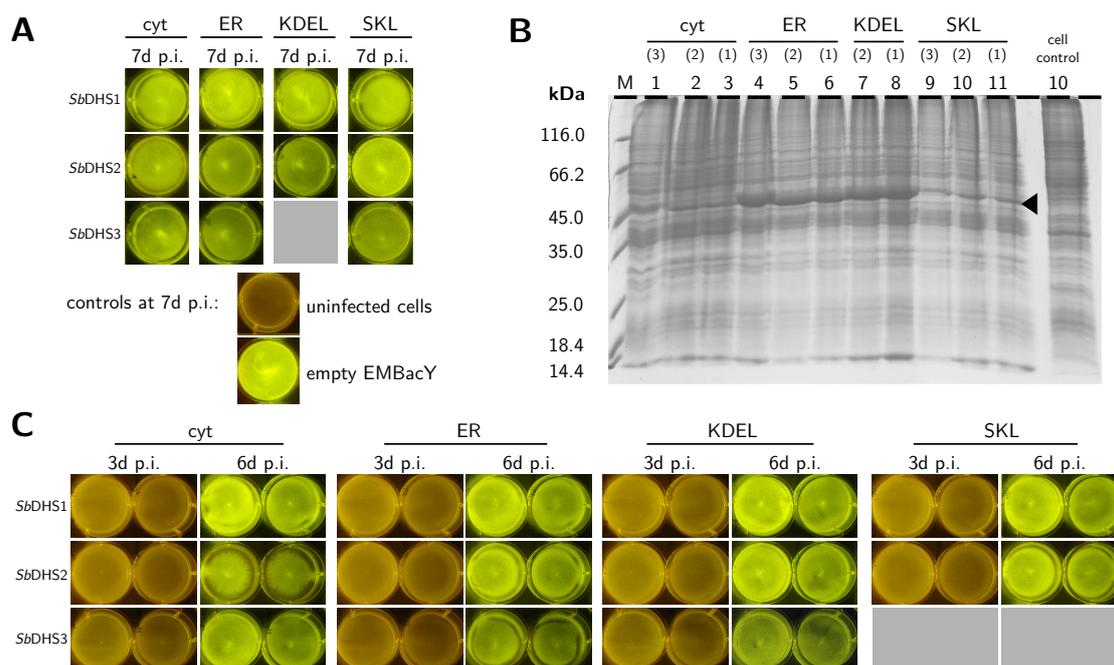


Figure 5.11.: Sf21 insect cells transfected with *SbDHS*1–3 constructs in bMON14272

Images of Sf21 insect cells infected with P1 (**A**) and P3 (**C**) EMBacY virus containing the 12 different *SbDHS* constructs. Uninfected cells (no fluorescence) and empty EMBacY (from another experiment) are depicted as control and visual orientation. The images were taken with a consumer camera on a Blue/Green LED Transilluminator (Nippon Genetics) with an Amber filter. The type of localization signal is denoted above the images. Grey boxes indicate failed transfection/infection. **B**: SDS-PAGE result from samples from **A** taken at day 8. The type of localization signal and the number of the *SbDHS* (in brackets) is denoted above the lanes. The black arrowhead indicates the bands of the DHS subunits at about 41.5 kDa for cytosolic constructs to 43.7 kDa for the longest ER-KDEL constructs. M: molecular weight marker (Pierce™ Unstained Protein MW Marker, Thermo Scientific).

Interestingly, although all cells exhibit similar EYFP fluorescence – which should in theory correlate with the amount of expressed heterologous protein – the amount of expressed *SbDHS*1, *SbDHS*2 and *SbDHS*3 from different constructs in the SDS-PAGE samples depicted in figure 5.11 **B** varies significantly. In this experiment, the ER and ER-KDEL samples exhibit an intense band around the height of the 45 kDa molecular weight marker band. This correlates with the expected molecular weight of the *SbDHS* subunits of 43.2 and 43.7 kDa for ER and ER-KDEL constructs, respectively. The samples from cytosolic and SKL constructs also show a band around the same height – which proves the expressibility of the corresponding constructs –, but it is much less distinct. Apart from the appearance of this SDS-PAGE, earlier preliminary results suggest that the differences in the amount of expressed protein are independent of the type of localization signal used (data not shown).

For all Sf21 insect cells infected with different virus generations of the EMBacY bacmids, no crystalline structures could be observed under an inverted light microscope (Olympus CKX41). Since cells that actually express a fair amount of *SbDHS* also did not exhibit any

crystal growth, this is probably not due to low expression levels. Subsequent experiments will have to include High FiveTM insect cells as expression hosts, which yielded crystals for the pestiviral autoprotease NS2 from the first part of this thesis, where Sf21 cells also did not support the growth of crystalline structures. Unfortunately, a bacterial contamination in the transfection reagent and the time-consuming trouble-shooting as a consequence thereof finally prevented further experiments in the course of this thesis.

5.4. Thermal stability determination of *SbDHS* and *SbeIF-5A* isoforms

5.4.1. Buffer variation and pH optimum

In several activity studies with DHS of different origin and plant HSS, the pH optimum of the catalyzed reaction was determined to be at a basic pH around 9.0 to 9.5 [179, 190, 151, 148]. However, under physiological conditions the DHS enzyme as well as its substrate protein eIF-5A are localized in the cytosol. While the enzymatic reaction reaches its optimum – especially for low spermidine concentrations – at a basic pH due to the significantly higher fraction of diprotonated and thereby optimally charged spermidine substrate (see section 4.2.5 for further details), the DHS (or HSS) enzyme itself and the eIF-5A substrate can be expected to be optimized for physiological conditions. Since buffer conditions and additives can have an important impact on protein stability during purification procedures, handling and storage, as well as on subsequent experiments, the thermal stability of the three *SbDHS* and the two eIF-5A homologs was characterized in more detail in the course of this thesis.

As already mentioned in 5.1.2, differential scanning fluorimetry (DSF) can be applied to easily determine the thermal stability of soluble proteins under a variety of conditions (see section 7.14.5 for experimental details). This biophysical method is based on the step-wise heating of a set of protein samples (e.g. up to 96 samples in the RT-PCR system used in the course of this thesis) under the examined conditions in the presence of the fluorescent dye SYPRO Orange[®]. The dye sensitively reports on the thermally induced unfolding of the protein by a significant increase in its fluorescence – initiated by the binding to exposed hydrophobic areas of the unfolding protein. The fluorescence signal intensity plotted over the sample temperature ideally results in a sigmoidal increase of the SYPRO Orange[®] fluorescence in the course of the protein's "melting" process, and a subsequent decrease in fluorescence due to self-aggregation of the unfolded protein, which displaces the bound dye molecules and quenches their fluorescence. From the sigmoidal portion of the melting curve, the protein's melting temperature $T_{m(DSF)}$ under the given conditions can be calculated.

In this thesis, the melting temperature $T_{m(DSF)}$ was derived from the melting curves by either of two ways: If the melting curve exhibits a nearly ideal, sigmoidal shape, the data can be fitted to the Boltzmann Equation, from which the melting point is calculated as the midpoint of the unfolding transition. For non-ideal curves, or a fast estimation, the first derivative of the fluorescence signal intensity can be plotted as a function of temperature. The $T_{m(DSF)}$ is represented as the inflection point of the original curve, which is indicated by the minimum of the first derivative. For curves which exhibit more than one transition, the latter method can give fast estimates of both transitions with sufficient accuracy, while fitting to the Boltzmann Equation would be much more elaborate. In this thesis, both methods

5.4. Thermal stability determination of *SbDHS* and *SbIF-5A* isoforms

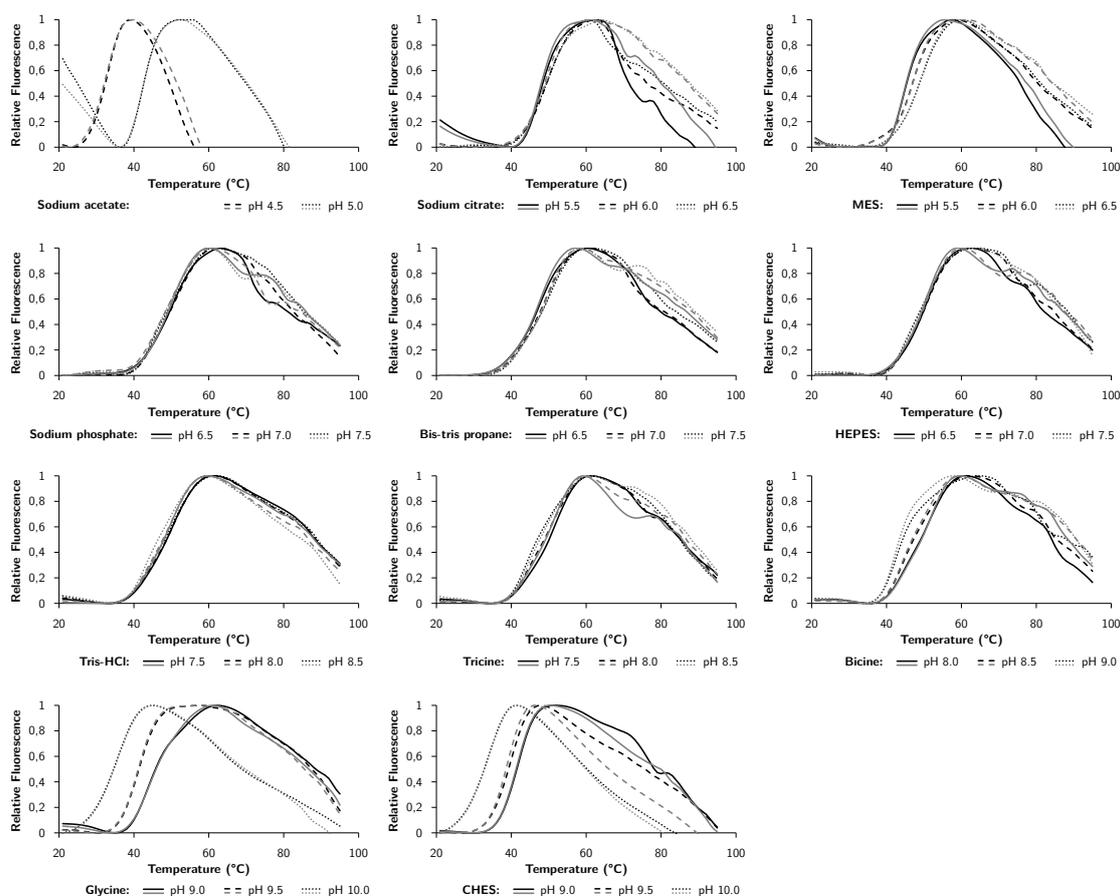


Figure 5.12.: Melting curves of *SbDHS3* (HSS) under varying buffer and pH conditions determined by DSF

The *StreptII-SbDHS3-His₆* protein samples were diluted to 0.2 mg mL^{-1} in the indicated buffer at the indicated pH (100 mM buffer substance, 150 mM NaCl) and measured in duplicates for each pH value (black and gray lines, respectively). The samples were heated and measured in 1 K steps from 21 to 95 °C.

were applied for data interpretation, depending on data quality. To ensure comparability, all melting curves inside a certain dataset were analyzed with the same mathematical approach, which is indicated below the corresponding figures.

In a first thermal stability experiment, a range of 11 buffer substances at three different pH values in 0.5 steps around their highest buffering capacity were tested at a physiological salt concentration of 150 mM with samples from *StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆* and *StreptII-SbDHS3-His₆*, as well as *SbIF-5A1-His₆* and *SbIF-5A2-His₆*.

In figure 5.12, the duplicate melting curves for all tested buffers and pH values are compiled exemplarily for *StreptII-SbDHS3-His₆* (HSS). The melting curves exhibit the ideal behavior described above, with the sigmoidal increase in fluorescence intensity during protein unfolding, followed by the decrease in fluorescence upon aggregation. For most buffer substances, the

5.4. Thermal stability determination of *SbDHS* and *SbeIF-5A* isoforms

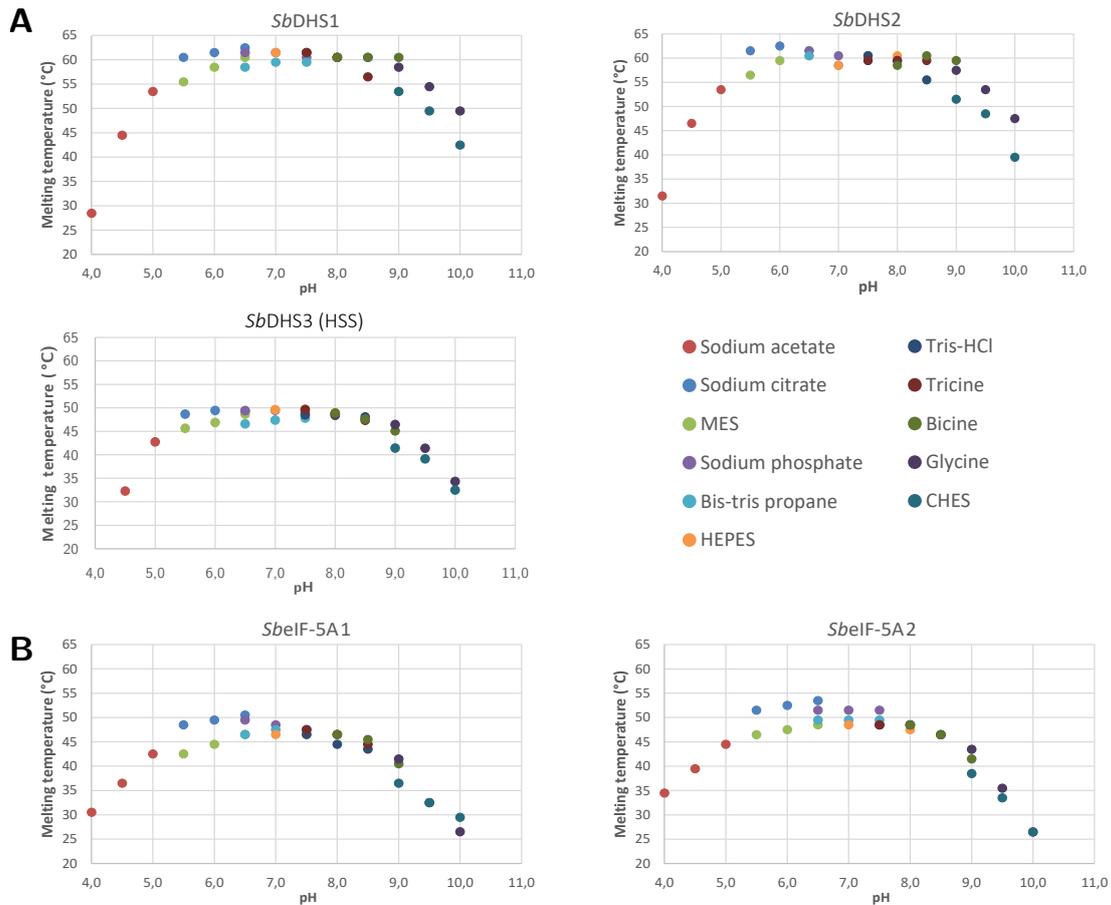


Figure 5.13.: Melting temperatures of *SbDHS* and *SbeIF-5A* homo/paralogs under varying buffer and pH conditions determined by DSF

The protein samples were diluted to 0.2 g L^{-1} in the indicated buffer at the indicated pH (100 mM buffer substance, 150 mM NaCl) and measured in duplicates for each pH value. The samples were heated and measured in 1 K steps from 21 to 95 °C (see figure 5.12 for exemplary melting curves). The melting temperatures were calculated from the melting curves by derivative analysis. **A:** Results for *StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆* and *StreptII-SbDHS3-His₆*. **B:** Results for *SbeIF-5A1-His₆* and *SbeIF-5A2-His₆*.

different pH values have no significant effect on protein stability, and the corresponding melting curves are almost identical. Exceptions are the buffers with very low (sodium acetate pH 4.0–5.0) or very high pH values (glycine-NaOH or CHES, pH 9.0–10.0). Here, the lower or higher the pH, respectively, the less stable is the *SbDHS3*.

From these melting curves for *StreptII-SbDHS3-His₆* and comparable curves from *StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆*, *SbeIF-5A1-His₆* and *SbeIF-5A2-His₆* (data not shown), melting temperatures were calculated by derivative analysis. In figure 5.13, the resulting melting temperatures for the different buffers are plotted over the examined pH values for *SbDHS* (A) and *SbeIF-5A* samples (B).

5.4. Thermal stability determination of *SbDHS* and *SbeIF-5A* isoforms

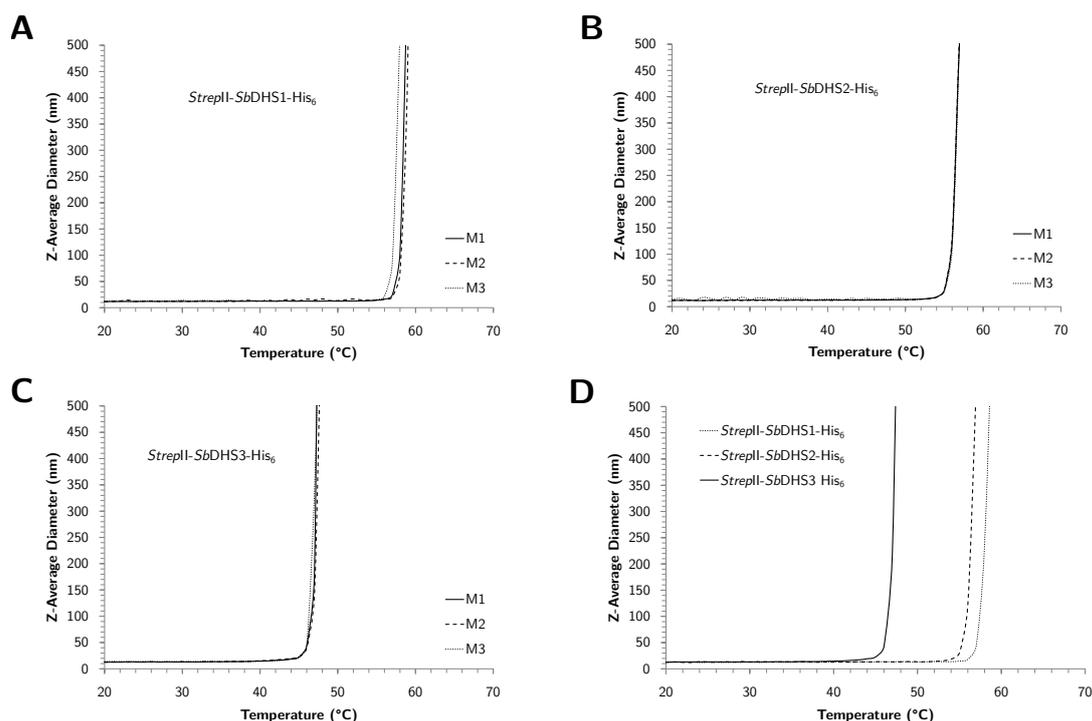


Figure 5.14.: Melting curves of *SbDHS*/HSS determined by DLS

A–C: Melting curves of *StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆* and *StreptII-SbDHS3-His₆* ("HSS") at 2 g L^{-1} in phosphate buffer (100 mM NaH_2PO_4 , 140 mM NaCl, 0.4 mM NAD^+ , pH 7.5). The graphs show the z-average diameter (nm) of the homotetramers in solution upon heating from 20 to 70 °C, calculated from DLS data at the corresponding temperatures. M1–M3: replicate measurements. **D:** Representation of the arithmetic mean of the z-average diameter of the measurement triplicates from **A–C** at the corresponding temperature.

All three *SbDHS* have their pH optimum around 7.2–7.6, with a stability range from pH 6.0 to 8.0, where buffer substance and pH do not significantly influence thermal stability. Below pH 6.0 and above pH 8.0, the stability decreases. At high pH, the *SbDHS* are more stable in glycine-NaOH or bicine than in CHES buffer, so that their stability in glycine-NaOH at pH 9.0 is still similar to their stability at the pH optimum. Overall, *SbDHS3* (HSS) reaches melting temperatures of around 50 °C, while DHS1 and DHS2 exhibit similar melting temperatures of about 60 °C around their pH optimum and are therefore noticeably more stable.

This result has been confirmed by additional dynamic light scattering (DLS) measurements. Therefore, the samples were also heated step-wise and the z-average diameter (intensity weighted harmonic mean size) was calculated from the autocorrelation function of the intensity fluctuation of laser light scattered by the randomly moving proteins in the sample. The z-average diameter increases considerably, when the protein starts to unfold. The onset of this increase is defined as the melting temperature $T_{m(\text{DLS})}$ of the protein (see section 7.14.6 for experimental and theoretical details).

In figure 5.14, the z-average diameter of the three *StreptII-SbDHS-His₆* (A–C) is plotted against the sample temperature for three independent measurements per sample. D shows the plot of the arithmetic mean of the three measurements for all three *StreptII-SbDHS-His₆*. All three measurements per sample give almost identical melting curves. While for *StreptII-SbDHS3-His₆* the melting temperature $T_{m(DLS)}$ is about 45 °C, *StreptII-SbDHS1-His₆* and *StreptII-SbDHS2-His₆* exhibit melting temperatures of approx. 56 and 53 °C, respectively. Although the melting temperatures derived from DSF and DLS measurements cannot be compared directly, the difference in the determined melting temperatures confirms the DSF results: *SbDHS3* (HSS) is considerably less stable than *SbDHS1* and *SbDHS2*.

Both *Sbelf-5A* exhibited similar stability in the tested buffers and at different pH (see figure 5.13 B), with *Sbelf-5A2* slightly more stable overall. Their pH optimum is localized around a pH of 6.8 with melting temperatures between 45–50 °C for *Sbelf-5A1* and around 50 °C for *Sbelf-5A2*. The buffer substance at a given pH has no significant influence on stability, with the exception of sodium citrate and MES at lower pH values of 5.5–6.5: Here, sodium citrate increases stability noticeably compared to MES. This effect is also visible for the DHS samples, although a bit less pronounced.

5.4.2. Effect of NaCl concentration on thermal stability

With buffer substances and pH values determined in the previous DSF experiment, the influence of the NaCl concentration on the thermal stability of the three *StreptII-SbDHS-His₆* and both *Sbelf-5A-His₆* was examined. Therefore, sodium citrate, sodium phosphate, HEPES, Tris-HCl and glycine-NaOH buffers spanning pH values from 6.5 to 9.0 and supplied with different NaCl concentrations between 0 and 500 mM were tested with DSF (see figure 5.15).

SbDHS1 and *SbDHS2* exhibit a very similar behavior (A). Their melting temperatures are not influenced by the NaCl concentration. *SbDHS3*, however, seems to be slightly less stable with higher NaCl concentrations. Nevertheless, it has to be noted that the data was not as consistent as the data for *SbDHS1* and *SbDHS2*, so that only an overall trend can be stated at this point.

In contrast to the *SbDHS* samples, the *Sbelf-5A* samples clearly exhibit a dependence of the thermal stability on NaCl concentration. The melting temperatures of both *eIF-5A* noticeably increase with increasing NaCl concentrations. The effect is most prominent for higher pH values, especially glycine-NaOH pH 9.0, where the samples with 500 mM NaCl are about 14 K more stable than samples without NaCl. For sodium citrate pH 6.5 on the other hand, the stabilizing effect is merely noticeable and increases the melting temperature by only 2–3 K.

Taken together, the three *SbDHS* and both *eIF-5A* have their pH optimum at physiological values. *SbDHS1* and *SbDHS2* are very stable and reach melting temperatures around 60 °C, but *SbDHS3* – which is considered an HSS – is considerably less stable with its highest melting temperatures around 50 °C. While the three *SbDHS* are quite robust over NaCl concentrations from 0–500 mM NaCl in all tested buffers, both *eIF-5A* are stabilized by higher NaCl concentrations – especially at higher pH values, at which activity assays are routinely performed.

5.4. Thermal stability determination of *SbDHS* and *Sbelf-5A* isoforms

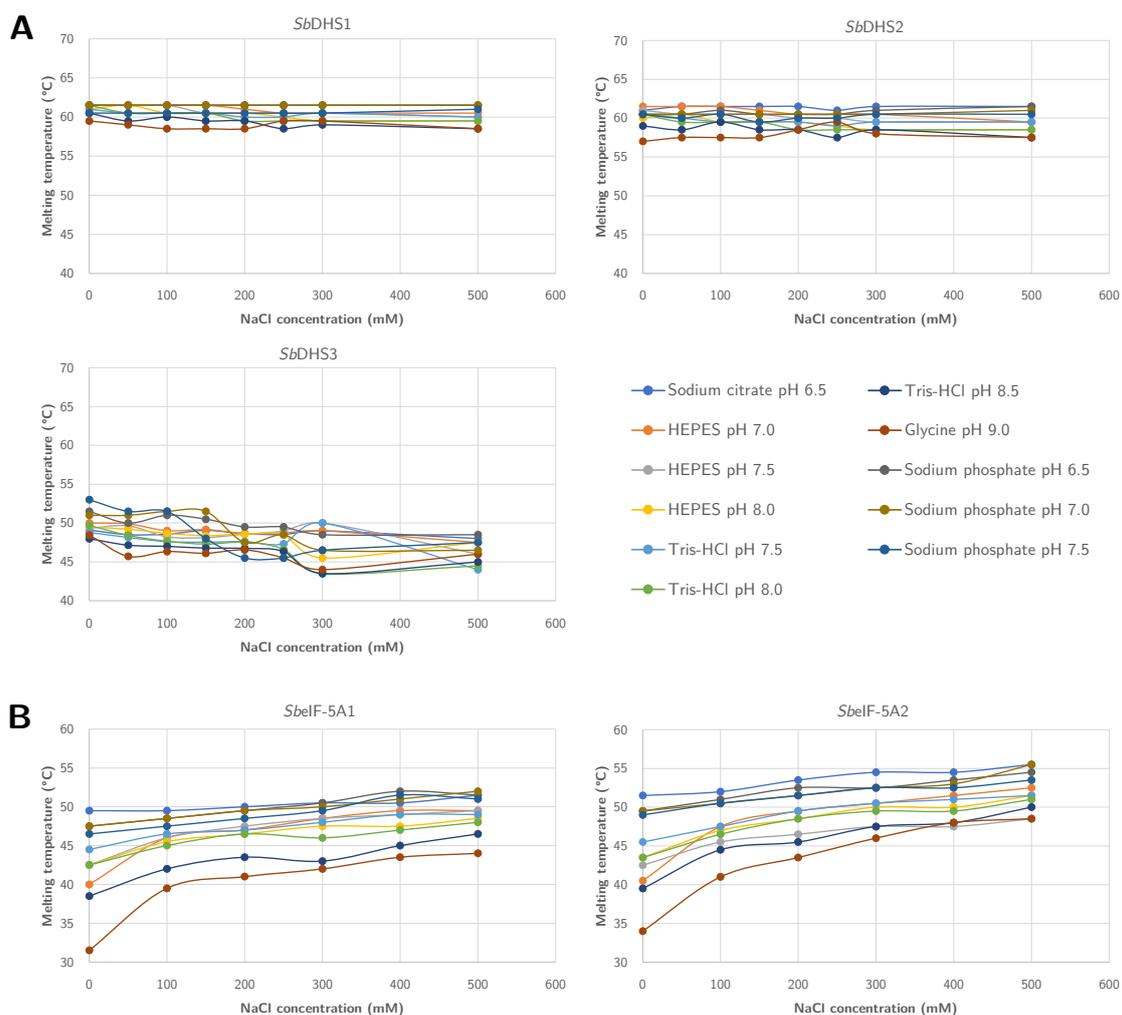


Figure 5.15.: Melting temperatures of *SbDHS* and *Sbelf-5A* homo/paralogs under selected buffer, pH and salt conditions determined by DSF

The protein samples were diluted to 0.2 g L^{-1} in 100 mM buffer at the indicated pH, mixed with different NaCl concentrations and measured in duplicates for each pH value. The samples were heated and measured in 1 K steps from 21 to $95 \text{ }^\circ\text{C}$. The melting temperatures were calculated from the melting curves by derivative analysis. **A:** Results for *StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆* and *StreptII-SbDHS3-His₆*. **B:** Results for *Sbelf-5A1-His₆* and *Sbelf-5A2-His₆*.

5.4.3. The stabilizing influence of NAD^+ and GC7 on *SbDHS* variants

The effect of the addition of NAD^+ and/or the inhibitor GC7 at concentrations of 0 , 2.5 and 5 mM on the thermal stability of the three *SbDHS* variants (*StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆* and *StreptII-SbDHS3-His₆*) was tested in four different buffers (sodium citrate pH 6.5 , HEPES pH 7.0 , sodium phosphate pH 7.5 and glycine-NaOH pH 9.0 ; 100 mM each) at three NaCl concentrations (0 , 150 and 300 mM). In figure 5.16 **A**, exemplary melting

curves for the results from sodium phosphate and glycine-NaOH buffer with 150 mM NaCl are compiled.

The separate addition of 2.5 and 5 mM NAD^+ had similar effects on all three *SbDHS*. The same was observed for the separate addition of 2.5 and 5 mM GC7 (see figure 5.16 A). Therefore, in the following it will not be distinguished between these cases for simplification. It should be noted that "0 mM NAD^+ " in these experiments means no additional NAD^+ , however, the original protein samples (about 10 mg mL^{-1}) already contained 2 mM NAD^+ for storage stability reasons (see 5.1). In the final DSF assays, this concentration was diluted to $40 \mu\text{M}$. As the results from the DSF assay with buffers containing different NaCl concentrations already indicated (see above), the NaCl concentration had also only small effects on the thermal stability of the different *SbDHS* in these DSF experiments, and no clear trend was deducible. Therefore, if not stated otherwise, results are summarized for all three NaCl concentrations.

Besides the overall difference in thermal stability between *SbDHS3* (HSS) and *SbDHS1/2* (see above), the shape of the curves also differed significantly under conditions without GC7 or NAD^+ : while the melting curves of *SbDHS3* exhibit a monophasic shape, the melting curves of *SbDHS2* and – to a much lesser extent – *SbDHS1* show a biphasic behavior, indicating two distinct unfolding or disassembling events.

Upon the addition NAD^+ , *SbDHS3* (HSS) is only marginally stabilized by up to 5 K in all four buffers, while the melting curves of *SbDHS1* and *SbDHS2* lose their biphasic shape in favor of the second transition, without further significant stabilizing effects (see figure 5.16 A).

In contrast, the separate addition of GC7 to the protein samples has a more pronounced influence on the thermal stability: similar to NAD^+ , GC7 shifts the biphasic melting curve behavior of *SbDHS1* and *SbDHS2* towards a monophasic curve at the second transition. In addition, the melting temperatures of *SbDHS3* (HSS) are increased by 7–12 K for sodium citrate, HEPES and sodium phosphate buffers, and up to about 15–18 K in glycine-NaOH buffer (see figure 5.16 A). For *SbDHS1* and *SbDHS2*, the separate addition of GC7 to the samples resulted in a thermal stabilization of 3–7 K for sodium citrate, HEPES and sodium phosphate buffers, and 10–12 K in glycine-NaOH buffer – with a slightly more stabilizing effect for lower NaCl concentrations (data not shown).

The influence of the combined addition of NAD^+ and GC7 on the thermal stability is primarily based on the significant effect of GC7, which is only slightly enhanced by the presence of additional NAD^+ (data not shown). For *SbDHS3* (HSS), the stabilizing effect is in the range of 10–12 K for sodium citrate and HEPES, 12–16 K for sodium phosphate and about remarkable 21 K for glycine-NaOH. Overall, the stabilizing effects of GC7 in combination with additional NAD^+ are less pronounced for *SbDHS1* and *SbDHS2* samples. In sodium citrate or HEPES buffers, the melting temperatures increase about 4–9 K, in sodium phosphate and glycine-NaOH 12–16 K.

The highest overall thermal stability is achieved for all *SbDHS* in glycine-NaOH buffers at pH 9.0 and sodium phosphate buffers at pH 7.5, both supplied with GC7 and NAD^+ – resulting in melting temperatures of about 66°C (*SbDHS3*) and 73°C (*SbDHS1* and *SbDHS2*).

Based on these experiments, the influence of GC7 on the thermal stability of the three *S. bicolor* DHS is more distinct than that of NAD^+ . However, the effect of 2.5 and 5 mM GC7

5.4. Thermal stability determination of *SbDHS* and *SbeIF-5A* isoforms

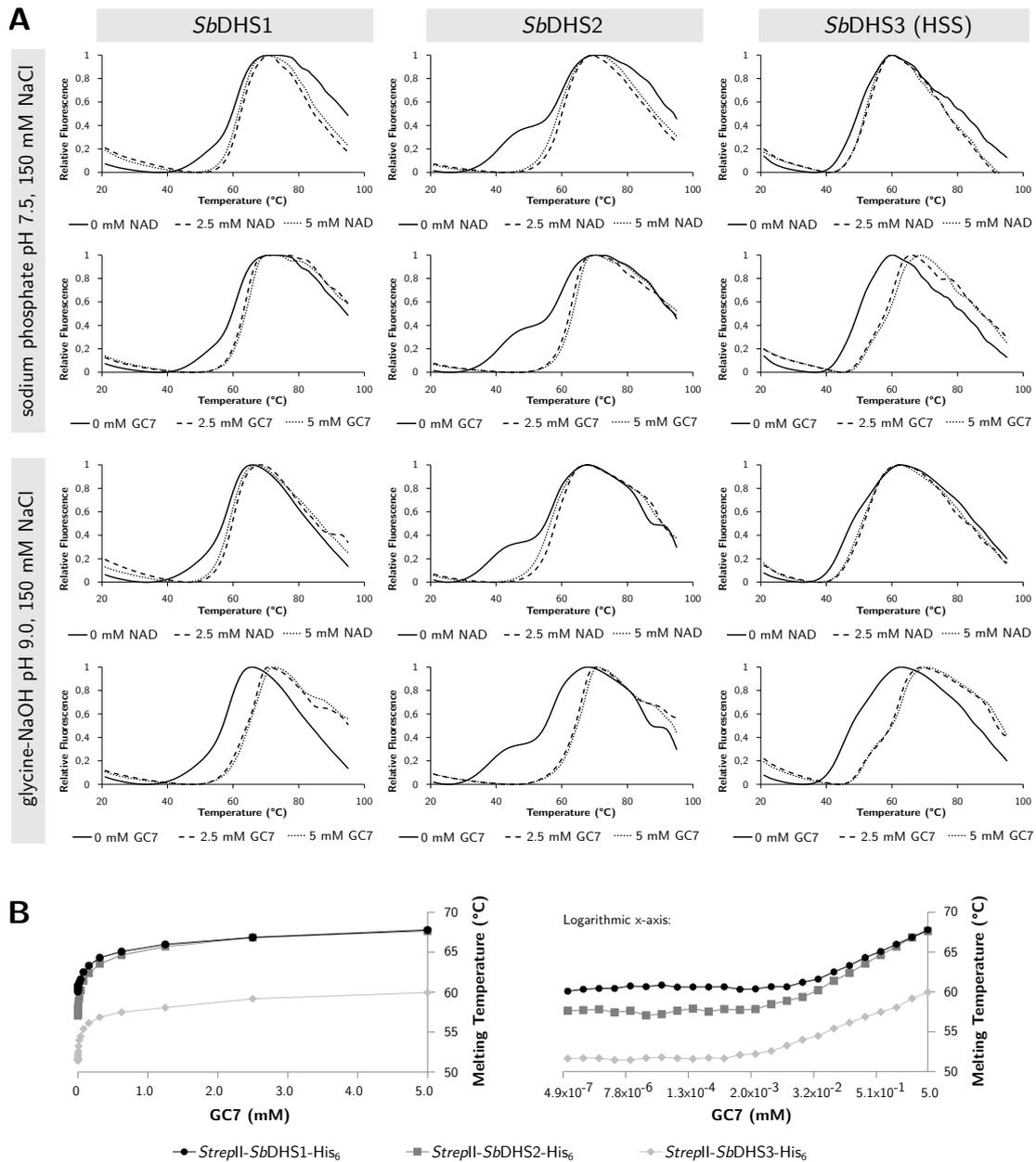


Figure 5.16.: Melting curves of *SbDHS* variants in correlation to NAD and GC7 concentration, determined by DSF

A: The protein samples were diluted to 0.2 g L^{-1} in the indicated buffer. The samples were heated and measured in 1K steps from 21 to 95°C . The melting curves are exemplary results for the three *SbDHS* in either 100 mM glycine-NaOH buffer at pH 9.0 or 100 mM sodium phosphate buffer at pH 7.5 and with and without the addition of NAD^+ or the inhibitor GC7. The relative fluorescence was calculated from raw data by normalizing to the minimum and maximum values of the corresponding melting curve. **B:** The protein samples were diluted to 0.2 g L^{-1} in 100 mM sodium phosphate buffer at pH 7.5 with 150 mM NaCl and 1 mM NAD^+ . The GC7 concentrations were varied from 0.6 nM to 5 mM and prepared as 1:2 serial dilutions. The samples were heated and measured in 1K steps from 21 to 95°C in duplicates. The melting temperatures were calculated as inflection point from a Boltzmann fit of the corresponding melting curves.

is very similar. To determine the minimum GC7 concentration that has a stabilizing effect, the three *SbDHS* were diluted in sodium phosphate buffer at pH 7.5 and supplied with a fixed concentration of 1 mM NAD^+ . GC7 was added in a 1:2 dilution series with concentrations ranging from 0.6 nM to 5 mM. Figure 5.16 **B** shows the resulting melting temperatures of the three DHS in correlation to the GC7 concentration. The non-logarithmic representation (left graph) clearly shows that the curve of the melting temperatures exhibits a hyperbolic or logarithmic shape trending against a plateau with maximal melting temperatures of about 60 °C for *SbDHS3* (HSS) and 68 °C for *SbDHS1* and *SbDHS2*. Although the shape of the curve is similar to the hyperbolic shape of a Michaelis-Menten kinetic, classic transformations of the data into a Lineweaver-Burk or Eadie-Hofstee plot immediately refute this assumption, since the resulting graphs are not linear (data not shown).

Due to the broad range of GC7 concentrations and the 1:2 dilution series, the lower concentrations are not sufficiently resolved in this representation. If the GC7 concentration is depicted in logarithmic scale (right graph), it becomes obvious that the melting temperature does not change significantly until the GC7 concentration reaches a threshold of about 1.2 μM for *SbDHS3* (HSS), and 2.4 and 9.8 μM GC7 for *SbDHS1* and *SbDHS2*. Above this threshold the curves can be described by a logarithmic function.

Interestingly, the threshold concentrations are in the same range as the concentrations of the corresponding DHS subunits in the samples of about 4.6 μM . Taking into account the low K_i of GC7 of 9.7 nM (determined for DHS from rat testis [114]) and assuming that this K_i also applies to all *S. bicolor* DHS/HSS, based on the law of mass action approx. 26.1, 52.1 and 99.8% of spermidine binding sites are occupied by the inhibitor GC7 at the indicated concentrations, respectively. However, it has to be taken into account that inhibitor binding most likely depends on the prior binding of NAD^+ , so that the fraction of DHS active sites able to bind GC7 might not be 100%.

5.5. Determination of binding kinetics for the different *SbDHS* isoforms by bio-layer interferometry

The highly specific deoxyhypusination of the eIF-5A by the DHS enzyme in eukaryotic cells is preceded by a similarly specific binding event between the two proteins. In angiosperms, the HSS – which developed from duplication of the DHS gene – is supposed to have lost its ability to bind the substrate protein eIF-5A, thereby specializing on the homospermidine synthase side reaction (see section 4.3 for further details). The three DHS copies of the Poaceae model plant *Sorghum bicolor*, which were biophysically characterized in the course of this thesis, might therefore represent different evolutionary stages on the way from DHS to HSS. Activity assays already revealed that *SbDHS3* only exhibits very low DHS activity, but improved homospermidine synthesis rates (A. Wesseling¹, internal communication). Although no significant differences between *SbDHS1* and *SbDHS2* neither in DHS nor HSS activity have been detected to date, it is possible that while one of these enzymes retains its original DHS

¹Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

function, the other is a more recent duplicate that has not yet undergone enough distinct changes to clearly display HSS characteristics.

Since the specific binding of eIF-5A to the active funnel area of the DHS enzyme is a crucial prerequisite for the deoxyhypusination reaction, the binding between the three *SbDHS* and the two eIF-5A substrate proteins was characterized in more detail in this thesis – to pick up on subtle differences that might not be detectable by activity measurements alone. Therefore, a biophysical technique termed "bio-layer interferometry" (BLI) was applied, which allowed the sensitive measurement of binding kinetics.

BLI exploits the interference pattern of white light reflected at an internal reference surface and the bio-layer surface of an optical fiber tip (biosensor). In figure 5.18 **A**, the general principle is illustrated in detail. From this interference pattern of the wavelengths of white light, the wavelength shift $\Delta\lambda$ – that results from the differences in the reflected light caused by the thickness of the bio-layer on the tip of the biosensor – can be calculated with high accuracy. This wavelength shift sensitively reports on the proceedings at the bio-layer: upon molecule association to the bio-layer, $\Delta\lambda$ increases, while the dissociation of molecules is reflected by a decrease in $\Delta\lambda$. The time-resolved change in $\Delta\lambda$ allows the determination of association and dissociation curves, from which a complete kinetic binding profile can be derived.

For the BLI binding measurements, one binding partner has to be immobilized as so-called "ligand" on the tip of the biosensor, while the other – the "analyte" – is supplied in solution at specific concentrations. Optimally, measurements are performed in both orientations, with both binding partners acting once as analyte and once as ligand. Since *StreptII*-tagged *SbDHS* as well as eIF-5A had already been produced in large amounts and high purity in the course of this thesis (see section 5.1), an antibody-based approach was chosen for the immobilization of the ligand to the sensor tip, which exploits the very high affinity of the mouse-derived *StrepMAB*-Immo antibody (iba) to an enhanced *StreptII*-tag variant, the serine-alanine-*StreptII*-tag. In figure 5.18 **B**, the measurement setup at the biosensor tip is illustrated schematically. To immobilize the *StrepMAB*-Immo antibody to the sensor tip, the Dip and Read™ Anti-Mouse IgG Fc Capture Biosensors (AMC; PALL fortéBIO) were applied. These

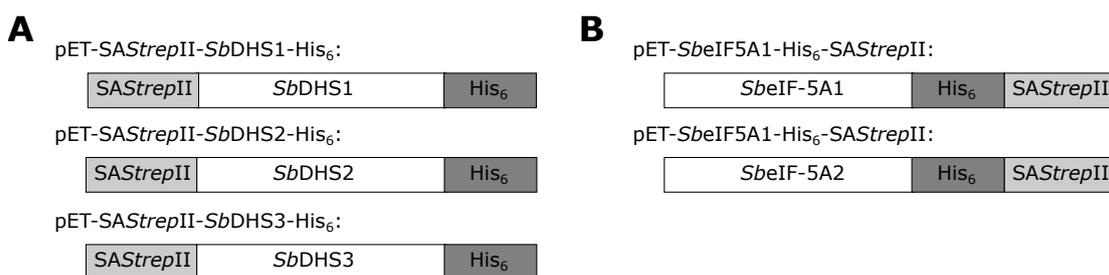


Figure 5.17.: *SbDHS* isoforms constructs for BLI measurements

A: *SbDHS* isoform 1, 2 or 3 with an *N*-terminal Ser-Ala-*StreptII*-tag and a *C*-terminal His₆-tag. **B:** *SbeIF-5A* isoform 1 or 2 with a *C*-terminal His₆-tag followed by a Ser-Ala-*StreptII*-tag. Constructs not to scale.

5.5. Determination of binding kinetics for the different *SbDHS* isoforms by bio-layer interferometry

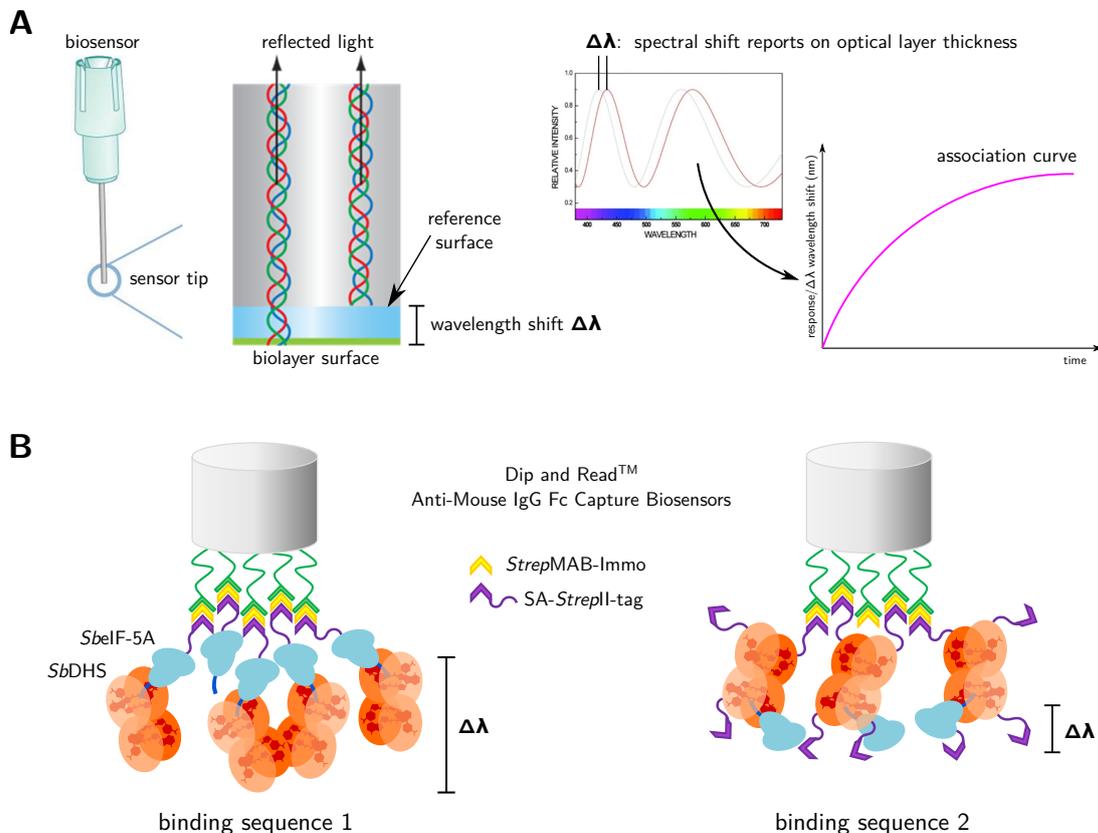


Figure 5.18.: Schematic representation of the BLI principle and the strategy for the *SbDHS*/eIF-5A binding kinetics measurements

A: Bio-layer interferometry (BLI) measurement principle: White light is conducted through an optical fiber to the tip of the biosensor, where it is reflected at the surface of the bio-layer on top of the tip and at an internal reference surface. This results in a wavelength shift $\Delta\lambda$ between the two differentially reflected light portions, which generates a distinct interference pattern for all wavelengths of the applied white light. From this interference pattern, the thickness of the layer between the reference surface and the bio-layer surface can be calculated with high fidelity. Upon binding of molecules to the bio-layer, this thickness increases – which can be visualized in a time-dependent manner as an association curve. The dissociation of molecules would in turn result in a decrease of the bio-layer thickness and a dissociation curve (not depicted). The figure was prepared from images of informative literature of PALL *fortéBIO*, the manufacturer of the Octet Red96e system. **B:** Possible measurement orientations for *SbDHS*/eIF-5A binding kinetics: In a first binding step, the *StrepMAB-Immo* antibodies (iba) are bound to the Dip and Read™ Anti-Mouse IgG Fc Capture Biosensors (AMC; PALL *fortéBIO*) via their Fc portion. Subsequently, either SA-*StreptII*-tagged *SbElF-5A* (binding sequence 1) or *SbDHS* (binding sequence 2) are specifically bound as ligands to the *StrepMAB-Immo* antibodies. The binding between the ligand and the corresponding analyte – *SbDHS* or *SbElF-5A*, respectively – can then be determined. The graphics were generated with Inkscape™ (version 0.92, <https://inkscape.org/>).

biosensors are layered with covalently bound antibodies that specifically recognize the Fc portion of mouse-derived antibodies – in this case that of *StrepMAB-Immo*.

For the stable binding of the chosen ligand to the immobilized *StrepMAB-Immo*, the ligand has to carry the SA*StrepII*-tag. Therefore, the existing *SbDHS* and eIF-5A constructs that already contained an *N*- or *C*-terminal *StrepII*-tag, respectively, were modified by insertion of a short serine-alanine coding sequence in front of the respective *StrepII*-tag (see sections 7.9.4 and 7.9.2 for detailed cloning strategies). In figure 5.17, the resulting constructs are depicted schematically.

5.5.1. Binding characteristics and kinetics of *SbDHS* homo/paralogs in basic glycine buffer

First BLI measurements to determine the binding kinetics between the three *SbDHS* and their substrates eIF-5A1 and eIF-5A2 were done in glycine-NaOH buffer (100 mM glycine, 150 mM NaCl, 1 mM NAD⁺, pH 9.0) to mimic the conditions of the activity assays².

The results of some of these measurements are depicted in figure 5.19. Initially, the Dip and Read™ Anti-Mouse IgG Fc Capture Biosensors (AMC; PALL *forté*BIO) were loaded with *StrepMAB-Immo* (iba) at a concentration of 10 mg L⁻¹ in PBS buffer (10 mM NaH₂PO₄, 140 mM NaCl, 0.02 % (v/v) Tween® 20, 1 % (w/v) BSA, pH 7.5). As can be seen in figure 5.19 **A**, the binding signal of the antibody was homogeneous and produced a wavelength shift of about 1.5 nm. While the *StrepMAB-Immo* antibody bound stably to the antibody immobilized on the biosensor in PBS buffer (figure 5.19 **A c**; further data not shown), the *StrepMAB-Immo* antibody unfortunately dissociated from the biosensor in glycine-NaOH buffer at pH 9.0, which was used for all subsequent steps (**A d**). Although this is not an ideal situation for the measurement of binding kinetics, measurement data can still be interpreted if a reference sensor (see e. g. **B** "no ligand") is used in parallel and subtracted during data analysis.

Following the association of the *StrepMAB-Immo* antibody, the SA*StrepII*-tagged ligands were loaded at a concentration of 10 mg L⁻¹ in glycine-NaOH buffer (**B–E**). All three *SbDHS* as well as both *Sb*eIF-5A ligands exhibited stable binding to *StrepMAB-Immo* (see figure 5.19 **B–E e** for *SbDHS2* and *Sb*eIF-5A1; data not shown for all other ligands). Due to their considerably smaller size compared to DHS, the eIF-5A1 ligand bound much faster, but only produced a wavelength shift of about 0.2 nm as compared to a shift of at least 0.5 nm for DHS2 ligand (**B, C**). As the "no ligand" reference sensorgram confirmed, ligand binding was specific at the applied concentrations and the glycine-NaOH buffer did not cause unspecific signals.

After ligand loading, a baseline was established (**f, g**) prior to analyte association. In these first measurements, the analytes were diluted to 100 nmol L⁻¹ (**B, C, E**) or 10 μmol L⁻¹ (**D**) for the association step. At 100 nmol L⁻¹, *SbDHS2* (**B, E**) and *SbDHS1* (**E**) exhibited the expected binding curve upon association to the immobilized *Sb*eIF-5A1 and *Sb*eIF-5A2 (data not shown). In contrast, *SbDHS3* (HSS, **E**) did not show any detectable binding signal at this

²Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

5.5. Determination of binding kinetics for the different *SbDHS* isoforms by bio-layer interferometry

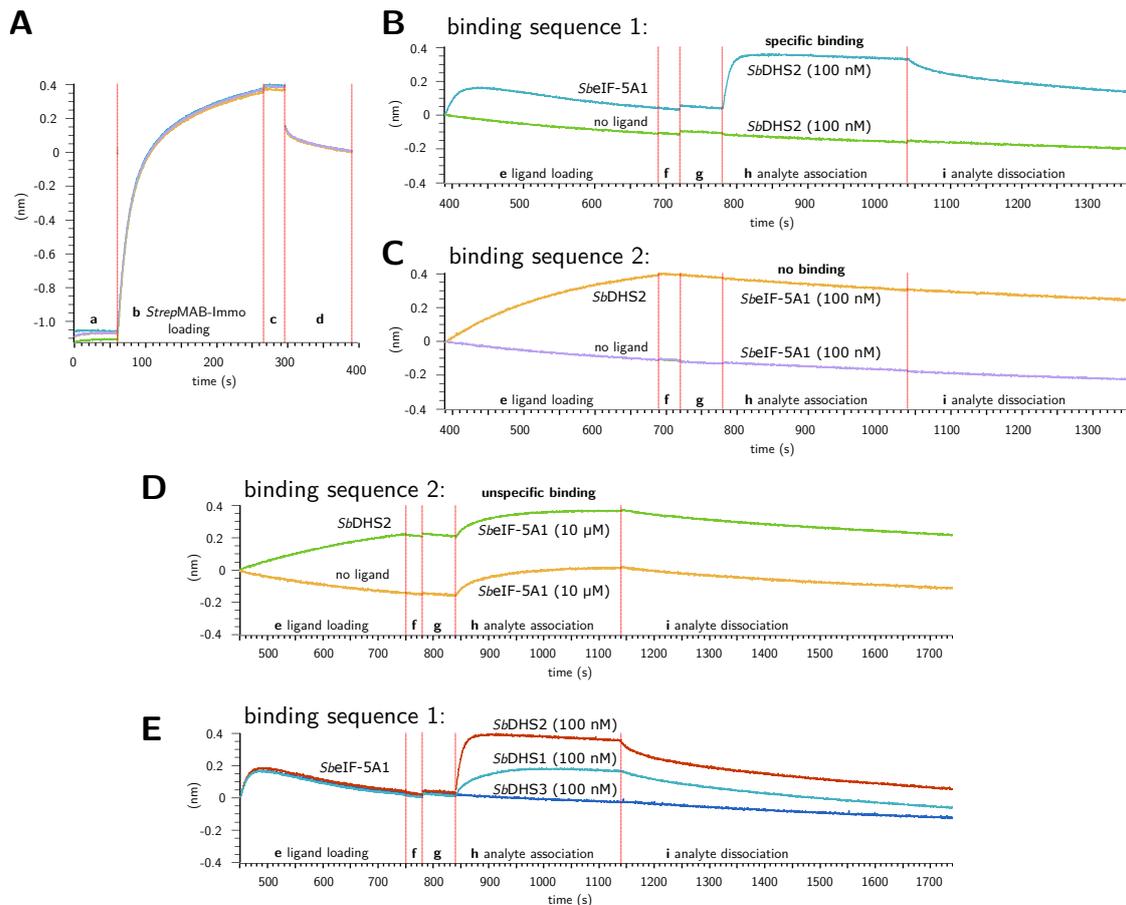


Figure 5.19.: Binding behavior of *S. bicolor* DHS and eIF-5A1 in basic glycine buffer

A a: Equilibration of Dip and ReadTM Anti-Mouse IgG Fc Capture Biosensors (AMC; PALL *forté*BIO). **b:** *StrepMAB-Immo* (iba) was loaded at a concentration of 10 mg L^{-1} in a PBS buffer (10 mM NaH_2PO_4 , 140 mM NaCl, 0.02% (v/v) Tween[®] 20, 1% (w/v) BSA, pH 7.5). Baseline steps in **c** PBS buffer and **d** glycine-NaOH (100 mM glycine, 150 mM NaCl, 1 mM NAD^+ , pH 9.0). **B and C:** Comparison of binding sequences 1 and 2 (see figure 5.18). **D:** Test for specific binding of analyte eIF-5A1 to ligand DHS2 at $10 \mu\text{mol L}^{-1}$. **E** Comparison of DHS1, DHS2 and DHS3 as analytes for ligand eIF-5A1. **B-E:** Ligands *SAS**StrepII-SbDHS2-His*₆ or *SbeIF-5A1-His*₆-*SAS**StrepII* were loaded at a concentration of 10 mg L^{-1} in glycine-NaOH buffer. **f, g:** Baseline steps in glycine-NaOH-buffer. **h:** Association of analytes *SbeIF-5A1-His*₆ or *SbDHS1-His*₆, *SbDHS2-His*₆ and *SbDHS3-His*₆ at 100 nmol L^{-1} in glycine-NaOH buffer, except for **D**. **i:** Dissociation of analytes in glycine-NaOH buffer. Measurements were performed with the Octet Red96e system and data was prepared with the Data Analysis HT software (version 10.0.0.48; PALL *forté*BIO).

concentration. At a higher concentration of $10 \mu\text{mol L}^{-1}$, the binding of all *SbDHS* analytes to the *SbeIF-5A* ligands became unspecific, which resulted in the same binding signals for sensors with and without ligands (data not shown, but comparable to the situation in **C**).

Interestingly, no binding signal could be detected for both *SbeIF-5A* analytes at 100 nmol L^{-1} binding to either *SbDHS1* or *SbDHS2* ligand (see **C** for *SbeIF-5A1* analyte and *SbDHS2* lig-

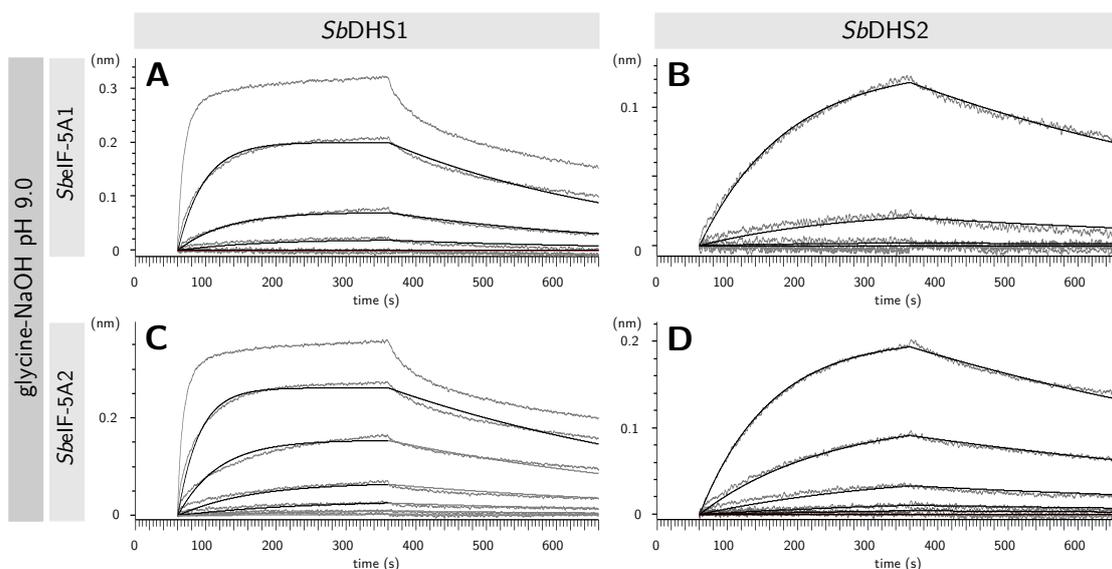


Figure 5.20.: Binding kinetics of *S. bicolor* eIF-5A homologs as ligands and DHS homologs as analytes in glycine-NaOH buffer pH 9.0

BLI determination of binding kinetics of *SbDHS1* and *SbDHS2* binding to either *SbelF-5A1* or *SbelF-5A2* in glycine-NaOH buffer (100 mM glycine, 150 mM NaCl, 1 mM NAD^+ , pH 9.0). Depicted are association and dissociation curves (gray lines). Ligands were loaded at a concentration of 10 mg L^{-1} . Analytes were supplied as 1:2 dilution series with samples in the range of 3.13–200 nM (*SbDHS1*; **A, C**) and 0.39–25 nM (*SbDHS2*; **B, D**). Data was fitted globally to a 1:1 binding model after subtraction of a reference signal, with the fitted curves indicated by black lines. Measurements were performed with the Octet Red96e system and data was prepared and analyzed with the Data Analysis HT software (version 10.0.0.48; PALL *forté*BIO).

and; data not shown for all other ligands and analytes). For an analyte concentration of $10 \mu\text{mol L}^{-1}$, the resulting binding of both *SbelF-5A* to either *SbDHS1* or *SbDHS2* was un-specific (see **D** for *SbelF-5A1* analyte and *SbDHS2* ligand, data not shown for all other ligands and analytes). Obviously, only one orientation – here termed binding sequence 1 (see also figure 5.18) – yielded interpretable binding data. The following experiments were therefore performed in this orientation.

Based on these results, a full kinetic profile was measured for *SbelF-5A1* and *SbelF-5A2* as ligands with *SbDHS1* and *SbDHS2* as analytes. Therefore, SA*StreptII*-tagged ligands were loaded onto biosensors prepared with *StreptMAB*-Immobilized at a concentration of 10 mg L^{-1} in glycine-NaOH buffer pH 9.0. Binding kinetics were measured for 7 concentrations of a 1:2 dilution series of *SbDHS1* (3.13–200 nM) and *SbDHS2* (0.39–25 nM). In figure 5.20 **A–D**, the resulting association and dissociation curves are depicted for each binding pair.

The binding curves of *SbDHS2* exhibit a nearly ideal exponential shape for the association as well as the dissociation step. Since literature reports a 1:1 stoichiometry in the DHS–eIF-5A complexes (see section 4.2.4 for further details), this is in accordance with the expected 1:1 binding between a DHS tetramer and the eIF-5A monomer. However, the binding curves of *SbDHS1* exhibit a small heterogeneous component, which seems to become more apparent for

higher concentrations. This might be due to unspecific binding caused by a very high analyte concentration in relation to the K_D (see below). The highest analyte concentration was therefore excluded from the corresponding fits (**A**, **C**). To account for the unstable binding of *StrepMAB-Immo* to the biosensor in the basic glycine buffer, the data was reference subtracted prior to the global fit to a 1:1 (Langmuir) binding model. For the measurement of *SbDHS2* with *SbelF-5A1*, the three lowest analyte concentrations were omitted from the fit due to poor signal quality around the detection limit.

For the 1:1 binding model, a global fit of the following exponential integrated rate equation to the association curves is performed – an approach introduced by O’Shannessy in 1993 [196]:

$$R_t = E - Ee^{-k_s t}, \quad (5.1)$$

with $E = \frac{k_a C R_{max}}{k_a C + k_d}$ and $k_s = k_a C + k_d$. R_t is the response at a time t , C is the analyte concentration, k_a the association and k_d the dissociation constant. The K_D is calculated from k_a and k_d as $K_D = \frac{k_d}{k_a}$.

For the fit of the dissociation phase, the following exponential integrated rate equation is applied, where R_0 and t are based on t_0 as the beginning of the dissociation phase:

$$R_t = R_0 e^{-k_d t}. \quad (5.2)$$

Table 5.1 shows the data analysis results. The affinity constants for the binding between the two *SbDHS* and both eIF-5A in a glycine-NaOH buffer at pH 9.0 are quite high, located in the low nanomolar range. Based on these results, eIF-5A2 binds slightly tighter to both DHS than eIF-5A1, with higher association and lower dissociation constants. In addition, *SbDHS2* exhibits higher affinities towards both eIF-5A than *SbDHS1* – again with higher association and lower dissociation constants. As a result, the highest affinity is established between *SbDHS2* and eIF-5A2 with 3.66 nM, and the lowest for *SbDHS1* and eIF-5A1 with 11.3 nM. However, the accuracy of the exact binding data highly depends on the accuracy of the analyte concentrations. Since the protein concentrations were not determined absolute, but calculated from 280 nm absorption with a theoretical extinction coefficient based on amino acid sequence, the exact affinity values should not be over-interpreted. Especially the small differences between *SbDHS1* and *SbDHS2* might arise from protein concentration inaccuracies. In addition, it should be noted that these measurements were only performed once.

Nevertheless, while the exact constants have to be interpreted with care, the overall low nanomolar affinities most likely reflect the actual binding situation. Based on these findings, the heterogeneous binding behavior in *SbDHS1* measurements (**A**, **C**) at high analyte concentrations is no surprise. In general, optimal kinetic measurements with the BLI should be performed in a concentration range between 0.1x and 10x of the expected K_D value as proposed by the manufacturer, a range that was exceeded by at least a factor of 2.

Table 5.1.: Binding kinetics of *S. bicolor* DHS1 and DHS2 to eIF-5A1 and eIF-5A2 in glycine-NaOH buffer

Ligand	Analyte	K_D (nM)	k_a (1/Ms)	k_d (1/s)
eIF-5A1	DHS1	11.3 [± 0.5 %]	2.42×10^5 [± 0.4 %]	2.72×10^{-3} [± 0.3 %]
eIF-5A1	DHS2	6.02 [± 0.7 %]	2.63×10^5 [± 0.6 %]	1.58×10^{-3} [± 0.4 %]
eIF-5A2	DHS1	6.16 [± 0.4 %]	3.13×10^5 [± 0.4 %]	1.93×10^{-3} [± 0.3 %]
eIF-5A2	DHS2	3.66 [± 0.3 %]	3.48×10^5 [± 0.3 %]	1.27×10^{-3} [± 0.3 %]

Affinity constant $K_D = k_d/k_a$; k_a : association constant; k_d : dissociation constant. Constants were derived from a data fit based on the 1:1 Langmuir binding model performed with the Data Analysis HT software (version 10.0.0.48; PALL fortéBIO). The percentage in square brackets reports the standard error of the corresponding fits. Buffer condition is specified in figure 5.20.

5.5.2. Binding characteristics and kinetics of *SbDHS* homo/paralogs in PBS at physiological pH

In a second series of experiments, the kinetic profiles for *Sb*eIF-5A1 and *Sb*eIF-5A2 – as SA*StreptII*-tagged ligands – with *SbDHS1* and *SbDHS2* as analytes were determined in a physiological PBS buffer (10 mM NaH₂PO₄, 140 mM NaCl, pH 7.5) with additional 0.02 % (v/v) Tween[®] 20 and 1 % (w/v) BSA to reduce unspecific binding. *Sb*eIF-5A ligands were loaded at a concentration of 10 mg L⁻¹, while the analytes were diluted in a 1:2 series from 100–1.55 nM for *SbDHS2* and either from 50–1.55 nM or from 25–0.78 nM for *SbDHS1*. The experiments were at least repeated three times with regenerated biosensors (see section 7.14.8 for details).

Representatives of the resulting binding curves are depicted in figure 5.21. For **A**, three repetitions of the kinetic measurements with regenerated biosensors are depicted in **B_{i-iii}**. They show that the data obtained from regenerated biosensors (**B_{ii}**, **B_{iii}**) is comparable to that from fresh biosensors (**B_i**).

Similar to the *SbDHS1* measurements in the previous experiment, all curves exhibit a slightly heterogeneous binding behavior. While the heterogeneous component is not as prominent in the measurements with *SbDHS2* as analyte, the curves of *SbDHS1* are clearly not represented by the applied 1:1 binding model, which was used to fit the data. Especially the binding curves of *SbDHS1* to *Sb*eIF-5A2 (**C**) do not exhibit the expected exponential association and dissociation curves. For *SbDHS1* to *Sb*eIF-5A2 (**C**), odd curve shape and spacing between the curves of certain concentrations was observed even after three repetitions with freshly prepared analyte dilutions (data not shown). Since the other measurements with the same charges of *SbDHS1* and *Sb*eIF-5A2 did not show comparable discrepancies – besides the heterogeneous behavior –, it was assumed there was a problem with the corresponding sensors, and the respective curves (6.125 and 0.78 nM) were excluded from analysis. Unfortunately, this measurement could not be repeated with fresh biosensors in the course of this thesis.

Table 5.1 compiles the results of a preliminary data analysis. A reference was subtracted prior to data fitting. Considering all previous experimental data, it can be assumed that one monomeric eIF-5A ligand (see section 5.6 for SEC-MALS data) binds one DHS tetramer. Although the DHS tetramer provides four identical active sites, DHS–eIF-5A complexes were

5.5. Determination of binding kinetics for the different *SbDHS* isoforms by bio-layer interferometry

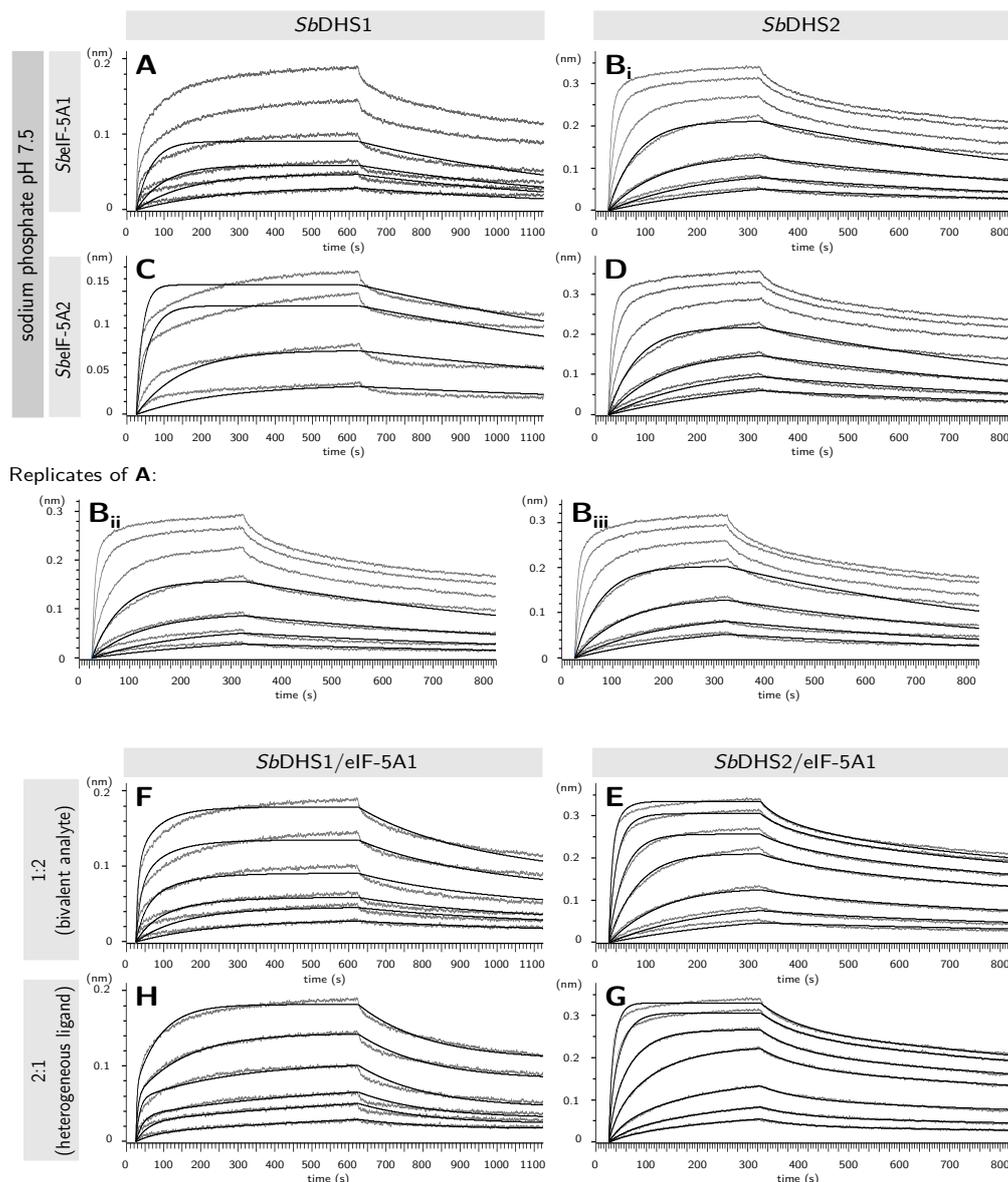


Figure 5.21.: Binding kinetics of *S. bicolor* eIF-5A homologs as ligands and DHS homologs as analytes in PBS buffer at pH 7.5

BLI determination of binding kinetics of *SbDHS1* and *SbDHS2* binding to either *SbelF-5A1* or *SbelF-5A2* in a PBS buffer (10 mM NaH_2PO_4 , 140 mM NaCl, 0.02 % (v/v) Tween[®] 20, 1 % (w/v) BSA, pH 7.5). Depicted are association and dissociation curves (gray lines). Ligands were loaded at a concentration of 10 mg L^{-1} . Analytes were supplied as 1:2 dilution series with samples in the range of 1.55–100 nM (*SbDHS2*; **B**, **D**), 1.55–50 nM (*SbDHS1*; **A**) and 0.78–25 nM (*SbDHS1*; **C**). Data was fitted globally to a 1:1 binding model (**A–D**), a 1:2 bivalent analyte model (**E**, **F**) and a 2:1 heterogeneous ligand model (**G**, **H**) after subtraction of a reference signal, with the fitted curves indicated by black lines. **E** and **G**, as well as **F** and **H** show the same data as **A** and **B**, respectively. **B_{i-iii}**: Exemplary measurement repetitions for **B** to show data reproducibility. Measurements were performed with the Octet Red96e system and data was prepared and analyzed with the Data Analysis HT software (version 10.0.0.48; PALL *forte*BIO).

Table 5.2.: Binding kinetics of *S. bicolor* DHS1 and DHS2 to eIF-5A1 and eIF-5A2 in PBS

Ligand	Analyte	Model	K_D (nM)	k_a (1/Ms)	k_d (1/s)	K_{D2} (nM)	k_{a2} (1/Ms)	k_{d2} (1/s)
eIF-5A1	DHS1	1	0.67 [± 3.2 %]	2.56×10^6 [± 11.3 %]	1.71×10^{-3} [± 12.2 %]	-	-	-
eIF-5A1	DHS2	1	0.87 [± 1.9 %]	1.51×10^6 [± 4.7 %]	1.07×10^{-3} [± 4.6 %]	-	-	-
eIF-5A2	DHS1	1	0.29 [± 1.4 %]	3.01×10^6 [± 13.3 %]	8.72×10^{-4} [± 14.5 %]	-	-	-
eIF-5A2	DHS2	1	0.75 [± 9.5 %]	1.84×10^6 [± 19.4 %]	1.30×10^{-3} [± 8.4 %]	-	-	-
eIF-5A1	DHS1	2	0.32 [± 2.2 %]	7.55×10^5 [± 0.4 %]	2.43×10^{-4} [± 2.2 %]	0.53 [± 2.0 %]	2.46×10^7 [± 1.9 %]	1.30×10^{-2} [± 0.7 %]
eIF-5A1	DHS2	2	0.60 [± 0.6 %]	8.68×10^5 [± 1.3 %]	5.23×10^{-4} [± 0.6 %]	3.22 [± 1.3 %]	4.01×10^6 [± 1.1 %]	1.29×10^{-2} [± 0.6 %]

Affinity constant $K_D = k_d/k_a$ of the individual binding kinetics; k_a , k_{a2} : association constants; k_d , k_{d2} : dissociation constants. Constants were derived from a data fit based on the 1:1 Langmuir binding model (1) or the 2:1 heterogeneous ligand model (2) performed with the Data Analysis HT software (version 10.0.0.48; PALL *forté*BIO). The percentage in square brackets reports the standard error of the contributing fits. For model 1 at least 3 measurements were averaged (eIF-5A1–DHS2: n=5; eIF-5A1–DHS1, eIF-5A2–DHS1, eIF-5A2–DHS2: n=3). For model 2, data analysis was performed exemplarily with one dataset each. Buffer condition is specified in figure 5.21.

reported to exhibit 1:1 stoichiometry (see section 4.2.4 for further details), so that one DHS tetramer will also only bind one eIF-5A monomer.

Although the curves' shape indicates heterogeneous binding behavior, a first analysis was performed using the 1:1 Langmuir binding model (see above), which theoretically describes the putative biological situation best. To minimize the effects of heterogeneity due to unspecific binding from higher analyte concentrations, these were excluded from the fit – except for the *SbDHS1*–*Sb*eIF-5A2 measurements. Similar to the kinetic constants derived from the data obtained in basic glycine-NaOH buffer, the calculated K_D values reflect a high affinity between both *S. bicolor* eIF-5A and its two DHS, with values even slightly below 1 nM (see table 5.1).

5.5.3. Identification of the cause for heterogeneous binding behavior

In the Data Analysis HT software (version 10.0.0.48; PALL *forté*BIO) applied for data analysis, three more complex fitting models are implemented besides the simple 1:1 binding model. While one model accounts for mass transport effects, the other two deal with analyte (1:2 bivalent analyte model) and ligand heterogeneity (2:1 heterogeneous ligand model).

Mass transport effects arise from the depletion of analyte molecules in the direct vicinity of the ligand surface due to the fast binding of analyte molecules to the ligand. If the diffusion and convection of analyte molecules – the "transport" – to the ligand surface is slower than the association rate, these transport processes become association rate limiting and mask the actual association rate. However, in the experiments described here, the 96-well plates containing the analyte were shaken at 1000 rpm to prevent mass transport effects. Moreover, the curves' shapes did not exhibit features typical for mass transport effects and the mass transport model did not fit the data well (data not shown). Therefore, the other two models were also tried to see if they are able to represent the data better than the 1:1 model.

The 1:2 bivalent analyte binding model assumes that the analyte (a DHS in these experiments) provides two independent binding sites for the ligand (eIF-5A). Although experimental evidence from literature proposes a 1:1 stoichiometry (see section 4.2.4 for further details), a parallel binding at two sites might not be completely out of the question. However, the 1:2 model only fitted the data slightly better than the 1:1 model (see figure 5.21 E and F for representative fits on the data from A and B).

In contrast, the 2:1 heterogeneous ligand binding model was able to fit the actual data quite well, except for the highest analyte concentrations (see figure 5.21 G and H for representative fits on the data from A and B). In table 5.2, the kinetic constants calculated based on these fits are also summarized. They contain the contributions of two putative binding events with affinity constants in the low nanomolar and high picomolar range, respectively.

The heterogeneous ligand model represents essentially three situations, which cannot be easily differentiated: 1) two-step conformational change, 2) heterogeneous ligand and 3) bivalent ligand. As mentioned above, literature reports a 1:1 stoichiometry for the DHS–eIF-5A interaction, although the DHS provides four seemingly identical active sites. It was therefore proposed that a conformational change upon binding of the first eIF-5A prevents the binding of further eIF-5A molecules. This would agree with the two-step conformational change model and could explain the heterogeneity that is more or less present in all binding curves measured in the course of this thesis.

The two other models – assuming a heterogeneous or bivalent ligand with two different or identical binding sites per molecule, respectively –, do not readily agree with the theoretical binding situation, since especially the ligand situation in these assays is well defined: The *Sb*eIF-5A isoforms are monomeric and highly pure as determined by SDS-PAGE and DLS (data not shown). In addition, the immobilization via the SA*StreptII*-tag provides specifically bound and homogeneously oriented ligands on the biosensor surface – in contrast to other immobilization methods like amine coupling for example, which is not specific for a certain ligand and utilizes the *N*-terminal amine group and the ϵ -amine groups of lysine residues for immobilization. This can cause an inhomogeneous orientation of ligand molecules on the sensor surface, resulting in highly accessible and less accessible binding sites due to different ligand orientations. However, the affinity-tag-based approach is especially well suited to prevent these effects by immobilizing the ligand via a single specific binding site.

Table 5.3.: Relationships of the sums and products of σ and γ and the individual values of σ to the analyte concentration C and the k_a and k_d values of the three biphasic models

Model	$\sigma_1 + \sigma_2$	$\sigma_1\sigma_2$	σ_1 and σ_2	$\gamma_1 + \gamma_2$	$\gamma_1\gamma_2$
1	$k_{a1}C + k_{d1} + k_2 + k_{-2}$ linear in C	$k_{a1}(k_2 + k_{-2})C + k_{d1}k_{-2}$ linear in C	non-linear in C	$k_{d1} + k_2 + k_{-2}$	$k_{d1}k_{-2}$
2	$(k_{a1} + k_{a2})C + k_{d1} + k_{d2}$ linear in C	$k_{a1}k_{a2}C^2 + (k_{a1}k_{d2} + k_{d1}k_{a2})C + k_{d1}k_{d2}$ quadratic in C	linear in C	$k_{d1} + k_{d2}$	$k_{d1}k_{d2}$
3	$(k_{a1} + k_{a2})C + k_{d1} + k_{d2}$ linear in C	$2k_{a1}k_{a2}C^2 + k_{a1}k_{d2}C + k_{d1}k_{d2}$ quadratic in C	non-linear in C	$k_{d1} + k_{d2}$	$k_{d1}k_{d2}$

Model 1: conformational change; model 2: heterogeneous ligand; model 3: bivalent ligand. k_a : association constant; k_d : dissociation constant; k_2 and k_{-2} : constants of the conformational change. The table was taken from [272].

Since it is not possible to derive the correct binding model from these considerations alone, the obtained data was further analyzed. In 2015, Tiwari et al. reported an analytical method to distinguish between these three models based on binding kinetic data [272], which is applied in the following to the above-described data.

The heterogeneous ligand model assumes that both binding events – or a binding and a conformational change –, each with their own k_a and k_d constants, contribute to the measured response, and that the measured curves correspond to a linear combination of these two contributions. This assumption is represented by the following double-exponential equation for the association curve:

$$R_t = E - Ee^{-\sigma_1 t} + F - Fe^{-\sigma_2 t}, \quad (5.3)$$

and the equation below for the dissociation phase:

$$R_t = Ee^{-\gamma_1 t} + Fe^{-\gamma_2 t}. \quad (5.4)$$

σ and γ are complex constants that contain information on the analyte concentration (σ) and the k_a and k_d values of the two contributing events (σ , γ), and they depend on the actual binding model. In table 5.3 the relationships between these values are summarized for the three above-mentioned binding models. Based thereon, the plot of $\sigma_1 + \sigma_2$, $\sigma_1\sigma_2$ and the individual σ_1 and σ_2 against the analyte concentration C can provide information on the underlying binding model.

Two binding kinetics were fitted with the double-exponential equation 5.3 for each of the four interactions. In figure 5.22, the σ values of one of these fitted kinetics per interaction are plotted against the corresponding analyte concentration C as described above. In addition, simulated curves based on the double-exponential fits are depicted to visualize the contributions of the two exponential parts (light and dark gray dots) to the overall curve shape (black dots) for four different analyte concentrations.

The obtained σ values did neither exhibit clear linear nor quadratic behavior, but are rather scattered and incoherent, except for the interaction of *SbDHS1* with eIF-5A2. Especially the values calculated at low analyte concentrations and with a low contribution of the corresponding exponential component to the overall curve shape (e. g. simulated curves of *SbDHS2* and both eIF-5A, light gray dots) could not be determined with certainty, since large changes in σ only resulted in minor improvements to the overall fit. For seven of the eight analyses, no direct correlation between σ values and analyte concentration C could be established.

For the interaction of *SbDHS1* with eIF-5A2, however, $\sigma_1 + \sigma_2$ and $\sigma_1\sigma_2$ as well as σ_1 seem to be linearly dependent on analyte concentration C , while σ_2 is not. Based on these results, the conformational change model (1) would have to be applied (see table 5.3). Consequently, the corresponding kinetic constants were calculated as described in Tiwari et al. 2015 [272]: The k_{a1} of the initial binding reaction is $5.04 \times 10^{-6} \text{ M}^{-1} \text{ s}^{-1}$ and its k_{d1} is $6.93 \times 10^{-3} \text{ s}^{-1}$, the constants of the conformational change are $k_2 = 5.76 \times 10^{-3} \text{ s}^{-1}$ and $k_{-2} = 5.09 \times 10^{-4} \text{ s}^{-1}$. The overall K_D of binding and conformational change is 0.11 nM, which is in the range of the K_D value determined by the 1:1 binding model (see table 5.2).

In addition, the simulated curves reveal another aspect of the heterogeneous binding curves. For *SbDHS2* binding to both eIF-5A, it seems as if the contribution of a faster binding

5.5. Determination of binding kinetics for the different *SbDHS* isoforms by bio-layer interferometry

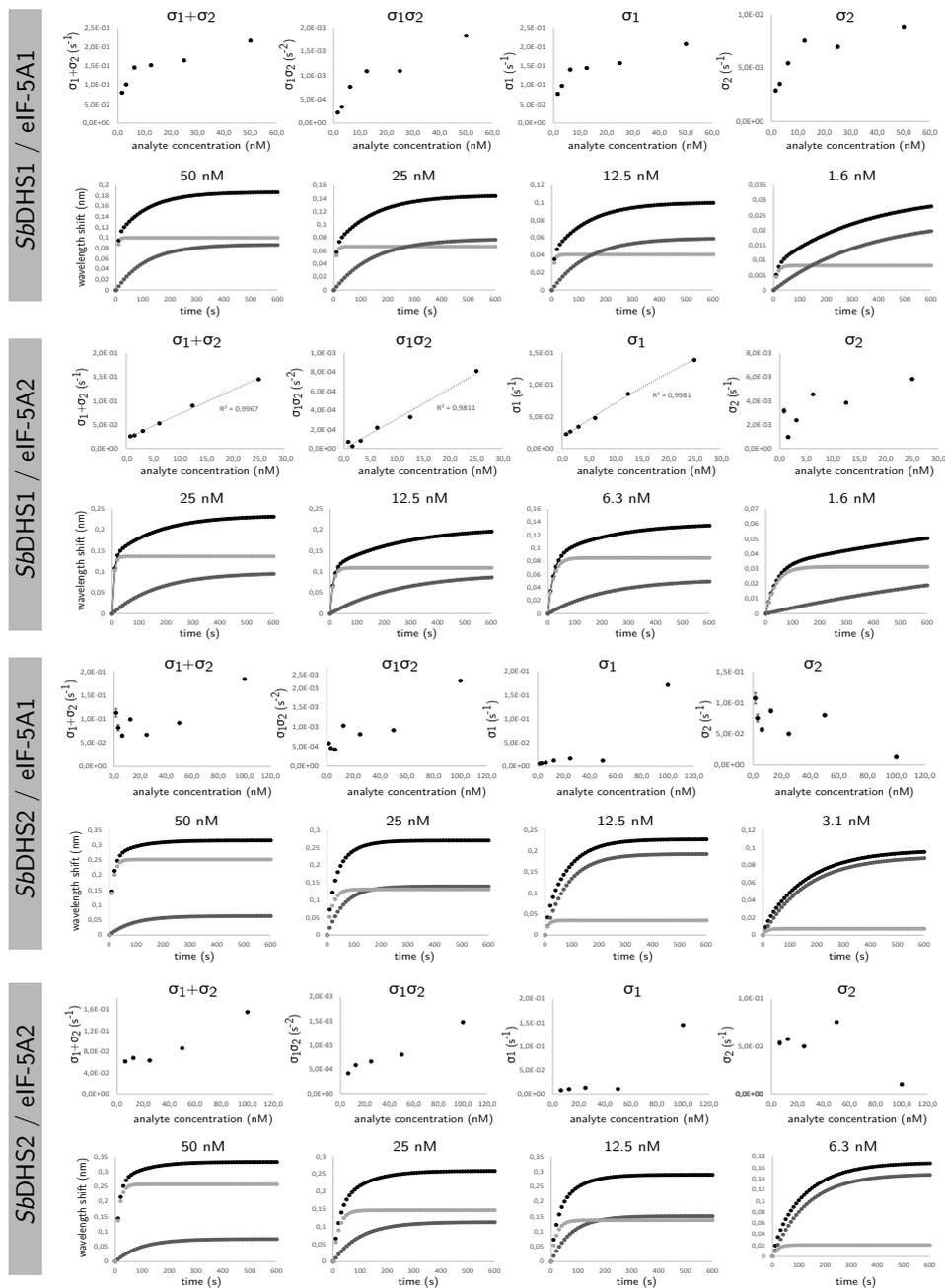


Figure 5.22.: Determination of binding model based on the analysis of a double-exponential fit according to Tiwari et al. 2015 [272]

Plot of $\sigma_1 + \sigma_2$, $\sigma_1\sigma_2$ and the individual σ_1 and σ_2 over the analyte concentration C for binding kinetics of either *SbELF-5A1* or *SbELF-5A2* with either of the two *SbDHS*. Error bars report the standard error of the corresponding fit. If possible, a linear fit of the data was performed (dashed line), with the corresponding coefficient of determination denoted beside. In addition, simulated binding curves created from the results of the fitting procedure represent the double-exponential curve (black dots) and the two exponential contributions therein, as determined by fitting to equation 5.3 (dark and light gray dots), for four different analyte concentrations. Data fitting was performed with OriginPro[®] 8 SR0 (version 8.0724, OriginLab Corporation).

component – if it is indeed a binding event – increases with analyte concentration (light gray dots), while a slower binding component (dark gray dots) decreases. This could hint in the direction of a problem derived from too high analyte concentrations. However, for *SbDHS1* this behavior was not observed.

Taken together, the obtained results indicate a high affinity in the low nanomolar or even high picomolar range of both DHS to both eIF-5A in PBS buffer at pH 7.5 without significant differences. The binding curves exhibit heterogeneous binding behavior. Based on the performed data analyses described above, the most likely cause for the heterogeneity in the binding behavior is a high analyte concentration in correlation to the low affinity, which often leads to heterogeneous effects in binding. In addition, a high concentration of immobilized eIF-5A molecules on the biosensor surface might also contribute to a perceived ligand heterogeneity. However, further influences from a putative underlying non-1:1 binding model could neither be proved nor excluded.

5.5.4. The influence of NAD^+ and the inhibitor GC7 on the interaction of *Sb*eIF-5A and DHS

The DHS enzyme is dependent on the cofactor NAD^+ – and the substrate spermidine – to perform the deoxyhypusination reaction on eIF-5A. As already demonstrated by DLS and DSF experiments in the course of this thesis (see sections 5.1.2 and 5.4), NAD^+ has a stabilizing effect on *SbDHS* and can prevent self-aggregation during purification and upon storage. The inhibitor GC7, a spermidine analog, exhibited even stronger stabilizing properties in thermal shift (DSF) experiments. Therefore, the influence of these two substances on the interaction of *SbDHS* with eIF-5A was also examined by BLI.

In preliminary experiments, the effect on the dissociation phase was analyzed. In figure 5.23 **A** and **B**, the sensorgrams for the interaction of *SbDHS2* with *Sb*eIF-5A2 and their dissociation behavior in the presence (or absence) of different concentrations of NAD^+ and/or GC7 are depicted. In **A**, it can easily be seen that the presence of low millimolar concentrations of GC7 alone significantly increases the dissociation rate (red and light/dark blue lines) by a factor of about 3 for 2 mM GC7 in contrast to the situation without inhibitor (green line). This effect is even enhanced by the presence of NAD^+ , as the direct comparison of a dissociation in 1 mM GC7 (**B**: yellow line) with a dissociation in 1 mM of both, GC7 and NAD^+ (**B**: green line), impressively shows. Here, the dissociation rate is nearly doubled (from about $8.5 \times 10^{-3} \text{ s}^{-1}$ to $14.9 \times 10^{-3} \text{ s}^{-1}$) by the additional presence of 1 mM NAD^+ . This is probably due to the fact that the binding of GC7 to the active sites of the DHS tetramers most likely depends on the prior binding of NAD^+ , so that the supplementation of 1 mM NAD^+ enhances the fraction of NAD^+ containing active sites available for GC7 binding. The addition of 5 mM GC7 and NAD^+ increased the dissociation rate even further by a factor of 8 to about $37.0 \times 10^{-3} \text{ s}^{-1}$, compared to the situation without either component.

Actually, based on these results, a dissociation step in a buffer containing the combination of 5 mM GC7 and NAD^+ was implemented for biosensor regeneration (see section 7.14.8). Prior experiments with classical regeneration conditions (e. g. low pH glycine buffer, high salt buffers) – to strip the secondary *Strep*MAB-Imm and the bound ligands and analytes off the immobilized primary anti-Mouse IgG antibody on the biosensor – only resulted in their

aggregation on the sensor tip (data not shown). With the combination of NAD^+ and GC7, however, the surface composed of *Strep*MAB-Imm and the bound eIF-5A ligand stayed intact and unchanged, while the bound DHS analyte was effectively removed.

Interestingly, the combined effect of GC7 and NAD^+ on the dissociation rate is in stark contrast to the effect of 1 mM NAD^+ alone: While the dissociation rate is significantly increased in combination with GC7, NAD^+ alone stabilizes the interaction of the DHS to eIF-5A, thereby drastically slowing down the dissociation process to about 2% of the original dissociation rate (B: red line).

However, in contrast to these results for the dissociation phase, the addition of up to 2 mM NAD^+ to the tested *SbDHS2* analyte solutions did not result in any significant change in the binding behavior (data not shown). Since all purified DHS samples – including the samples for BLI measurements – were supplied with 2 mM NAD^+ prior to storage to prevent aggregation (see section 5.1.2), the absence of a detectable effect might be due to the fact that most DHS molecules were already saturated with NAD^+ .

The effect of similar concentrations of the inhibitor GC7 on the association phase was much more drastic: At a concentration of 1 mM, it completely prevented binding of *SbDHS2*

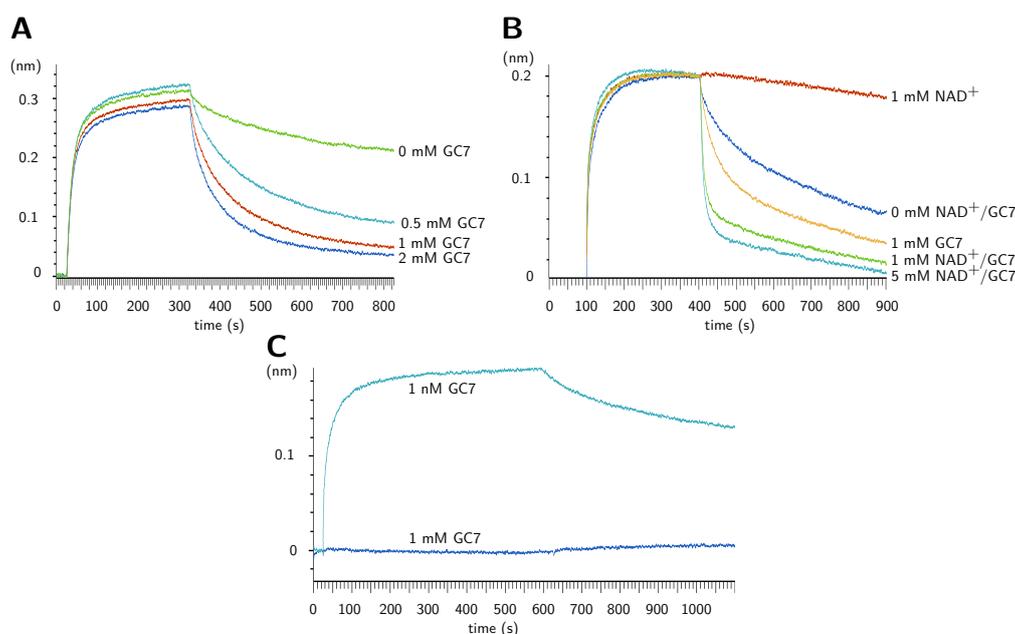


Figure 5.23.: Association or dissociation of *S. bicolor* DHS2 to/from eIF-5A2 in the presence of NAD^+ and/or the inhibitor GC7

BLI determination of *SbDHS2* binding to *Sb*eIF-5A2 in a PBS buffer (10 mM NaH_2PO_4 , 140 mM NaCl, 0.02% (v/v) Tween[®] 20, 1% (w/v) BSA, pH 7.5), with the association step (C) or dissociation step (A, B) in the presence of the indicated concentrations of NAD^+ and/or the inhibitor GC7. Depicted are association and dissociation curves. Ligands were loaded at a concentration of 10 mg L^{-1} , analytes were supplied at 100 nM (A), 50 nM (B) or 25 nM (B). Measurements were performed with the Octet Red96e system and data was prepared with the Data Analysis HT software (version 10.0.0.48; PALL *forté*BIO).

(25 nM), while a concentration of 1 nM did not exhibit any distinct effect on the binding behavior (see figure 5.23 C). This is in accordance with the calculated saturation of the DHS with GC7 molecules at the indicated concentrations by the law of mass action: Based on the K_i of GC7 of 9.7 nM (determined for DHS from rat testis [114]), at 1 mM all DHS molecules should be saturated to 100% with GC7, while at 1 nM only 0.8% of the DHS tetramers are in a complex with the inhibitor. Further experiments will have to screen different inhibitor concentrations in-between these values.

In addition, even the residual carryover on the sensor tips from the regeneration step with 5 mM GC7 and NAD^+ completely abolished DHS binding to eIF-5A ligands in reused 100 nM analyte solutions in the course of four repetitions, while fresh analyte solution restored binding to expected levels afterwards (data not shown). To prevent this effect in further analyses with regenerated biosensors, they were washed successively in 12 different wells with fresh PBS buffer after each regeneration (see section 7.14.8).

Overall, the presence of the cofactor NAD^+ at millimolar concentrations on the one hand significantly stabilized the interaction between *SbDHS* and its substrate protein eIF-5A by lowering the dissociation rate. The inhibitor GC7, on the other hand, drastically increased dissociation rates at millimolar concentrations and prevented DHS binding to eIF-5A even at much lower concentrations. The combination of NAD^+ and GC7 enhanced the dissociation process even further, while their combined effect on the association phase was not tested yet.

5.6. SEC-MALS characterization of *SbDHS1*, *SbDHS2* and *SbeIF-5A*

The results from the binding experiments on the OCTET system described in the previous section indicated that the presence of NAD^+ significantly stabilizes the interaction between *SbDHS2* and *SbeIF-5A*. For the crystallization and subsequent X-ray diffraction-based structure determination of a *S. bicolor* DHS–eIF-5A complex, the stable and long-lasting binding between the enzyme and its substrate would be a crucial prerequisite.

To determine if the presence of NAD^+ in the buffers sufficiently stabilizes the interaction to allow the isolation of the enzyme–substrate complex by size-exclusion chromatography (SEC), several SEC runs – followed by an online multi-angle light scattering (MALS) analysis – were performed on samples from *S. bicolor* DHS1, DHS2, eIF-5A1 and combinations thereof in the presence or absence of additional NAD^+ . Therefore, samples of IMAC- and SEC-purified SA*StreptII-SbDHS1*-His₆, SA*StreptII-SbDHS2*-His₆ and *SbeIF-5A1*-His₆ at concentrations of 2 g L^{-1} ($11.4 \mu\text{mol L}^{-1}$; DHS) or 1 g L^{-1} ($54.1 \mu\text{mol L}^{-1}$; eIF-5A) as well as a 1:5 mixture of either of the two DHS tetramers with *SbeIF-5A1*-His₆ were separated on a Superdex 200 10/30 GL SEC column (GE Healthcare) in a PBS buffer at pH 7.4. The separated molecule species were subsequently characterized by online MALS. The measurements were performed once in the absence of externally added NAD^+ and then again with 1 mM NAD^+ supplied in sample and running buffer (for further experimental details, see section 7.14.7).

In figure 5.24, the chromatograms of the different SEC runs with and without additional NAD^+ performed for the *SbDHS2* and eIF-5A1 samples are depicted. The chromatograms show the relative UV absorbance measured over the duration of the SEC run and the results

5.6. SEC-MALS characterization of *SbDHS1*, *SbDHS2* and *SbeIF-5A*

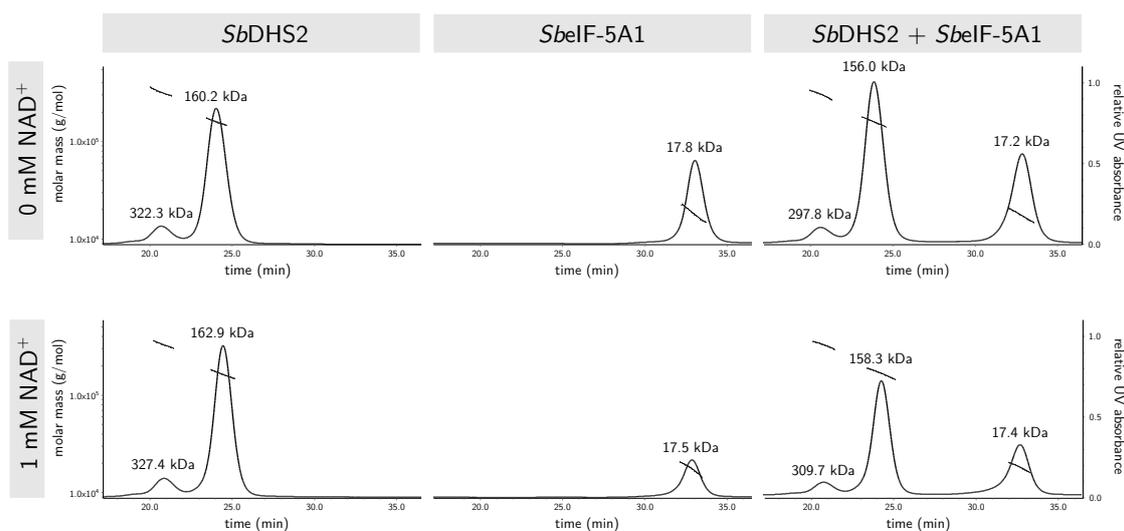


Figure 5.24.: SEC-MALS measurements of *S. bicolor* DHS2 and eIF-5A

Samples of *SAbstreptII-SbDHS2-His₆*, *SbeIF-5A1-His₆* or a 1:5 mixture of DHS2 tetramer to eIF-5A1 monomer were separated on a Superdex 200 10/30 GL SEC column (GE Healthcare) in a PBS buffer at pH 7.4, either with or without additional 1 mM NAD⁺. An online MALS analysis was performed following SEC. The molecular weights M_w of the different molecule species calculated from MALS data are denoted above the corresponding peaks.

of the molar mass determination by MALS for the analyzed peaks. The results for *SbDHS1* were comparable and are therefore not shown.

The separately run *SbDHS2* samples (left chromatograms) in the absence and presence of additional NAD⁺ exhibit a very similar running behavior. The main peaks – with mass fractions of 87.6 and 89.2 %, respectively – consist of a monodisperse molecule species (dispersity \mathcal{D} : $M_w/M_n=1.002$ or 1.003) with a calculated molecular weight of 160.2 or 162.9 kDa. This is in the range of the expected molecular weight of the active tetramer (175.2 kDa) calculated from the amino acid sequence. In addition to the main peak, a second peak with a mass fraction of 12.4 or 10.8 %, respectively, is also visible at lower retention times in both *SbDHS2* samples. This second peak consists of a monodisperse molecule species ($\mathcal{D}=1.003$ or 1.002) with a calculated molecular weight of 322.3 or 327.4 kDa, respectively. Interestingly, this molecular weight equals the weight of a dimer of the active DHS2 tetramer ("octamer"). A similar running behavior has already been observed previously for the polishing SEC step during purification (data not shown). However, after the polishing SEC, only fractions from the main peak with the approximate molecular weight of the active tetramer had been pooled. Obviously, small amounts of higher order oligomers reformed upon storage, thawing or sample preparation.

The *SbeIF-5A1* samples run in the absence or presence of NAD⁺ (figure 5.24, middle chromatograms) only show a single, monodisperse peak ($\mathcal{D}=1.013$ or 1.009) with a calculated molecular weight of 17.8 or 17.5 kDa, respectively. This correlates well with the expected molecular weight of 18.5 kDa for the eIF-5A1 monomer.

The chromatograms from samples of *SbDHS2* 1:5 mixed with *SbeIF-5A1* run in the absence or presence of NAD^+ (right chromatograms) exhibit three separate, monodisperse peaks ($D=1.004\text{--}1.008$), that correlate with the two peaks from separately measured *SbDHS2* – with 156.0 or 158.3 kDa for the tetramer and 297.8 or 309.7 kDa for the putative "octamer" – and the single peak from *SbeIF-5A1* with 17.2 or 17.4 kDa, respectively. For the complex of the *SbDHS2* tetramer with one eIF-5A1 monomer (see section 4.2.4 for further information on DHS/eIF-5A binding stoichiometry), a molecular weight of approx. 193.7 kDa can be expected. However, the data do not indicate the presence of a molecule species of that molecular weight – neither in the absence nor in the presence of NAD^+ .

Therefore it can be assumed that although the presence of NAD^+ has a distinct stabilizing effect on the interaction between *SbDHS2* and *SbeIF-5A1* during binding measurements with the OCTET system, the interaction is not stable enough to stay intact over the course of a SEC experiment.

The productive interaction of *N*-terminally *StreptII*-tagged *SbDHS2*-His₆ with *ZmeIF-5A*-His₆ has been proved by activity measurements (A. Wesseling³, internal communication). Although it is therefore unlikely that the *N*-terminal SA*StreptII*-tag would interfere with the binding of *SbeIF-5A1*-His₆, it has to be noted that SA*StreptII*-tagged *SbDHS1*-His₆ and *SbDHS2*-His₆ – as ligands – did not yield interpretable association/dissociation curves with *SbeIF-5A1*-His₆ or *SbeIF-5A2*-His₆ in the binding experiments with the OCTET system described in the previous section.

5.7. Interaction of DHS and HSS subunits

Eukaryotic as well as archaeal DHS is an obligate homotetramer. Since the active sites are made up from the combined interfaces of two subunits (see section 4.2.2), the integrity of their interaction is crucial for DHS functionality. The evolution of an HSS paralog in angiosperm species – which lacks DHS functionality and exhibits enhanced HSS activity – is therefore restricted by the necessity to maintain the active sites in a functional state. Directly after the duplication of the DHS gene, the duplicates are able to interact with each other, forming paralogous heteromers. Through this interaction, the new duplicate inherits the selective pressure of the original DHS. Since changes in the new duplicate, which negatively affect eIF-5A binding or deoxyhypusine synthesis activity, would probably also be detrimental for the functionality of the original DHS – a phenomenon termed "paralog interference" [12, 121] – the evolution of HSS from the duplicate DHS is most likely driven by a purifying selection process, which prevents changes that impair the function of heterologous complexes. In the case of angiosperm DHS and HSS, the HSS needs to escape the selection pressure of the DHS to optimize its HSS activity. This can either be done by a separation of the paralogs at the expression level, which minimizes the possibilities for negative effects, or by prevention of paralog interaction by changing the responsible interface.

As explained above, the interface is crucial for DHS or HSS functionality. It is therefore most likely that DHS and HSS will retain their mutual interfaces and are still able to interact

³Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

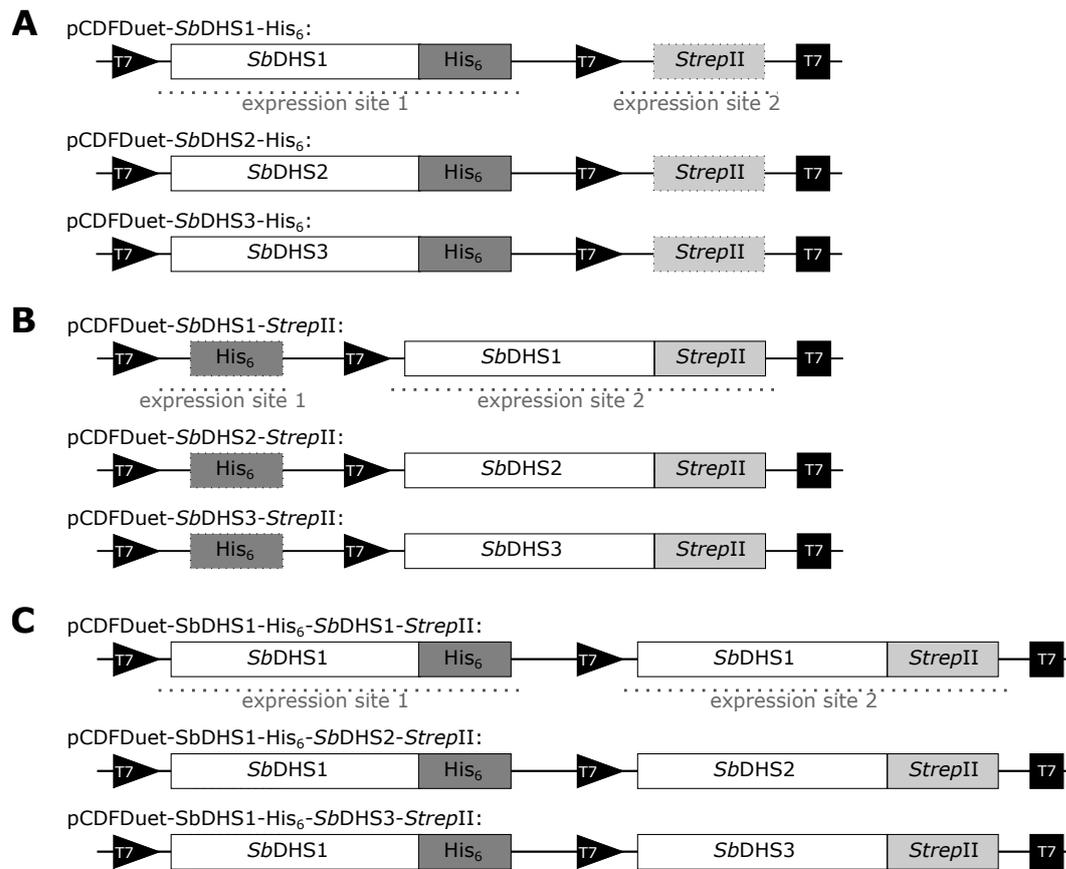


Figure 5.25.: Coexpression and control constructs of *SbdHS/HSS*

These coexpression and control constructs are based on the pCDFDuet™ 1 vector, which features two multiple cloning sites (MCS) with their own T7 promoter and a common T7 terminator sequence. The original S-tag in MCS2 was replaced with a *StrepII*-tag. **A:** Control constructs with C-terminally His₆-tagged *Sorghum bicolor* DHS1, DHS2 and DHS3 (HSS) in MCS1 and an empty MCS2. **B:** Control constructs with C-terminally *StrepII*-tagged DHS1, DHS2 and DHS3 (HSS) in MCS2 and an empty MCS1. **C:** Coexpression constructs with C-terminally His₆-tagged DHS1 in MCS1 in combination with either C-terminally *StrepII*-tagged DHS1, DHS2 or DHS3 (HSS) in MCS2.

with each other. However, the interaction of paralogous angiosperm DHS and HSS has not been described to date. To verify if DHS and HSS can indeed form paralogous heteromers as a proof of this concept, coexpression constructs of the two DHS homologs (DHS1 and DHS2) and their HSS paralog (DHS3) from the model plant *Sorghum bicolor* have been designed, cloned and tested in the course of this thesis. In figure 5.25 these constructs are depicted schematically (see section 7.9.5 for the detailed cloning strategy).

Three coexpression constructs were designed, which each contain *SbdDHS1* with a C-terminal His₆-tag in MCS1 and either DHS1, DHS2 or DHS3 (HSS) with a C-terminal *StrepII*-tag in MCS2 (see figure 5.25 C). The coexpression construct with DHS1 in both MCS was intended as positive control: Since DHS1 naturally interacts with itself, heteromers

consisting of His₆- and *Strepll*-tagged DHS1 must occur – unless the different tags impair their interaction. In addition to the coexpression constructs, six control constructs were created, which contained DHS1, DHS2 or DHS3 (HSS) in either of the two MCS for the separate expression of the His₆- and *Strepll*-tagged homo/paralogs (see figure 5.25 **A** and **B**).

The coexpression and control constructs were expressed in *E. coli* BL21-Gold(DE3) and subsequently affinity purified via the His₆-tag (or *Strepll*-tag for separately expressed, *Strepll*-tagged DHS/HSS) and polished by SEC, according to the established DHS expression and purification protocols (see sections 7.12.1 and 7.12.2). After this initial purification, the coexpressed proteins and the positive and negative controls were purified a second time on a small scale. Therefore, the negative controls were prepared in advance by mixing *SbDHS1*-His₆ with either *SbDHS1-Strepll*, *SbDHS2-Strepll* or *SbDHS3-Strepll* and incubated at 4 °C overnight. These negative controls as well as the coexpressed proteins were subsequently bound to Ni-NTA resin and rigorously washed with 40 CV of a buffer containing 1 M NaCl and 2% (v/v) Tween[®] 20 to interrupt all unspecific protein–protein interactions (see section 7.12.3). Afterwards, the bound protein was eluted and analyzed by denaturing SDS-PAGE and subsequent western blot (see sections 7.7.1 and 7.7.3).

For the SDS-PAGE, similar protein amounts were applied for each sample. Besides the three coexpression samples, the separately expressed *SbDHS1*-His₆ (*Strepll*-tag negative control), the purification negative controls and the separately expressed *SbDHS1-Strepll*, *SbDHS2-Strepll* and *SbDHS3-Strepll* (*Strepll*-tag positive controls) were analyzed. On the western blot, the *Strepll*-tag of the DHS or HSS that formed a heterotetramer with the *SbDHS1*-His₆ (or of the *Strepll*-tag positive controls) was detected with a color reaction of AP-conjugated *Strep*-Tactin[®] (see section 7.7.4).

The western blot results of two experiments (samples from the same initial purifications, but separately handled in the second, small-scale purification step) are depicted in figure 5.26. In both blots (**A** and **B**), the Ponceau S stain confirms similar amounts of protein in the different lanes. As expected, all bands are located at about the height of the 40 or 42 kDa band of the respective molecular weight marker. This corresponds to the calculated molecular weight of the separate subunits: 41.8 kDa for *SbDHS1*-His₆ and 42.0, 42.1 and 42.5 kDa for *Strepll*-tagged *SbDHS1*, 2 and 3, respectively.

In lane 1 – the *Strepll*-tag negative control of both blots – a distinct Ponceau S stained red band is visible, but no purple stain of the AP color reaction (middle image). In contrast, the *Strepll*-tag positive controls in lanes 2–4 (**A**) and 8–10 (**B**) exhibit distinct purple bands. This confirms that the detection with AP-conjugated *Strep*-Tactin[®] on the western blot is specific for the *Strepll*-tag, excluding false positive results.

Lanes 5–7 in both blots contain the coexpressed protein samples of *SbDHS1*-His₆ with either *SbDHS1-Strepll*, *SbDHS2-Strepll* or *SbDHS3-Strepll*. As expected, the samples from *SbDHS1*-His₆ coexpressed with *SbDHS1-Strepll* (lanes 7) show a distinct purple band. This demonstrates that homologous subunits, which only differ in their C-terminal affinity tag, readily form "hetero"tetramers, and confirms that the C-terminal affinity tag does not influence the assembly of the tetramers. In addition to the DHS1/DHS1 complex, the clearly visible purple bands in lanes 5 and 6 indicate the presence of DHS1/DHS3 (HSS) and DHS1/DHS2 heteromers. Both bands are of similar intensity as the band of DHS1/DHS1 in **A** and slightly

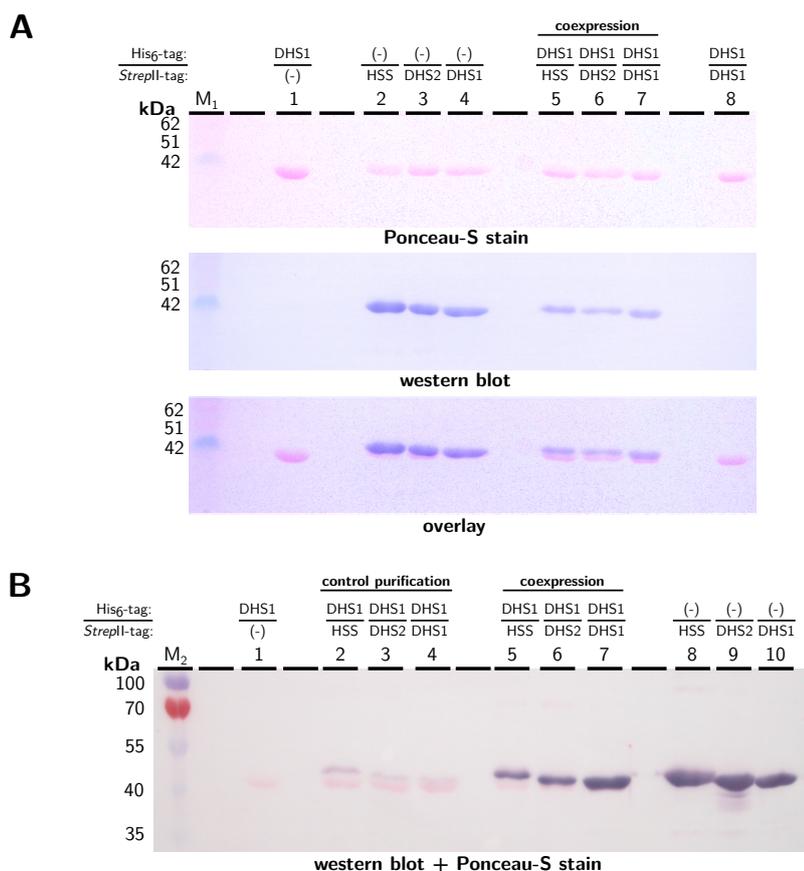


Figure 5.26.: Western blots of coexpressed *SbDHS/HSS*

A: Ponceau-S stain, western blot detection of *StreptII*-tagged *SbDHS/HSS* with AP-conjugated *Strep-Tactin*[®] and overlay. Lane 1: negative control, sample from a separate control expression of His₆-tagged DHS1 purified via His₆-tag; lanes 2–4: positive control, samples from separately expressed *StreptII*-tagged DHS/HSS purified via *Strep-Tactin*[®]; lanes 5–7: samples from the coexpression of His₆- and *StreptII*-tagged DHS/HSS purified via His₆-tag; lane 8: purification control, sample from separately expressed, subsequently mixed and His₆-tag-purified His₆- or *StreptII*-tagged DHS1. **B:** Western blot detection of *StreptII*-tagged DHS/HSS with AP-conjugated *Strep-Tactin*[®], stained with Ponceau-S. Lane 1: negative control, sample from a separate control expression of His₆-tagged DHS1 purified via His₆-tag; lanes 2–4: purification control, samples from separately expressed, subsequently mixed and His₆-tag-purified His₆-tagged DHS1 with *StreptII*-tagged DHS1, DHS2 or HSS (DHS3); lanes 8–10: positive control, samples from separately expressed *StreptII*-tagged DHS/HSS purified via *Strep-Tactin*[®]. M: Molecular weight marker (M₁: Roti[®]-Mark BI-PINK, Roth; M₂: PageRuler Prestained Protein Ladder, Thermo Scientific).

less pronounced in **B**. These differences might be due to differences in overall protein load, but could also hint at a less favorable complex formation between heterologous subunits. However, these experiments do not allow an unequivocal distinction.

Interestingly, the purification controls in lane 8 (**A**) and lanes 2–4 (**B**) also show a faint purple band in addition to the Ponceau S stain. This is unexpected, since the washing procedure of the second small scale purification step was rather rigorous. However, these bands are

very weak compared to the bands from the coexpressed samples. It can therefore be assumed that the *Strepll*-tag signal detected from the coexpression samples actually corresponds to the presence of heterologous tetramers. Since even huge complexes are not static and the samples of the purification control have been incubated overnight, it might be possible that subunits of the initial homotetramers have been shuffled over time, resulting in a small fraction of heteromers after incubation.

Taken together, these results confirm the presence of paralogous heteromers for an angiosperm DHS/HSS pair – namely *SbDHS1* and *SbDHS3* – for the first time and present evidence for the influence of paralog interference on HSS evolution.

6. Discussion

Eukaryotic and archaeal deoxyhypusine synthase (DHS) is a highly conserved enzyme, which catalyzes the unique NAD⁺-dependent deoxyhypusination reaction on a single specific lysine residue of the eIF-5A precursor protein. This modification is the first of two steps in the formation of hypusine at this residue. The (deoxy)hypusination of eIF-5A is crucial for its function in essential cellular processes – thereby rendering the (deoxy)hypusination machinery indispensable for all eukaryotic and archaeal cells. Best characterized to date is the role of the modified eIF-5A in the translation of ribosome-stalling motifs such as polyproline stretches, and in the translation termination process, although it has also been implicated in pathologic processes like viral reproduction, some forms of cancer or diabetes (see section 4.1).

The DHS enzyme is an obligate homotetramer, that contains four identical active sites (see section 4.2.2). Two antiparallel active sites are located at the end of a deep tunnel at the interface of two subunits, the "active" dimer. These dimers congregate to form the actual tetramer. In the deoxyhypusination reaction, an aminobutyl moiety is transferred from the polyamine substrate spermidine via an enzyme–imine intermediate at the catalytic lysine onto the acceptor lysine on a protruding finger-like loop of eIF-5A (see section 4.2.1). The deoxyhypusination reaction requires the presence of NAD⁺ in the active site, which is reduced in the first step of the reaction, stays bound to the enzyme during the following steps and finally regenerates by reducing the reaction product to deoxyhypusine – a mechanism termed "transient hydride transfer" [82].

Besides the deoxyhypusination reaction, under natural conditions the DHS enzyme also catalyzes the formation of homospermidine from spermidine and the small diamine putrescine with low efficiency (see 4.2.5). In several angiosperm lineages, gene duplication gave rise to a second DHS coding sequence, which has been subsequently optimized for this homospermidine synthase (HSS) side reactivity – resulting in a paralogous HSS with enhanced homospermidine synthase activity and loss of DHS activity (see section 4.3). In many of these plant lineages, the new HSS enzyme became the first specific enzyme of the pyrrolizidine alkaloid biosynthesis pathway, which utilizes homospermidine as starting product for the synthesis of defense compounds against herbivores.

In the model plant *Sorghum bicolor*, gene duplication events created three distinct DHS-like sequences. Activity assays¹ performed on cloned and affinity purified protein samples from these DHS-like proteins revealed that while two of these enzymes (DHS1 and DHS2) still exhibit DHS activity alongside HSS activity, the third enzyme (DHS3) showed significantly impaired DHS functionality, but enhanced homospermidine synthesis rates. It was therefore proposed that these three DHS variants represent different evolutionary stages in the development of a separate HSS from a DHS gene duplication – with a well-advanced HSS (DHS3), the original DHS that has to be maintained due to the essential nature of the catalyzed

¹Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

deoxyhypusination reaction, and a more recent duplicate, which might exhibit features of an intermediate between DHS and HSS.

To gain a deeper understanding of common features and differences of these three *S. bicolor* DHS homo/paralogs, they were characterized in more detail by several different biophysical approaches in the course of this thesis.

6.1. Supplementation of NAD⁺ is necessary to prevent protein loss during purification and aggregation upon storage

In the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober at the Botanical Institute of Kiel University, plant DHS and HSS are routinely expressed and purified in medium scale according to established protocols to provide protein samples for activity assays and the determination of enzyme characteristics. However, for the biophysical experiments and crystallization trials performed in the course of this thesis, larger amounts of homogeneous, well-behaved *S. bicolor* DHS and HSS had to be produced.

Therefore, *E. coli* expression cultures were upscaled based on the original protocol, and the produced protein was IMAC and SEC purified according to the established procedure (see section 5.1.1). While the increase of expression culture volume on the one hand resulted in higher amounts of heterologously expressed DHS and HSS protein, the subsequent IMAC purification on the other hand revealed that the obtained protein exhibited inhomogeneous behavior. After the binding of the His₆-tagged DHS or HSS to the Ni²⁺-matrix of the affinity column, the bound protein is regularly washed with a buffer containing 1.5 M NaCl to remove unspecifically bound protein impurities from the immobilized target protein. In protein purified from small- to medium-scale expressions, this wash step has no detrimental effect on the bound protein. In contrast, vast amounts of the His₆-tagged DHS from large-scale expressions were washed off the affinity matrix directly following the high salt step – leaving only a small fraction of the tagged protein bound to the column. After all, the stably bound protein behaved as expected, but the total protein yield was not much higher than that from medium-scale expressions. Since an overloading of the affinity matrix material could be excluded, the cause for huge protein loss at the established high salt wash step had to be attributed to an inhomogeneous situation in the heterologously expressed protein pool.

As already mentioned above, the homogeneity of the purified protein is a crucial prerequisite for the performance of biophysical measurements and especially for protein crystallization for structure determination purposes. Therefore, several strategies were tried to improve protein quality. In a first attempt, the expression conditions were varied in the direction of slower protein production over an extended period to allow for proper folding and to reduce the risk of the production of truncation fragments. In addition, a different expression strain – RosettaTM 2(DE3) – was applied, which provides additional tRNAs for rare codons. Since a codon usage analysis revealed several critical codons in the DHS coding sequences, this was also supposed to reduce the occurrence of truncation fragments from abortive translation, which would miss the C-terminal His₆-tag. Tetramers containing one or more of such truncated proteins might therefore exhibit different binding behavior than tetramers containing four His₆-tags. In addition, based on SDS-PAGE results the *E. coli* chaperone GroEL was

suspected to bind to a fraction of the expressed protein, and buffers containing ATP and Mg²⁺ were applied to stimulate the release of chaperone-bound protein. However, neither of these efforts had a significant impact on the binding behavior of the His₆-tagged DHS in the IMAC purification procedure.

Furthermore, constructs containing an additional *N*-terminal *StreptII*-tag were designed, cloned and expressed in parallel to the original constructs. While the *N*-terminal *StreptII*-tag significantly increased overall protein expression rates compared to the original constructs with the native DHS/HSS *N*-termini, they exhibited similar behavior upon IMAC purification as their original counterparts – although the purification via the *StreptII*-tag did not exhibit any irregularities.

Subsequently, DHS protein samples were analyzed by differential scanning fluorimetry (DSF [252]) to ensure optimal buffer conditions during cell lysis and the following IMAC purification. DSF gives an estimate of the thermal stability of a protein under the analyzed conditions and can be a valuable tool in the identification of stabilizing additives or the optimal pH and buffer substance for the handling of a certain protein. In a 96 well format, it allows the parallel screening of a broad range of different conditions with a relatively low amount of sample protein (see sections 5.1.2 and 5.4 for further details). Here, a variety of pH values and buffer substances as well as different salt concentrations were tested. However, while these experiments confirmed that the *S. bicolor* DHS variants are similarly stable over a broad pH range from about pH 5.5 to pH 9.0, and do not exhibit a distinct buffer substance preference or a pronounced dependence on salt concentrations, they provided no indication for an optimization of the lysis and purification conditions. Nevertheless, based on these results some aspects of buffer composition were changed just for protocol simplification, but without effect on protein homogeneity in the IMAC purification step (see sections 5.1.2 for details).

Finally, the routinely performed dynamic light scattering (DLS) measurements – which are applied to assess the quality of purified protein prior to flash freezing and storage and after thawing – delivered the critical clue for the solution of the "inhomogeneity enigma". DLS is a biophysical technique that allows the determination of the apparent hydrodynamic diameter of randomly moving particles in solution or suspension based on the intensity fluctuations of scattered laser light (see section 7.14.6 for further details). For freshly purified DHS samples, DLS measurements revealed a nearly monodisperse particle size distribution around an average diameter of about 11 nm, which corresponds to the expected size of a DHS tetramer. After flash freezing, storage and thawing, the same sample exhibited again a rather narrow distribution of particle sizes – albeit around an average diameter of 18.6 nm, which corresponds to an estimated protein molecular weight of approx. 624 kDa. This indicates that the DHS tetramers might assemble – or aggregate – into well-defined higher-order oligomers. Based on the obtained average particle size this oligomer is most likely a hexadecamer as a tetramer of tetramers. However, this assumption remains speculative.

The DHS enzyme depends on NAD⁺ as a cofactor and spermidine as aminobutyl donor for the deoxyhypusination reaction. While the binding of NAD⁺ and spermidine to DHS is independent of the binding of the eIF-5A precursor, the binding of spermidine most likely requires the prior binding of NAD⁺ (see section 4.2.3), which is assumed to induce a conformational adaption of the active tetramer. Therefore, NAD⁺ and the spermidine analog and DHS

inhibitor *N*¹-guanyl-1,7-diaminoheptane (GC7, see figure 4.5), which had already been successfully applied as crystallization additive (see section 4.2.2), were tested for their stabilizing and anti-aggregative effect on the DHS enzyme in further DLS experiments. Interestingly, immediately upon the addition of 2 mM NAD⁺ and GC7 to the thawed and oligomerized DHS sample, the narrow particle size distribution shifted back to the average diameter of freshly purified DHS prior to flash freezing and storage. The separate addition of NAD⁺ subsequently revealed that this effect was solely due to the addition of NAD⁺, while GC7 alone did not induce any changes in the observed particle size distribution. According to the law of mass action, the saturation of DHS active sites with the inhibitor should be 100% at the applied enzyme and inhibitor concentrations, if the *K_i* of 9.7 nM determined for DHS from rat testis is assumed to apply also to *S. bicolor* DHS [114]. However, considering that GC7 most likely requires the presence of NAD⁺ in the active site for its own binding, the lack of any detectable effect is to be expected if the active sites are only partially occupied by NAD⁺ – as concluded from results obtained from the supplementation with NAD⁺ alone.

These DLS experiments indicated that the presence of NAD⁺ plays a significant role in the preservation of a defined and homogeneous oligomeric state of *S. bicolor* DHS and can reverse aggregation upon flash freezing and storage. It was therefore hypothesized that the inhomogeneous behavior of a large portion of the expressed DHS enzyme during IMAC purification was due to an inhomogeneous saturation of the individual DHS tetramers with NAD⁺ – which led to a similar aggregation or oligomerization process during lysis and purification as was observed for the purified (and originally well-behaved) DHS upon flash freezing and storage. Initially, oligomerized/aggregated enzyme might therefore bind weakly to the IMAC matrix, but is afterwards washed off by the high salt wash step – probably due to impaired accessibility of the His₆-tag in the aggregates. The aggregates/oligomers most likely stay bound to the column and the other bound proteins during the high salt wash step due to hydrophobic interactions, which are promoted by the high ionic strength of the wash buffer. If the salt concentration is lowered to the initial value, these hydrophobic interactions are reduced and the aggregates elute from the column.

Subsequent purification experiments confirmed that supplementation of lysis and purification buffers with low millimolar amounts of NAD⁺ completely prevented inhomogeneous binding behavior and enabled the purification of large amounts of well-behaved, homogeneous *S. bicolor* DHS and HSS enzymes from all constructs used in the course of this thesis. In addition to the significantly improved purification procedure, the addition of 2 mM NAD⁺ to all storage buffers reliably prevented oligomerization/aggregation upon flash freezing and storage as further DLS measurements revealed.

Based on these results it has to be assumed that the *E. coli* cells were not able to provide all DHS enzymes with the necessary amount of cofactor to prevent oligomerization or aggregation upon strong overexpression for an elongated period. Furthermore, the effect of NAD⁺ on the aggregation state of the DHS enzymes shows that the cofactor must have a distinct effect on the conformation of the tetramer – enabling or preventing the congregation of tetramers into higher order assemblies. This is in accordance with results from fluorescence measurements of intrinsic tryptophan fluorescence, which was quenched upon binding of NAD⁺ and thereby indicated a conformational adaptation during DHS–NAD⁺ complex formation [148, 294, 149].

6.2. Two *Sorghum bicolor* eIF-5A isoforms are readily expressed and purified based on previously established protocols

The DHS enzyme and its original substrate, the eIF-5A precursor, are highly conserved between eukaryotic and even archaeal species. Especially the *N*-terminal domain of eIF-5A, which carries the modified lysine residue in a finger-like loop and probably mediates binding to the surface funnel of the DHS enzyme for the correct positioning of the lysine residue in the active site, is very well conserved (see section 4.1). It is therefore no surprise that the DHS enzyme not only accepts its "own" eIF-5A, but also readily modifies eIF-5A precursors from other species. Examples are rat testis DHS catalyzing deoxyhypusination of eIF-5A precursor from Chinese hamster ovary cells (CHO cells), yeast DHS modifying human eIF-5A precursor or *Senecio vernalis* DHS accepting *Phalaenopsis* eIF-5A [293, 124, 187].

However, although a C-terminally His₆-tagged eIF-5A construct from the Poacea *Zea mays* was already available in the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober (Botanical Institute of Kiel University), the authentic *Sorghum bicolor* eIF-5A was chosen over *Zea mays* eIF-5A to represent the native situation in the subsequent biophysical characterizations. Therefore, data bank research was performed in 2015 in the course of this thesis to identify eIF-5A-like coding sequences from *Sorghum bicolor* genomic sequence data. At that point, two eIF-5A-like coding sequences from different gene loci were available for *S. bicolor*, which were termed *Sb*eIF-5A1 (GenBank Acc. No.: KXG34270.1) and *Sb*eIF-5A2 (GenBank Acc. No.: KXG37504.1) in this thesis. In 2017, an additional eIF-5A-like coding sequence was identified (GenBank Acc. No.: KXG36613.1; see section A.5 figure A.13 for an alignment of all three sequences).

With the sequence information from the above-specified data bank entries, primers were designed for the targeted amplification of the first two eIF-5A variants from cDNA, which was obtained by RNA isolation from fresh *S. bicolor* leafs (Botanical Garden, Kiel University) and subsequent reverse transcription. Afterwards, the coding sequences were cloned into the *Z. mays* eIF-5A vector – substituting the original *Z. mays* sequence. In addition to these C-terminally His₆-tagged constructs, two constructs containing a C-terminal serine-alanine-*StrepII*-tag were created. These were intended for the binding kinetic experiments, since the SA*StrepII*-tag enabled stable and specific binding to the *StrepMAB*-Imm antibodies (iba), which were used for ligand immobilization on the biosensors (see section 5.5).

After confirmation of construct sequence integrity by sequencing, all constructs were heterologously expressed in *E. coli* BL21(DE3) or BL21-Gold(DE3) and subsequently purified by IMAC based on previously established protocols² (see sections 7.11.1 and 7.11.2 for details).

In contrast to the upscaled *S. bicolor* DHS expression and purification described in the previous section, expression and IMAC or *StrepII*-tag-based purification of both *S. bicolor* eIF-5A isoforms were straightforward and yielded high amounts of very pure protein (see section 5.2.2 figure 5.7). To ensure not only purity, but also homogeneity, the IMAC-purified protein samples were subjected to a polishing SEC step, where they exhibited a single peak around the expected retention volume – indicating a homogeneous sample with a defined molecular mass

²Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

and hydrodynamic diameter. This was also confirmed by routinely performed DLS measurements with freshly purified and thawed protein, which exhibited a single molecule species with nearly monodisperse average particle size distribution (data not shown). However, the average molecule diameter derived from these measurements was around 5.9 nm – corresponding to a calculated molecular weight of approx. 42 kDa.

In literature, homodimerization of eIF-5A monomers was reported for several species. Early studies on human eIF-5A isolated from erythrocytes already demonstrated homodimerization [55]. For yeast eIF-5A, an L-shaped homodimer with an overall structure similar to that of bacterial EF-P was identified by small-angle X-ray scattering (SAXS) experiments [66]. Other experiments with yeast or human eIF-5A indicated that the homodimerization is dependent on the hypusination and requires the presence of RNA, which is bound by the eIF-5A homodimer [154]. Without hypusination, the examined yeast eIF-5A was predominantly monomeric. In addition, the dimerization could be reversed by the addition of the reducing agent DTT, although the involvement of disulfide bridges was ruled out [86]. The only crystal structure of plant eIF-5A to date is from *Arabidopsis thaliana* and also features a distinct dimerization pattern – albeit with a parallel orientation that is completely different to the L-shaped homodimer reported for yeast [274]. Sequence comparisons revealed several residues involved in hydrogen bonding at the *N*- and *C*-terminal interfaces (Y19, P20, Q21, S22, S107, T110), which are highly conserved in plant eIF-5A sequences, and are also present in the two *S. bicolor* eIF-5A isoforms examined in this thesis.

With these results from literature in mind, it is very well imaginable that the *S. bicolor* eIF-5A isoforms are not monomeric, but are mostly dimerized – as the results of the DLS data analyses with an approximate molecular weight of 42 kDa suggest. However, molecular weight calculations based on the apparent hydrodynamic diameter obtained from DLS measurements always assume a globular molecule shape. The determined hydrodynamic diameter only represents the diameter of a sphere that would move in the same way as the scattering particle. The eIF-5A monomer, in contrast, is not of globular shape, but has an elongated, dumb bell-like structure. The calculated molecular weight from DLS data analysis might therefore overestimate the actual size – and thereby the molecular weight – of the scattering eIF-5A molecules. Indeed, later SEC-MALS measurements showed that IMAC- and SEC-purified *S. bicolor* eIF-5A is most likely monomeric, rather than a homodimer (see section 5.6).

Taken together, all four new *S. bicolor* eIF-5A constructs behaved similar to eIF-5A from other plant species (e.g. *Z. mays*) and could be obtained in high amounts from heterologous expression in *E. coli* and subsequent IMAC purification in combination with a polishing SEC or *StreptII*-tag-based affinity chromatography step. The final protein charges exhibited excellent purity and homogeneity suitable for the following biophysical characterizations or crystallization trials.

6.3. Crystallization of individual *S. bicolor* DHS homo/paralogs and in complex with eIF-5A precursor

As described in detail in section 4.3, the three *S. bicolor* DHS – or more precisely the HSS-like DHS3 and the two DHS isoforms 2 and 3 – exhibit a very high amino acid sequence identity.

Activity assays already showed that DHS3 only displayed severely reduced DHS functionality, but significantly improved HSS activity (A. Wesseling³, internal communication). In the course of this thesis it could be demonstrated that the reduced DHS activity is caused by the loss of eIF-5A binding to DHS3 (see sections 5.5.1 and 6.5.1). Based on the few amino acid changes between the two DHS and the HSS-like DHS3 derived from the sequence alignment and an analysis of the position of these changes in a model structure based on the structure of human DHS (see section 4.3 for further details), only a handful of residues have to be responsible for the improvement of the HSS functionality and the loss of eIF-5A binding. While the identity of these residues can on the one hand be further narrowed down by future mutational analyses (as discussed in section 6.5.1), crystal structures of the three DHS – and at best complex structures of the two DHS with the eIF-5A precursor – would on the other hand allow much deeper insight into the underlying structural differences of this evolutionary "change of function". While the crystal structure of human DHS is already available in the databases and also several structures of eIF-5A from different species have been solved to date, no crystal structure of a complex between a DHS and the eIF-5A precursor has been determined yet.

Therefore, two different crystallization approaches were applied in this thesis to generate protein crystals for structure determination purposes: the "classical" crystallization approach, which is aimed at the production of single crystals from careful supersaturation of a solution of purified protein, and the *in vivo* crystallization technique, where the protein crystallization is initiated inside a living cell by the overexpression of the target protein. Recent developments in data acquisition and analysis, like XFEL and serial crystallography, enabled the application of these *in vivo* grown crystals for the determination of protein structures and thereby opened up new routes for the generation of structural information [229, 85, 76, 34].

6.3.1. The classical crystallization approach yielded diffracting crystals with a high number of overlaid crystal lattices for the DHS enzymes

In the course of this thesis, a variety of different initial crystallization screens were used, mostly in a high throughput approach at the specialized HTX platforms at EMBL Hamburg and EMBL Grenoble, to identify promising crystallization conditions for the individual *S. bicolor* DHS enzymes and the complexes of DHS1 and DHS2 with either of the two eIF-5A isoforms (see section 5.3.1 for details).

For the individual enzymes, selected conditions – which already yielded small crystals in the initial screens – were reproduced in our laboratories and further optimized in small scale grid screens. However, the crystals did not grow as homogeneous single crystals, but rather intergrown as star-like structures, and plates with several layers of crystal lattices that are slightly displaced with respect to one another (see figure 5.8 for impressions). Although these crystals diffracted sufficiently well upon illumination with synchrotron-generated X-ray beam (beamline P14 of EMBL Hamburg at the PETRA III storage ring; DESY, Hamburg, Germany), the obtained diffraction data could not be interpreted due to the high number of overlaid

³Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

crystal lattices. This indicates that the crystal package and the contacts between the protein molecules are well in the two dimensions that make up the crystal layers, but that in the third dimension the interactions are weak and not sufficiently defined to promote three-dimensional crystal growth. Therefore, further efforts have to be made: on the one hand to optimize the existing crystallization conditions – e. g. by screening of additives that might improve crystal packing –, and on the other hand to explore other promising conditions, which have not been examined in more detail yet and might provide a better environment for the growth of three-dimensional single crystals. In addition, the protein constructs should also be reevaluated. At the moment, DHS constructs with *N*-terminal *StreptII*-tag and *C*-terminal His₆-tag were tried for crystallization, since they were obtained in much larger amounts from *E. coli* overexpression than only *C*-terminally His₆-tagged constructs. However, the enhancement of the already flexible putative *N*-terminal ball-and-chain structure (see section 4.2.2), might complicate the assembly of homogeneous crystals. By implication, the same could be true for the *C*-terminal His₆-tag, so that the application of constructs with only one tag, or the complete removal of tags prior to crystallization by the introduction of a specific protease cleavage site should also be considered.

For the crystallization of complexes between DHS1 and DHS2 and the two eIF-5A isoforms, promising conditions were also identified from a number of initial screenings in the presence of low millimolar concentrations of NAD⁺ and inhibitor GC7, which were supposed to stabilize the DHS enzyme and the interaction between enzyme and precursor based on previous DLS and DSF results (see sections 5.1.2 and 5.4). Since the eIF-5A constructs used for crystallization did not contain any tryptophan residues, UV images of the crystalline matter indeed indicated the presence of DHS enzyme in these structures by its fluorescence, but the presence of eIF-5A precursor could not be confirmed. Unfortunately, subsequent bio-layer interferometry binding experiments revealed that while inhibitor GC7 stabilizes the DHS enzyme, it actually induces a conformation that prevents the binding of the precursor and even harshly interrupts existing interactions (see section 6.5.4). Based on these results it is almost certain that the spherulites and crystalline matter that was observed for the complex crystallization trials solely consist of DHS enzyme molecules. Therefore, initial screens for complex crystallization have to be repeated in the absence of the inhibitor.

In addition, the production of variant constructs of the DHS enzyme should be considered. As summarized in section 4.2.4, for human DHS two mutations have been described that significantly enhanced the binding of the eIF-5A precursor: a D243A exchange in the vicinity of the tunnel entrance and the mutation of the catalytic lysine K329 to alanine. Both variants exhibited a stabilization of the DHS–eIF-5A complex and also an alteration of the stoichiometry from 1:1 to 1:4. Since both residues are also conserved in plant DHS – which at least is no surprise for the catalytic lysine residue –, they might facilitate the crystallization of the complex, and in comparison to the structure of the native complexes they would provide valuable insights into the structural mechanism that dictates 1:1 stoichiometry under natural conditions.

Besides the crystallization approach, alternative techniques for the determination of the complex structure should also be taken into account. Recent advancements in cryo-electron microscopy (CryoEM) and data processing now allow the determination of structures of pro-

teins and their complexes with near-atomic resolution from electron microscopy images of vitrified protein solutions (see [49] for a description of the method). Since structures are already available for human DHS (PDB Acc. No.: 1RQD, 1DHS) and e.g. the *Arabidopsis thaliana* eIF-5A (PDB Acc. No.: 3HKS), these structures could be modeled into three-dimensional data obtained from CryoEM and would at least reveal the stoichiometry, shape and interacting surfaces of the complex of DHS and the eIF-5A precursor.

6.3.2. Expression of DHS isoforms in Sf21 insect cells did not yield *in vivo* crystals

As already mentioned above, *in vivo* crystallization of proteins is not an isolated incident, but occurs rather often – especially in cell lines heterologously overexpressing a target protein. In the first part of this thesis, crystal structures were reported for the pestiviral transmembrane autoprotease NS2 upon overexpression in High FiveTM insect cells (see section 2.2.5). After these promising results, the *S. bicolor* DHS were also prepared for the expression in insects cells. On the one hand, the structure determination directly from *in vivo* grown crystals would ultimately simplify the overall route from protein construct to protein structure by skipping laborious protein purification and crystallization steps. On the other hand, the experience gained from a variety of different proteins crystallizing *in vivo* could finally lead to a more detailed understanding of the underlying cellular processes, which would also be of high value.

In this thesis, all three *S. bicolor* DHS isoforms were cloned into four vectors that contain signal sequences for the targeting of the overexpressed protein into different cellular compartments, namely the cytosol, ER and peroxisomes. These vectors were already applied for the *in vivo* crystallization of the autoprotease NS2 in the first part of this thesis. The construct identity was confirmed by sequencing and the integrity of the resulting artificial baculovirus genome (bacmid) with the desired expression cassettes was evaluated by colony PCR (see section 5.3.2 for further details). Subsequently, Sf21 insect cells were transfected with the different bacmids and the virus titer was increased by successive production of four virus generations in Sf21 insect cells (see section 7.13.3 for method details). The overexpression experiments were performed with the third and fourth virus generation in Sf21 cells. Unfortunately, overexpression in High FiveTM insect cells – which yielded crystals for the NS2 autoprotease in the first part of this thesis –, was ultimately prevented by time-consuming trouble-shooting due to a contamination of the High FiveTM cells.

However, since the EMBacY bacmid that was used here as carrier for the DHS expression cassette also contained an expression cassette for the fluorescing EYFP protein under the same promoter as the target DHS, the expression of the DHS enzymes could be monitored by the increasing EYFP fluorescence in the Sf21 insect cells (see figure 5.11 for impressions). From the 12 tested DHS constructs, two DHS3 constructs failed to exhibit any EYFP expression or other signs of infection due to inefficient transfection or infection in the course of the virus amplification. For all other constructs, distinct yellow fluorescence was observed already at day 4 after infection, indicating the simultaneous expression of EYFP and the corresponding DHS. Nevertheless, SDS-PAGE results showed that although all constructs exhibited similar EYFP fluorescence, the amount of expressed DHS varied significantly and without correlation to the cellular localization of the expressed DHS. To identify the cause for this discrepancy, the EYFP fluorescence intensity should be determined in more detail in relation to the amount of

6.4. The DHS isoforms tolerate a broad range of buffer conditions and they are significantly stabilized in the presence of NAD⁺ and the inhibitor GC7 especially under basic conditions

cells and expressed DHS protein, for example with a sensitive fluorescence plate reader. This approach could unveil if the impression of "comparable" EYFP fluorescence from the images of the 6 well plates might indeed be misleading, and rather represent different magnitudes of fluorescence that the applied consumer camera and the human eye cannot resolve.

However, hitherto no crystalline structures could be identified in the Sf21 cells expressing *S. bicolor* DHS isoforms. In the light of the inconsistent expression described above, it might be necessary to evaluate and optimize the efficiency of the transfection/amplification procedure and the reliability of the expression monitoring by EYFP. In addition, the expression in High FiveTM insect cells should also be tried, since these cells generally express even higher amounts of heterologous protein than Sf21 cells. To prevent any competing situation with the EYFP during protein expression – which might reduce DHS protein yield and could thereby interfere with crystallization – the bMON14272 bacmid, which does not contain the EYFP expression cassette, should also be tested. DHS constructs of this bacmid have already been prepared in the course of this thesis (see section 5.3.2)

6.4. The DHS isoforms tolerate a broad range of buffer conditions and they are significantly stabilized in the presence of NAD⁺ and the inhibitor GC7 especially under basic conditions

The determination of a protein's thermal stability under a variety of different conditions – e. g. by dynamic scanning fluorimetry (DSF, see sections 5.4 and 7.14.5) – is a valuable tool to identify stabilizing and destabilizing agents and parameters. In this thesis, DSF was applied to optimize buffer conditions during purification and identify some hallmarks of *S. bicolor* DHS and HSS enzyme stability. In addition, the effect of the presence of the cofactor NAD⁺ and the DHS-specific inhibitor GC7 on the stability of the enzymes was assessed.

In several DSF measurements with a variety of buffer substances and pH values, the pH optimum of the *S. bicolor* DHS isoforms was determined. In contrast to the activity optimum reported for eukaryotic DHS and plant HSS enzymes of about pH 9.0 to 9.5 [179, 190, 151, 148], both DHS1 and DHS2 as well as the HSS-like DHS3 exhibited their highest thermal stability around the physiological pH, from about pH 7.2 to 7.6. However, the enzymes are stable over a broad range of pH values, with significantly decreasing stability only below pH 5.5 and above pH 9.0 (see section 5.4.1 figure 5.13). For the two eIF-5A isoforms, the pH optimum is also around pH 6.8. This is in accordance with the cytosolic localization of the DHS enzyme and its eIF-5A substrate protein, and supports the assumption that the activity optimum at basic pH is not defined by the DHS enzyme, but is based on the fact that the fraction of diprotonated spermidine – the catalytically relevant substrate – is highest at this pH (see section 4.2.5 for further details). At physiological pH, the triprotonated form is predominant, so that much higher overall spermidine concentrations are necessary to reach comparable activity levels as at basic pH.

Interestingly, the overall thermal stability of the HSS-like DHS3 is with its highest melting temperatures around 50 °C about 10 K less stable than DHS1 and DHS2, which both reach melting temperatures between 60 and 65 °C. This difference was independently confirmed by DLS measurements, where the unfolding of the protein at increasing temperature is monitored

6.4. The DHS isoforms tolerate a broad range of buffer conditions and they are significantly stabilized in the presence of NAD⁺ and the inhibitor GC7 especially under basic conditions

by the increase of its hydrodynamic diameter (see section 7.14.6 for further theoretical details). Since the DHS enzyme activity is crucial for cellular survival and normal cellular function (see sections 4.1), it seems reasonable that the highly conserved DHS enzyme is optimized for high stability. The duplicate DHS3 (HSS), however, is more flexible in terms of evolutionary changes, and since in *S. bicolor* no secondary pyrrolizidine alkaloid pathway is dependent on the stability and functionality of the HSS enzyme, the reduced stability will probably have no detrimental effect on the plant.

In further experiments, the effect of NaCl concentrations between 0 and 500 mM on the three DHS and both eIF-5A isoforms was examined in selected buffers from the previous experiments representing the most stable conditions between pH 6.5 and 9.0. For DHS1 and DHS2, no correlation between the NaCl and the thermal stability could be detected in the applied concentration range and buffer conditions (see section 5.4.2 figure 5.15 A). For DHS3, a small overall decrease in thermal stability was observed for higher NaCl concentrations, although the data was more variable than for the other two DHS. In contrast, for the eIF-5A isoforms the thermal stability was enhanced by the presence of increasing NaCl concentrations. The effect was also pH-dependent and more pronounced for high pH values than for the lower ones (see figure 5.4.2 B).

From these experiments it can be concluded that the DHS enzymes and the eIF-5A precursors are relatively tolerant concerning general buffer conditions and therefore can be handled in a range of buffers featuring different buffer substances and pH values as well as NaCl concentrations. For the eIF-5A isoforms, at least physiological salt concentrations should be applied to stabilize the proteins in basic buffer conditions, otherwise they are also relatively tolerant.

Since previous DLS experiments in the course of the purification experiments already revealed the importance of the presence of NAD⁺ to prevent DHS aggregation during purification and upon flash freezing and storage, low millimolar concentrations of NAD⁺ and the DHS inhibitor GC7 were also tested for their stabilizing potential in different buffers (see section 5.4.3).

Overall, the addition of NAD⁺ and GC7 had a stabilizing effect on all three DHS isoforms in the four tested buffer conditions, which included a physiological PBS condition and the basic glycine buffer that was routinely applied for activity measurements. The effect of NAD⁺ was less pronounced than that of GC7. However, since the storage buffer of the enzymes already contained 2 mM NAD⁺ to prevent aggregation upon storage (see section 5.1 for further details), it might be possible that the active sites were already occupied to a certain extent by NAD⁺ molecules from the storage buffer. Nevertheless, both substances exhibited their highest stabilizing potential under basic glycine buffer conditions, resulting in a difference in melting temperatures of up to 18 K for DHS3 and up to 12 K for DHS1 and DHS2. In further experiments, different inhibitor concentrations ranging from 0.6 nM to 5 mM were tested against a constant NAD⁺ concentration of 1 mM NAD⁺. The stabilizing effect of GC7 became apparent above a threshold of 1.2 μM for *SbDHS3* (HSS), and 2.4 and 9.8 μM GC7 for *SbDHS2* and *SbDHS1*. Based on the law of mass action, at these concentrations approx. 26.1, 52.1 and 99.8% of spermidine binding sites are occupied by the inhibitor GC7 – if calculated with a common K_i of 9.7 nM as determined for DHS from

rat testis [114]. However, these slightly different values could also reflect differences in the K_d values for the three DHS variants. Since the stability increases further with increasing inhibitor concentrations above the threshold value, it is most likely that the effect already becomes apparent upon binding of the first GC7 molecule to the tetramer (relating to 25 % occupancy) and is enhanced by the successive binding of further inhibitor molecules upon rising GC7 concentrations. Taken together, these calculations suggest that the stabilizing effect of GC7 is indeed due to the specific binding to the active site and that a stabilizing effect can already be achieved with low micromolar concentrations of the inhibitor at enzyme concentrations of 0.2 g L^{-1} (corresponding to about $4.6 \mu\text{M}$ active sites). For significant effects, however, low millimolar inhibitor concentrations should be applied.

6.5. Characterization of the binding between the different *S. bicolor* DHS and eIF-5A isoforms by bio-layer interferometry (BLI)

The deoxyhypusination reaction catalyzed by the DHS enzyme is a unique, but ubiquitous reaction in eukaryotes and archaea, and it exhibits a high specificity towards its only protein substrate, the initiation factor 5A (eIF-5A). The binding of eIF-5A to the DHS enzyme is most likely mediated by electrostatic interactions between the surface area of the funnel-like entrance to the active site of the DHS tetramer on the one side and on the other side the surface of the *N*-terminal domain of eIF-5A, which carries the finger-like loop with the modified lysine residue (see sections 4.2.2 and 4.2.3 for more information).

In several angiosperm lineages, the side reactivity of the DHS enzyme – the production of homospermidine from the aminobutyl moiety of spermidine and the small diamine putrescine as acceptor instead of the lysine residue of eIF-5A – evolutionary developed into a separate homospermidine synthase (HSS) upon duplication of the original DHS coding sequence (see section 4.3 for further details). In many of these plant lineages, the separate HSS became the first pathway-specific enzyme of the pyrrolizidine alkaloids biosynthesis, which produces toxic pyrrolizidine alkaloids to protect the plants from herbivores. The specialization towards the HSS functionality is in most cases accompanied by a loss of DHS activity – which is assumed to be due to the disappearance or alteration of the eIF-5A-specific binding surfaces on the DHS tetramer [192].

As already mentioned earlier, three DHS-like enzymes from the Poaceae model plant *S. bicolor* were examined for their biophysical characteristics in the course of this thesis. These three copies are supposed to represent three different stages of angiosperm DHS/HSS evolution – with a nearly full-fledged HSS (DHS3), and two DHS isoforms, which might represent the original DHS and an intermediate enzyme. While activity assays exhibited very low DHS activity and improved homospermidine synthesis rates for *S. bicolor* DHS3, which is therefore better classified as HSS (A. Wesseling⁴, internal communication), no distinct differences in the DHS and HSS activity of the two other DHS (DHS1 and DHS2) were detected. To identify more subtle changes between these two DHS – especially differences in the DHS

⁴Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober, Botanical Institute of Kiel University

surfaces that might influence efficient eIF-5A binding – this interaction was studied in more detail by bio-layer interferometry (BLI).

BLI is a sensitive technique for the time-resolved measurement of association and dissociation kinetics of proteins and even small molecules (see sections 5.5 and 7.14.8 for further details). It does only require comparatively small amounts of sample protein and the applied Octet Red96e system (PALL *forté*BIO) allows up to eight parallel measurements. While the BLI method itself does not need any molecule labeling, it is based on the immobilization of one binding partner – the ligand – to the tip of the biosensor, and the association and dissociation of a free analyte applied in different concentrations in solution. For the immobilization of the ligand, several established procedures are available, e. g. covalently via amine coupling, reversible by affinity-tag-based approaches (His₆-tag) or semi-reversible via covalently attached antibodies, which bind the ligand stably and with high specificity. In this thesis, an antibody-based approach was implemented, as detailed in section 5.5.

Therefore, *SbDHS1*, -2 and -3 as well as the two eIF-5A isoforms were equipped with an SA*StrepII*-tag, which is bound with very high affinity by an antibody that was specifically designed for the high demands of BLI and surface plasmon resonance applications (SPR, a competing technique to BLI). The BLI measurements were designed in both orientations – termed binding sequences 1 and 2 –, either with DHS as ligand and eIF-5A as analyte (2) or *vice versa* with eIF-5A immobilized on the biosensor by the antibody, and DHS in solution (1) (see section 5.5 figure 5.18).

The experiments were performed at two distinct buffer conditions. The first condition was inspired by the conditions routinely used for activity measurements in the research group of Prof. Dr. Dietrich Ober, with a glycine buffer at pH 9.0 and a physiological salt concentration. However, although the DHS reaction demonstrably reaches its maximum activity at basic pH, this is an artificial optimum created by the highest portion of optimally charged spermidine substrate at this pH (see section 4.2.5 for further information [11, 114, 153, 280]). In contrast, thermal stability experiments discussed earlier in this thesis revealed that DHS itself has its stability optimum around the physiological pH of about 7.4 – as can be expected for a cytosolic protein (see section 5.4). Therefore, the second buffer condition – a PBS buffer at pH 7.5 – was chosen to provide a more "natural" environment for the DHS and its binding partner eIF-5A.

6.5.1. *S. bicolor* DHS3 ("HSS") does not bind to either of the two eIF-5A isoforms

Initial binding experiments were performed in both buffers to determine cornerstone values for the following experiments. In section 5.5.1, some of these initial experiments are described for the basic glycine buffer system. A crucial prerequisite for the generation of solid binding kinetic data is the stable binding of the ligand over the course of the experiment. In the chosen system, this is achieved by two highly specific and affine binding events: the interaction of the covalently coupled anti-mouse Fc antibody with the Fc portion of the *StrepMAB*-Immo antibody, and the binding of the actual SA*StrepII*-tagged DHS or eIF-5A ligand to the *StrepMAB*-Immo antibody. While the latter interaction was sufficiently stable under both buffer conditions, the first interaction did not tolerate the basic glycine buffer and the *StrepMAB*-Immo antibody – and thereby the bound ligand – was constantly dissociating

during the experiment (see section 5.5.1 figure 5.19). A sensor prepared without analyte was applied in parallel to account for this signal loss and still enabled data interpretation, but the basic buffer conditions impaired the binding characteristics of the covalently immobilized anti-mouse Fc antibodies and prevented biosensor reuse. Therefore, binding kinetics in glycine buffer were only determined once in the course of this thesis.

As described above, the binding experiments with *S. bicolor* DHS and eIF-5A isoforms were designed in two orientations, with either of the interacting partners once as ligand and once as analyte. By measuring both possible orientations, the obtained kinetic values can be confirmed and external contributions and artifacts from the measurement setup are prevented or at least more easily identified. However, in the preliminary experiments with both buffer systems, a specific binding signal could only be detected between eIF-5A ligand and DHS analyte (binding sequence 1) at 100 nM. In the opposite orientation, no binding signal was measurable. Higher analyte concentrations of 10 μ M for both eIF-5A and DHS analytes led to the unspecific binding of analyte molecules to sensors with and without ligand to a comparable extent.

The difference in binding behavior depending on the bound analyte – eIF-5A or DHS – is most likely due to an adverse immobilization of the DHS molecule as analyte on the biosensor. The SAStrepII-tag is localized at the *N*-terminal end of each subunit, a position that did not interfere with enzyme activity in solution (A. Wesseling, internal communication), and that even boosted overexpression in *E. coli* (see section 5.1.2). However, the *N*-terminal end of the DHS subunit comprises a flexible so-called ball-and-chain motif – a terminal two-turn α -helix ("ball") attached to an extended loop structure ("chain") –, which was assumed to play a regulatory role (see section 4.2.2). A deletion of the first 20 amino acids – the "ball" motif – completely abolished DHS activity, emphasizing the important role of this *N*-terminal portion [123]. It is therefore well conceivable that the immobilization of the DHS tetramer via the *N*-terminally attached SAStrepII-tag interferes with the binding of the eIF-5A analyte. Since the tetramer is composed of four subunits that all possess the *N*-terminal affinity tag, the tetramer might actually be bound via all four tags and thereby be rigidly oriented in a position that does not allow access to the active site funnel area. In addition, the ball-and-chain motif is most likely held in a fixed position by the affinity tag-mediated interaction with the biosensor surface. This fixed and maybe even strained position could directly block the access to the surfaces necessary for eIF-5A binding or exert a more indirect effect on the tetramer conformation, which might lock the enzyme in an inactive conformation – similar to what was proposed for the cooperativity effect induced by the binding of eIF-5A (see section 4.2.4).

For future experiments, the SAStrepII-tag could be attached to the *C*-terminus of the DHS subunits to prevent detrimental effects by interference with the flexibility of the ball-and-chain motif, which might be crucial for successful interaction with the eIF-5A precursor. However, if the lack of binding between DHS ligand and eIF-5A analyte is due to an adverse orientation of the complete DHS tetramer rather than to the restraints of the ball-and-chain motif, a *C*-terminal SAStrepII-tag at all four subunits could result in a comparable inaccessible orientation.

Nevertheless, since neither of the available *N*-terminal SAStrepII-tagged DHS immobilized as ligands yielded any eIF-5A binding signal, all subsequent binding experiments and kinetic measurements were therefore performed with SAStrepII-tagged eIF-5A isoforms as ligands and the DHS isoforms as analytes.

Further preliminary binding experiments addressed the binding behavior of the three *S. bicolor* DHS isoforms to both eIF-5A isoforms (see section 5.5.1 figure 5.19 E for glycine buffer results). For the two eIF-5A isoforms, DHS1 and DHS2 – which also showed indistinguishable DHS activity in activity assays (A. Wesseling, internal communication) –, exhibited binding at a concentration of 100 nM in both tested buffer conditions, while DHS3 in contrast did not bind to either of the eIF-5A isoforms. Binding experiments at higher analyte concentrations of 10 μ M resulted in unspecific binding signals for sensors loaded with and without ligands, as already mentioned above. This confirms that the loss of DHS activity in the *S. bicolor* "HSS" variant DHS3 is indeed due to the loss of the binding capability for the eIF-5A precursor, as was already assumed from earlier affinity chromatography-based binding experiments and non-denaturing electrophoresis performed with *Senecio vernalis* HSS [192].

Considering the sequence alignment of the three DHS-like sequences from section 4.3 figure 4.11 and the position of the deviating amino acid residues in DHS3 as compared to DHS1 and DHS2 (see figure 4.12), only very few amino acid residues can be responsible for the abolished eIF-5A binding. In previous sequence analyses of a variety of HSS and DHS sequences from different angiosperm lineages, especially two residues have been proposed as telltale residues for the distinction between DHS and HSS: 274, which is an isoleucine in all characterized plant DHS, and a valine (or alanine) in the corresponding HSS, and 278, which is an asparagine in DHS and an aspartic acid in HSS enzymes (numbering based on *SbDHS2*) [120, 155]. In *SbDHS3*, however, only the first I274V substitution is present, while the residue at position 278 is still an asparagine. The only further changes in the vicinity of the entrance to the active site are R262N and F309L, which could directly contribute to a change in the binding surface. All other amino acid changes are less likely to be involved in the surface that binds the eIF-5A precursor. Future binding experiments with V274I, N262R and L309F variants of *SbDHS3* would therefore be of high value. The V274I variant could easily clarify if this telltale substitution alone is indeed responsible for the loss of eIF-5A binding – indicating a more indirect, maybe cooperative effect due to the buried position of this residue (see figure 4.12) – or if other substitutions are also necessary, like the N262R and L309F variants mentioned above, which could directly impair the binding interface.

Since the *N*-terminal ball-and-chain motif was already demonstrated to be of vital importance for the interaction of the human DHS enzyme with the eIF-5A precursor (see above and also section 4.2.2, [123]), another conceivable *SbDHS3* variant could encompass the substitution of the HSS *N*-terminal part by the *N*-terminus of one of the DHS enzymes. According to the sequence alignment 4.11, this part of the HSS sequence is almost completely different than that of both DHS sequences. If the ball-and-chain motif is indeed also important for proper binding of the eIF-5A precursor in plant DHS, these rather large discrepancies could be also responsible for the missing interaction.

6.5.2. Analyses of binding kinetics reveal low nanomolar to sub-nanomolar affinities between *S. bicolor* DHS and eIF-5A isoforms

In further experiments, the kinetic constants of the binding between *S. bicolor* DHS1 and DHS2 and the two eIF-5A precursor proteins were determined under basic glycine and physiological PBS buffer conditions. Therefore, the binding curves of a 1:2 dilution series of DHS1 or DHS2 analyte were measured against each of the two eIF-5A ligands, with 6–7 concentrations in the range of 0.39–200 nM – depending on the interacting partners (see sections 5.5.1 and 5.5.2 for measurement details). As explained in the previous section, experiments in glycine buffers were only performed once since the biosensors could not be reused. Experiments in PBS buffers, however, were each performed once with fresh biosensors and at least twice with regenerated ones.

Based on literature data, the stoichiometry of the DHS(tetramer)–eIF-5A(monomer) complex was expected to be 1:1. Although the DHS enzyme contains four seemingly identical active sites, it was proposed that a conformational change upon binding of the first eIF-5A molecule prevents binding of further molecules (see section 4.2.4). The binding curves obtained by BLI measurements were therefore fitted in a first step by the classical 1:1 (Langmuir) binding model, which assumes that the association and dissociation curves can be described by two single exponential functions, the integrated rate equations (see section 5.5.1 equations 5.1 and 5.2 [196]). The association and dissociation rate constants k_a and k_d were derived from a global fit of the measured curves of the individual interactions, and the affinity constant K_D was calculated from k_a and k_d .

For the measurements performed in glycine buffer, the association constants k_a were in the range of 2.4×10^5 to $3.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, and the dissociation constants k_d were determined to be about 1.3×10^{-3} to $2.7 \times 10^{-3} \text{ s}^{-1}$. With these values, affinity constants K_D were calculated to be in the range of 3.7 to 11.3 nM (see table 5.1). The highest association and lowest dissociation constants – resulting in the lowest K_D value – were observed for the interaction between DHS2 and eIF-5A2, the lowest association constant in combination with the highest dissociation constant – and therefore the highest K_D value – was determined for the interaction between eIF-5A1 and DHS1. Overall, the binding of both DHS to eIF-5A2 was minimally tighter than to eIF-5A1, and both eIF-5A exhibited a slightly higher affinity towards DHS2 than towards DHS1. However, since the corresponding measurements were only performed once and the accuracy of the actual values also depends on the accuracy of the initial protein concentrations used for the 1:2 dilution series, the exact constants should not be over-interpreted.

The measurements performed in PBS buffer at pH 7.5 resulted in even higher affinity constants K_D in the range of 0.3 to 0.9 nM, with association constants k_a from 1.5×10^6 to $3.0 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ and dissociation constants k_d from 0.9×10^{-3} to $1.7 \times 10^{-3} \text{ s}^{-1}$ (see table 5.2). These values – as well as the constants obtained from glycine buffer experiments – are in accordance with the K_D of about 0.5 nM (or even lower) determined for human DHS by equilibrium ultracentrifugation analysis [151]. However, a more thorough interpretation of the data is not suitable. As elaborated already in some detail in section 5.5.2, all binding curves do not exhibit a perfect exponential curve shape, as should be expected for a true 1:1 binding event. The heterogeneity is negligible in the curves measured in glycine buffer, but becomes more apparent in the curves measured in PBS buffer, especially in the binding

curves of *SbDHS1*. Since for BLI measurements optimal analyte concentrations should range between 0.1x and 10x of the expected K_D value, the observed heterogeneity can in part be attributed to the relatively high analyte concentrations in comparison to the nanomolar or even sub-nanomolar K_D . For data analysis, the highest and most heterogeneous curves were therefore already excluded (see sections 5.5.1 and 5.5.2 for details). Still, even the lower curves exhibit characteristics of a combination of a slower and a faster binding and dissociation event.

6.5.3. Heterogeneity in binding behavior cannot be attributed to a distinct binding model

In an attempt to determine the cause for the observed heterogeneity, further binding models besides the simple 1:1 model were also tested. After the exclusion of mass transport effects, which would limit the observed association rates to the diffusion/convection rates in the analyte solution, the 1:2 bivalent analyte model and the 2:1 heterogeneous ligand model were applied to the data. Since the DHS analyte can in theory provide up to four binding sites – for the moment taking putative cooperativity effects and the assumed 1:1 stoichiometry out of the equation –, the 1:2 bivalent analyte model seemed a sensible choice for data analysis. However, the model did not represent the data much better than the originally applied 1:1 model (see figure 5.21 E and F).

In contrast, the 2:1 heterogeneous ligand model provided a very good representation of the data (see figure 5.21 G and H). This model encompasses essentially three distinct situations: a conformational change upon binding (1), a heterogeneous ligand that provides two different bindings sites (2), and a bivalent ligand with two identical binding sites (3). Based thereon, the association and dissociation curves can be described by the sum of two exponential integrated rate equations for the two binding events – or binding and subsequent conformational change (see section 5.5.3 equations 5.3 and 5.4). Considering the previous experiments reported in literature, which propose a conformational change in the DHS tetramer upon binding of the eIF-5A precursor that prevents binding of further precursor molecules (see section 4.2.4), the conformational change model might indeed represent the actual biological situation better than the simple 1:1 binding model.

From the biological system's point of view, the other two models assuming a heterogeneous ligand situation are much less likely. The SEC-MALS experiments discussed in section 6.6 clearly showed that the eIF-5A isoforms are not dimerized – as was first speculated based on DLS data (see section 5.1.2) –, but monomeric. In addition, the eIF-5A ligands are uniformly oriented on the biosensor surface by the defined immobilization via the single C-terminal SA*StrepII*-tag, which is the best strategy to obtain a homogeneous surface and prevent heterogeneity in the measured interactions.

Nevertheless, the observed heterogeneity cannot unequivocally be attributed to one of these binding models solely based on these assumptions. Therefore, a mathematical approach presented by Tiwari et al. in 2015 was applied, that claimed to enable the distinction between these three 2:1 binding models – if they were indeed causing the heterogeneous binding behavior [272]. Therefore, the association and dissociation curves are fitted with the above-mentioned double-exponential function – the sum of the two integrated rate equations – and

the complex constants σ and γ from the exponents of the rate equations are determined. Depending on the underlying binding model, these constants contain information on the analyte concentration and the k_a and k_d values of the two contributing events (see table 5.3 for relationships). If the sums and products of σ constants and their individual values from each measured binding curve are plotted against the analyte concentration C of the corresponding measurements, the resulting data points should either exhibit a linear or quadratic correlation. Based thereon, the underlying binding model could be identified as summarized in table 5.3.

This analysis was performed with two kinetic measurements for each of the four interactions. Unfortunately, seven of the eight analyses did not yield coherent results, as depicted exemplarily for four of the analyses in section 5.5.3 figure 5.22. Neither a clear linear correlation, nor a quadratic one could be assigned to the plots of the sums, products and individual values of σ over analyte concentration C . Only one analysis seemed to deliver sensible results: For the binding between DHS1 and eIF-5A1, the plots of sum and product, as well as one of the individual σ values over C exhibited linear behavior, while the second individual σ values seemed to be non-linear. Based thereon, the underlying binding model would be assumed to be the two-step conformational change model, which would at least be plausible considering the previously proposed cooperativity effect upon eIF-5A binding to DHS. However, since only one of eight analyses yielded interpretable results, this assumption is still only speculative.

Based on these analyses, the actual cause for binding heterogeneity observed in the kinetic measurements performed in the course of this thesis cannot be assigned with certainty. If the characteristics of the binding between the two DHS and the two eIF-5A isoforms need to be interpreted in more detail, it is necessary to perform further experiments: on the one hand to determine the exact stoichiometry of the *S. bicolor* DHS–eIF-5A complexes – to become independent from assumptions based on experiments with e.g. human or yeast DHS and enable an evidence-based assignment of the appropriate binding model –, and on the other hand to exclude any unspecific effects derived from experimental setup.

To determine the stoichiometry of the complex *S. bicolor* DHS–eIF-5A complexes, analytical ultracentrifugation or non-denaturing electrophoresis could be applied, as was done for human DHS–eIF-5A complexes [151, 56]. Furthermore, small-angle X-ray scattering (SAXS) could be performed to obtain information on complex shape, stoichiometry and also provide complementary data on the K_D of the complex formation (e.g. as reviewed in [277]). In addition, CryoEM structures of the complexes could be determined as already proposed in section 6.3, and of course a complex structure determined by X-ray crystallography would add detailed information on the actual interaction and putative cooperative changes in the DHS structure upon complex formation. Since preliminary SEC-MALS experiments indicated that the complex might not be stable enough to be isolated by SEC or electrophoretic methods (see section 6.6), chemical cross-linking of the complex could also be tried to simplify complex isolation and stoichiometry determination. Native mass spectroscopy – maybe even in combination with complex cross-linking – would also allow further insights into the complex composition [103, 102].

To optimize the experimental setup for further BLI measurements, the ligand density on the biosensor should be evaluated and optimized, if necessary. Since the experiments performed in this thesis point in the direction of an inhomogeneous situation at ligand level, it might

be worthwhile to reduce the density of eIF-5A ligand on the biosensor surface. Therefore, two parameters should be addressed: the density of the capturing *Strep*MAB-Imm antibody, and the density of the captured eIF-5A ligand. In both cases, less density can be achieved by lowering the applied loading concentration and decreasing the loading time. Less ligand on the surface will further contribute to surface homogeneity, as the occurrence of highly accessible and less accessible binding sites due to putative steric hindrance can be reduced. Considering the measurement results discussed in this section, this approach seems to be the most promising. In addition, careful consideration must be taken in the choice of analyte concentrations, since higher analyte concentrations – especially concentrations way beyond 10x of the expected K_D value – might give rise to curve heterogeneity due to unspecific binding.

6.5.4. The inhibitor GC7 prevents binding of eIF-5A to DHS and interrupts existing complexes

In the course of this thesis, DLS and DSF experiments already demonstrated that the cofactor NAD^+ and the inhibitor GC7, a spermidine analog, exert stabilizing effects on the analyzed DHS tetramers. In addition, the supplementation of NAD^+ rescued higher order DHS aggregates and shifted a putative equilibrium situation between the active tetramer, putative octamers and even hexadecamers back to the predominance of the tetrameric form (see sections 5.1.2 and 5.4). Consequently, the influence of these two substances in the complex formation between *S. bicolor* DHS and the eIF-5A precursor were also examined by BLI.

In individual experiments, the effect of NAD^+ and GC7 on the association and dissociation phases were evaluated. For the inhibitor GC7, significantly increased dissociation rates were observed for the tested inhibitor concentrations in the range of 1 to 5 mM. This effect was even enhanced by the additional supplementation of NAD^+ at the same concentrations: at 5 mM GC7 and NAD^+ , the dissociation was accelerated about 8x compared to the situation without either component (see section 5.5.4 figure 5.23 A and B). This indicates a dependence of the GC7 effect on the prior binding of NAD^+ and confirms literature reports, in which spermidine and GC7 binding required prior binding of NAD^+ to the active site of the enzyme [148, 149]. In the association phase, binding of 25 nM DHS to eIF-5A was completely abolished in the presence of 1 mM GC7, which correlates to an inhibitor saturation of the DHS tetramers of 100% according to the law of mass action. In contrast, 1 nM GC7 had no detectable effect in the DHS–eIF-5A interaction, which is in accordance with an inhibitor saturation of only 0.8%.

These results are especially interesting: While literature attributes the inhibitory effect of GC7 (K_i about 9.7 nM [114]) to its competition with the actual spermidine substrate (K_D about 1 μM [148]), this data allows further insights into the underlying inhibition mechanism. Based on these binding data, the inhibitor GC7 not only competes with spermidine for the binding site and thereby displaces the productive substrate, but obviously induces a conformational change in the DHS enzyme that interrupts the existing interaction between a DHS molecule and the eIF-5A precursor and even prevents binding of further precursor molecules. To my knowledge, this is the first time that the interaction between a DHS and the eIF-5A precursor has been examined in the presence of the GC7 inhibitor. Whether this behavior is

specific to *S. bicolor* or plant DHS, or is a general feature of the GC7 inhibition mechanism, should be addressed in further experiments with DHS from other eukaryotic species. Furthermore, other inhibitory spermidine analogs could be tested whether they also affect the DHS–eIF-5A complex (see [114] for a comprehensive overview on DHS inhibitors). If this is in fact a general feature, crystal structures with and without the inhibitor would provide valuable insight into the structural details of this binding inhibition. In the light of these findings, the interpretation of the currently available structure of human DHS in complex with GC7 could also be reviewed.

As a side note, the drastic increase of dissociation rates in the presence of 5 mM GC7 and NAD^+ was of high value for the experiments at hand, since they enabled reliable biosensor reuse – after classic regeneration approaches failed (see section 5.5.4 for further details).

While the supplementation of GC7 alone has a rather drastic effect on both association and dissociation phase, the separate addition of NAD^+ only influenced the dissociation rates, as far as tested in the course of this thesis. As described above, NAD^+ significantly increased the dissociation rates in combination with GC7, but on its own it had the contrary effect: The supplementation of dissociation buffers with 1 mM NAD^+ slowed the dissociation process down to about 2% of the original rate. The association phase was not influenced by the additional presence of NAD^+ in these experiments, however, it has to be noted that the protein storage buffers – from which the dilution series were prepared – already contained 2 mM of the cofactor to prevent aggregation (see section 5.1). A lack of any detectable effect might therefore be due to the fact that the DHS tetramers still contained NAD^+ from the storage buffer. Further experiments could address this by prior removal of the cofactor via dialysis.

6.6. Complex between *S. bicolor* DHS and eIF-5A precursor does not survive size exclusion chromatography

The biophysical experiments described in the previous sections all conclusively show the importance of NAD^+ for the stability and homogeneity of DHS enzyme samples. During lysis and purification procedures, the supplementation of the utilized buffers ensured homogeneous binding behavior and prevented oligomerization/aggregation into higher order molecule assemblies upon flash freezing and storage, as DLS measurements revealed (see section 5.1.2). In the thermal stability assays performed with DSF, the presence of NAD^+ – derived from the storage buffer or additionally supplemented – was most likely the basis for the stabilizing effects that the inhibitor GC7 could exert, assuming that the binding of GC7 is probably dependent on the prior binding of NAD^+ to the active sites of the DHS enzyme (see section 5.4.3). Finally, the binding studies with BLI revealed that the presence of low millimolar concentrations of NAD^+ drastically decreased dissociation rates between the immobilized *Sb*eIF-5A ligand and the DHS analyte compared to the situation without additionally supplied NAD^+ – thereby significantly stabilizing the DHS–eIF-5A complex (see section 5.5.4).

This stable interaction would be of great value for the crystallization and subsequent X-ray diffraction-based structure determination of a *S. bicolor* DHS–eIF-5A complex. For human DHS in complex with eIF-5A it could be shown that the complex was stable enough to at

least partially withstand a non-denaturing gel electrophoresis or the capture via His₆-tag-based IMAC [151, 56]. In this thesis, the individual *S. bicolor* DHS1, DHS2, eIF-5A1 proteins and mixtures thereof were subjected to size-exclusion chromatography (SEC) and a subsequent on-line multi-angle light scattering (MALS) analysis step to determine if the constant presence of 1 mM NAD⁺ in sample and running buffers stabilized the DHS–eIF-5A complexes enough to enable their isolation (see section 5.6).

The analysis of scattering intensities obtained at several fixed angles (MALS) in combination with the upstream particle separation by size exclusion chromatography (SEC) allows a quite accurate determination of the absolute molar mass and root mean square (rms) radius of molecules or molecule assemblies in solution. This is in contrast to results obtained from batch dynamic light scattering experiments (DLS) – as performed in section 5.1.2 –, which determine a distribution of particle sizes in solution and yield an average diameter, which represents a sphere that would move in the same way as the scattering particles. With the assumption of a globular particle, the molecular weight is calculated from this theoretical sphere. Unfortunately, this calculation can be misleading – depending on the actual shape of the particle.

However, the SEC-MALS results for the two individual DHS tetramers in the presence or absence of NAD⁺ in sample and running buffers confirmed the DLS measurements obtained in earlier experiments: The main peak of the SEC corresponded to a molecule with a molecular weight of approx. 161 kDa, which is about the estimated weight of the homotetrameric DHS enzyme. In addition to this main peak, a small peak with a molecular weight of approx. 325 kDa is also visible – indicating a dimer of DHS tetramers (octamer). This phenomenon had already been observed for the polishing SEC step during DHS purification (data not shown), together with a peak of even higher molecular weight, which might correspond to the "tetramer of tetramer" situation detected for oligomerized/aggregated DHS enzyme after flash freezing and storage in the absence of NAD⁺ (see section 5.1.2). Since the presence or absence of NAD⁺ in the SEC running buffer had no effect on the occurrence of the "dimer" (or octamer) peak, it seems as if the formation of higher order oligomers is an inherent feature of the *S. bicolor* DHS and a kind of equilibrium state, that can be shifted towards the predominance of the tetrameric species by the addition of NAD⁺, however with a small portion of octamers – and most likely an even smaller amount of hexadecamers – still present.

For the separately analyzed *S. bicolor* eIF-5A1, the SEC-MALS analysis unambiguously answered the question posed earlier, whether *S. bicolor* eIF-5A is in fact a homodimer upon *E. coli* expression and affinity purification – as indicated by the calculated molecular weight of about 42 kDa obtained by DLS data analysis (see 5.2.2) – or a monomer. The SEC chromatogram only exhibits a single, monodisperse peak with a determined molecular weight of 17 kDa, which clearly agrees with the expected molecular weight of the eIF-5A1 monomer of 18.5 kDa. The DLS data is therefore overestimating the actual molecular weight of the eIF-5A molecules, most likely due to their non-globular shape.

The SEC-MALS analyses of the putative complexes of *S. bicolor* DHS1 or DHS2 with eIF-5A1, however, were disillusioning: While the SEC chromatograms in the presence and absence of NAD⁺ in the SEC running buffer exhibited the same peaks with comparable molecular

weights as obtained for the separate analysis of DHS and eIF-5A1 samples, no further peak corresponding to a complex of a DHS and the eIF-5A1 could be detected. This indicates that despite the high affinity between DHS and eIF-5A determined by BLI measurements and although the presence of NAD⁺ significantly increased complex stability and therefore its overall lifetime, the DHS–eIF-5A complexes were not stable enough to survive the size exclusion chromatography step. Based on these results, the formation of complexes for future crystallization experiments will still have to be induced *in situ* in the crystallization solution. Therefore, the DHS enzymes – or variants thereof that might exhibit improved precursor-binding, like D243A or K329A (see section 6.3.1) – have to be combined with carefully selected concentrations of eIF-5A and NAD⁺, but obviously in the absence of the inhibitor GC7, as these BLI experiments impressively demonstrated.

6.7. Evidence for the existence of paralogous heteromers comprised of DHS and HSS subunits

The interaction of proteins in well-defined complexes is often a crucial prerequisite for their functionality. In the case of eukaryotic and archaeal DHS enzyme, the individual subunits form dimers, which comprise two antiparallel active sites at their interface. Two of these dimers congregate to form the obligate DHS homotetramer – a "dimer of dimers" (see section 4.2.2). While the first dimerization is obviously essential to create the active sites, the formation of the tetramer might have a regulatory role: Literature reports a 1:1 stoichiometry of DHS(tetramer)–eIF-5A(monomer) complexes, although the DHS tetramer provides four seemingly identical binding sites. It was therefore assumed that the binding of the first eIF-5A substrate protein induces a conformational change, that prevents binding of further initiation factors (see section 4.2.4).

Several times during angiosperm evolution, independent duplications of the DHS coding sequence occurred. These duplicates allowed the development of a separate homospermidine synthase (HSS) from a side reactivity of the original DHS in these plant lineages (see section 4.3). The new HSS lost the DHS-specific affinity for the eIF-5A substrate (see previous section), optimized homospermidine synthase activity and – in many cases – became the trailhead enzyme of the pyrrolizidine alkaloid (PA) biosynthesis pathway, which produces plant defense compounds against herbivores. However, since the homospermidine synthesis is performed at the same active site as the original deoxyhypusination reaction, the evolution of the HSS enzyme is constrained to maintain integrity and functionality of the active sites – and thereby the interacting interfaces of the respective subunits that encompass them.

In the first period after the duplication event, the original DHS and its paralogous duplicate are virtually identical, and their subunits are able to interact with each other. An evolutionary change of the HSS duplicate could severely impair the essential modification of the eIF-5A precursor by the DHS enzyme via their mutual complexes, which subsequently might have devastating effects on all cellular processes that depend on modified eIF-5A (see section 4.1). The HSS therefore inherits the selective pressure of the DHS, and its evolution is restricted by the cross-interaction of the two paralogs – a phenomenon called "paralog interference" [12, 121]. In this period, two evolutionary scenarios can take place: either both paralogs

co-evolve by positive selection and thereby maintain and optimize the original DHS function, or changes in either of the paralogs are prevented by a purifying selection, since detrimental changes in one paralog could affect homomeric complexes and paralogous heteromers alike. For the DHS/HSS pair, it is most likely that the paralogs are subjected to the purifying selection process, since the integrity of the interaction between the four subunits is essential for the creation of the active site, the formation of the eIF-5A binding site and the regulation of enzyme activity.

Based on this assumption, it seems reasonable that evolutionary changes in the HSS paralog were carefully selected to maintain the integrity of the interfaces, and thereby also preserving the interaction between DHS and HSS subunits. However, until today, the interaction between paralogous angiosperm DHS and HSS subunits was only speculative. Therefore, experiments based on western blot detection of paralogous heteromers were performed in the course of this thesis to confirm the existence of paralogous heteromers as a proof of the described evolutionary scenario (see section 5.7).

In these experiments, two DHS homologs (DHS1 and DHS2) and their HSS paralog (DHS3) from model plant *Sorghum bicolor* have been cloned into three coexpression constructs that enable the parallel expression of two homo- or paralogous subunits in *E. coli*. In these constructs, DHS1 carrying a C-terminal His₆-tag was combined with either C-terminally *StreptII*-tagged DHS3 (HSS) or DHS2, and with DHS1 as a positive control for the interaction and experimental set-up. To exclude any construct-derived effects, six control constructs were created, which contained DHS1, DHS2 or DHS3 (HSS) for the separate expression of the His₆- and *StreptII*-tagged subunits.

All constructs were expressed in *E. coli* BL21-Gold(DE3) and initially purified by IMAC (all coexpression constructs and the His₆-tagged control constructs) or AC (*StreptII*-tagged subunits), and SEC as established in section 7.12.2. From the separately expressed His₆- and *StreptII*-tagged constructs negative controls were prepared by mixing DHS1-His₆ with either of the *StreptII*-tagged constructs and incubate them overnight. Afterwards, coexpressed protein as well as the positive and negative controls were purified a second time in a small scale via the His₆-tag and washed rigorously to remove any unspecifically bound subunits from the captured complexes. The eluted samples were separated by SDS-PAGE, blotted and the *StreptII*-tag detected by a color reaction with AP-conjugated *Strep-Tactin*[®].

The western blot revealed what had been already suspected: The *StreptII*-tag could be clearly detected in the DHS1–DHS1 positive control and in both coexpression samples, which either contain the homologous subunits DHS1 and DHS2 or the paralogs DHS1 and DHS3 (HSS). The signal intensity of the paralogous heteromers was similar to that of the positive control – maybe somewhat less pronounced –, so that a distinct preference for the original DHS1 homotetramer over either of the heteromers could not be deduced from these results. This is in accordance with the previously described evolutionary concept and the fact that the interacting interfaces are highly conserved between the DHS and HSS subunits and only contain few amino acid substitutions (see figure 4.12). These results demonstrate the existence of angiosperm DHS/HSS paralogous heteromers for the first time.

Interestingly, the negative controls, which contained separately expressed and afterwards mixed subunits, also exhibited a faint *StreptII*-tag signal in the western blot. Samples from

separately expressed His₆-tagged DHS or HSS did not exhibit any signal, confirming the specificity of the detection method. The signal is very weak compared to the signal from coexpressed samples, so that it can be assumed that the *StreptII*-tag signal detected from the latter samples indeed corresponds to the presence of heterologous tetramers. Since the samples immobilized by their His₆-tag were washed intensively with high amounts of high salt and detergent containing buffer, these signals most likely also come from a heterologous tetramer, that formed upon mixing of the separate subunits during overnight incubation. Although the interaction in the DHS tetramer are mediated through extensive interfaces (see section 4.2.2), the complex formation is still a dynamic equilibrium situation between subunit association into complexes and their dissociation into free subunits, however with the emphasis strongly on the side of the tetramer formation. It is therefore possible that single subunits – or maybe even complete active dimers – might be shuffled over time, thereby resulting in the faint *StreptII*-tag signal in the western blot.

In further experiments, this phenomenon could be addressed to determine the subunit exchange kinetics. Mass spectrometry has already been applied successfully to gain insights into the formation of stable protein complexes and the exchange rate of their subunits over time. Therefore, even small differences in the mass of the analyzed subunits are sufficient [263]. Another approach could exploit the time-resolved Förster resonance energy transfer (FRET), monitoring the development of a FRET signal upon mixing of separately labeled DHS or HSS tetramers [54] – or the decrease of the FRET signal if tetramers prepared for FRET are mixed with unlabeled tetramers.

However, although the western blot experiments performed in the course of this thesis demonstrate the presence of heteromeric complexes of *S. bicolor* DHS and HSS paralogs, their exact composition remains unknown. Since the DHS – or HSS – are composed as a "dimer of dimers" from two dimers, which comprise each two antiparallel active sites at their interface, it is conceivable that the subunits are not only exchanged separately, but rather as dimeric complexes, and that the heterotetramers do not consist of random ratios of the two homo- or paralogous subunits, but maybe of a defined situation of two homomeric dimers from two different homo- or paralogs forming a heterotetramer.

In the future, activity assays could be performed to determine if heteromers of HSS and DHS subunit exhibit different or impaired activity for deoxyhypusine and homospermidine synthesis. However, as long as the composition of the heteromers is not known, such activity results will only represent an average of the underlying situation. Therefore, prior experiments are reasonable to clarify the heteromeric situation. The already existing affinity tags could be used to supply one of the paralogous subunits with an additional mass – for example a tightly and specifically binding anti-*StreptII*-tag antibody (as used in BLI ligand immobilization, see section 5.5). The resulting mixture could subsequently be analyzed by SEC-MALS to determine the ratio of tetramers carrying no *StreptII*-tag, or one, two, three or four *StreptII*-tags. If only tetramers without *StreptII*-tag, or with two or four *StreptII*-tags are present, it would be very likely that not individual subunits are exchanged, but pre-assembled dimers. In the other case, it might be possible to correlate the ratio of the different species with activity results.

Taken together, the western blot results obtained in the course of these experiments confirm the presence of the predicted heterotetramers composed of subunits from an angiosperm DHS and its paralogous HSS for the first time. This proves that although the HSS (DHS3) already lost its ability to bind the eIF-5A precursor during evolution due to amino acid changes (see section 5.5), its subunits are still able to interact with those of an original DHS enzyme (DHS1 and DHS2), which is evidence for the influence of the postulated paralog interference on HSS evolution [12, 121]. In addition, the western blot results indicate that the tetramers are not static, but are in a dynamic assembling/disassembling equilibrium. Nevertheless, further experiments will have to be conducted to elucidate the influence of heteromer formation upon enzyme activity, to determine the exact heterotetramer composition and to gain further insight into the exchange kinetics of the subunits between separate paralogous tetramers and heterotetramers.

6.8. Conclusion and future prospects

The development of the specialized angiosperm HSS enzyme from a duplicate of the ubiquitous and highly conserved eukaryotic DHS enzyme is a good example for a rare evolutionary strategy, the so-called "evolution by change of function" [192]. In this thesis, biophysical differences and common features of two DHS homologs (DHS1, DHS2) and their HSS-like paralog (DHS3) from the model plant *Sorghum bicolor* were extensively characterized, and correlated to previous knowledge about DHS and HSS from other species.

All three enzymes – as well as the DHS substrate, the eIF-5A precursor – tolerated a broad range of pH values and salt concentrations around physiological conditions in thermal stability assays. Compared to the two DHS enzymes, however, the overall thermal stability of the HSS-like DHS3 was significantly lower, as determined by complementary DSF and DLS measurements. In further experiments, the DHS/HSS cofactor NAD^+ was identified as a key element in the prevention of unwanted aggregation or higher order oligomerization of the enzyme tetramers, which were observed in the course of this thesis during protein purification procedures and upon storage. Furthermore, NAD^+ notably increased the thermal stability of all three enzymes – especially at basic pH. BLI binding experiments revealed that the presence of millimolar concentrations of NAD^+ not only stabilized the DHS enzymes themselves, but also demonstrably enhanced the half-life of the DHS–eIF-5A interaction by drastically decreasing its dissociation rate. Although subsequent SEC-MALS experiments indicated that the supplementation of NAD^+ in the buffers did not suffice to isolate stable DHS–eIF-5A complexes, the stabilizing effect of NAD^+ on these complexes will still be of great value, especially for the future X-ray crystallographic determination of the complex structure.

The characterizing feature of a specialized angiosperm HSS in contrast to the paralogous DHS enzyme is the lack of DHS activity, frequently accompanied by increased homospermidine synthesis rates. BLI binding experiments confirmed that the missing DHS activity in the HSS-like DHS3 can indeed be attributed to the loss of binding between the enzyme and the eIF-5A precursor, as was previously already proposed for HSS from other species [192]. For the two homologous DHS, the kinetic constants of the interaction with two eIF-5A isoforms from

S. bicolor were determined at physiological pH as well as at pH 9.0, where the enzymatic reaction reaches its optimum. The K_D values for both enzymes ranged from 0.3 to 0.9 nM at pH 7.5, and from 3.7 to 11.3 nM at basic pH, which correlates with K_D values for human DHS obtained by equilibrium ultracentrifugation [151] and confirms the high affinity between these proteins. Unfortunately, although the underlying association and dissociation constants indicated differences in the interaction between the two DHS homologs and the two eIF-5A isoforms, the measurement conditions need to be optimized further to allow a more detailed data interpretation.

The deoxyhypusination reaction as well as the homospermidine synthesis require the triamine spermidine as donor for the transferred aminobutyl moiety. In this thesis, the spermidine analog and potent competitive DHS inhibitor GC7 was applied to substitute for the original substrate. In thermal stability assays, GC7 supported and even enhanced the stabilizing effect of NAD^+ on both DHS enzymes as well as on the HSS-like DHS3. However, in accordance with previous findings from human DHS, the inhibitor also depended on the presence and prior binding of the cofactor NAD^+ to the active sites of the enzymes [148, 149]. Additional BLI binding experiments revealed a rather astonishing characteristic of the inhibitory effect of GC7 – that has, to my knowledge, not been described until now. GC7 is commonly known as a competitive inhibitor that effectively displaces the spermidine molecule from its binding site due to its low nanomolar K_i , and that has already been confirmed in the spermidine binding site of the crystal structure of human DHS [114, 280]. The BLI binding experiments performed in the course of this thesis, however, clearly show that the inhibitor not only competes with spermidine for the active site, but also efficiently prevents the interaction of the DHS enzyme with the eIF-5A precursor, and even interrupts existing complexes. These results are especially valuable for future complex crystallization trials, since so far GC7 has always been added to the crystallization conditions for its stabilizing effect on DHS and HSS enzymes – a strategy that obviously had the contrary effect on the desired DHS-eIF-5A complexes.

During the development of a specialized angiosperm HSS from the duplicate of a DHS enzyme, the evolutionary fate of the HSS was proposed to stay linked to that of the DHS due to the fact that the formation of the active site – and thereby DHS as well as HSS activity – relies on the interaction of two individual subunits. It was hypothesized that the interface therefore is also conserved in the HSS enzyme, since any drastic changes would also impair HSS activity. This common interface would allow interaction between DHS and the paralogous HSS subunits, and transfer the selective pressure from the essential DHS enzyme to the HSS – a process called paralog interference [12, 121]. In this thesis, western blot experiments with *S. bicolor* DHS1 and HSS-like DHS3 confirmed the actual interaction of DHS and HSS subunits for the first time and provided direct evidence for the hypothesized paralog interference.

Taken together, the experiments performed in the course of this thesis allow a more detailed understanding of the DHS and HSS enzymes and provide a solid base for future experiments. A detailed three-dimensional structure of the DHS and HSS enzymes – and especially a complex of DHS and eIF-5A – would be the ultimate goal of all further efforts, since they can provide profound insights into the molecular mechanisms that underly the interaction between DHS and eIF-5A, the structural differences between DHS and HSS, and the inhibitory effect

of the spermidine analog GC7. Besides the optimization of the "classic" and *in vivo* crystallization experiments with the already described constructs, additional DHS variants should be generated with mutations that might stabilize the DHS–eIF-5A interaction (D243A, K329A; see sections 6.3.1 and 6.6). In addition to these X-ray crystallography-based approaches, other techniques could also be implemented to obtain more detailed information on structure and stoichiometry of the enzyme–precursor complexes, like native mass-spectrometry [103, 102], small-angle X-ray scattering (SAXS, [277]) and the increasingly powerful cryo-electron microscopy (CryoEM, [49]).

These experiments could be complemented by additional BLI binding experiments under optimized conditions and with further DHS and HSS variants. These should include HSS variants of the "telltale" amino acids (e. g. I274V/A, N278D), exchanges of specific amino acids in the vicinity of the enzymes' funnel area (R262N, F309L), as well as chimeric enzymes that introduce the DHS "ball-and-chain" motif into an HSS enzyme (see section 6.5.1). In addition, another eukaryotic, non-plant DHS could be examined in parallel. Here, the human DHS would be a reasonable choice, since already a lot of literature data is available for this enzyme. Consequently, this approach would also allow to relate experimental results from plant DHS to metazoan DHS via comparative analyses.

7. Material and methods

7.1. Bacterial strains and insect cell lines

Table 7.1.: Bacterial strains and insect cell lines

Strain/cell line	Genotype/features	Origin
Sf21	immortalized ovarian cell line from fall armyworm <i>Spodoptera frugiperda pupae</i> , susceptible to baculovirus infection	(1)
High Five TM	immortalized ovarian cell line from cabbage looper <i>Trichoplusia ni</i> , susceptible to baculovirus infection	(1)
XL1 Blue	F' <i>recA1 endA1 gyrA96 thi-1 hsdR17(r_K⁻m_K⁺) glnV44 relA1 lac</i> [F' <i>proA⁺B⁺ lacI^q Δ(lacZ)M15 Tn10(Tet^r)</i>]	(3)
BL21 (DE3)	<i>E. coli</i> B F ⁻ <i>ompT dcm gal hsdS_B(r_B⁻m_B⁻)</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>])	(3)
BL21 (DE3)-Gold	<i>E. coli</i> B F ⁻ <i>ompT dcm⁺ gal endA hsdS_B(r_B⁻m_B⁻) Tet^r</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) Hte	(3)
BL21-CodonPlus(DE3)-RIL	<i>E. coli</i> B F ⁻ <i>ompT dcm⁺ gal endA hsdS_B(r_B⁻m_B⁻) Tet^r</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) Hte [<i>argU ileY leuW Cam^r</i>]	(3)
BL21 (DE3)pLysS	<i>E. coli</i> B F ⁻ <i>ompT dcm gal hsdS_B(r_B⁻m_B⁻)</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) / pLysS (Cam ^r)	(3)
Rosetta TM 2(DE3)	<i>E. coli</i> B F ⁻ <i>ompT dcm gal hsdS_B(r_B⁻m_B⁻)</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) pRARE2(Cam ^r)	(3)
Rosetta TM 2(DE3)pLysS	<i>E. coli</i> B F ⁻ <i>ompT dcm gal hsdS_B(r_B⁻m_B⁻)</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) pLysSRARE2(Cam ^r)	(3)
Origami TM 2(DE3)	Δ(<i>ara-leu</i>)7697 Δ <i>lacX74</i> Δ <i>phoA</i> <i>PvuII phoR araD139 ahpC galE galK rpsL</i> [F' <i>lac⁺ lacI^q pro</i>] λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) <i>gor522::Tn10 trxB</i> (Str ^r , Tet ^r)	(3)
ArcticExpress TM (DE3)	<i>E. coli</i> B F ⁻ <i>ompT dcm⁺ gal endA hsdS_B(r_B⁻m_B⁻)</i> λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) Hte [<i>cpn10 cpn60 Gent^r</i>]	(3)
DH10Bac	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80Δ(<i>lacZ</i>)M15 <i>recA1 endA1</i> Δ(<i>ara-leu</i>)7697 Δ <i>lacX74 araD139 galU galK rpsL nupG</i> λ ⁻ / bMON14272 (Kan ^r) / pMON7124 (Tet ^r)	(3)
DH10EMBacY	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80Δ(<i>lacZ</i>)M15 <i>recA1 endA1</i> Δ(<i>ara-leu</i>)7697 Δ <i>lacX74 araD139 galU galK rpsL nupG</i> λ ⁻ / EM-BacY (Cam ^r , Kan ^r) / pMON7124 (Tet ^r)	(3)

(1) Eukaryotic Expression Facility at the European Molecular Biology Laboratory (EMBL) Grenoble.

(2) Research group Structural Infection Biology applying new Radiation Sources of Prof. Dr. Lars Redecke, Hamburg and Lübeck University.

(3) Research group Structural biology of Prof. Dr. Axel Scheidig, Kiel University.

7.2. BVDV NS2 expression constructs

Table 7.2.: NS2 constructs for expression in *E. coli*

No.	Vector	Features	Source
1	pET-His ₆ -SUMO-[1286]NS2*	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2; Kan ^r	(1)
2	pET-His ₆ -SUMO-[1351]NS2*	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2; Kan ^r	(1)
3	pET-His ₆ -SUMO-[1393]NS2*	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2; Kan ^r	(1)
4	pET-His ₆ -SUMO-[1286]NS2 _{Jiv90} *	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2 with NADL-like Jiv90 insert; Kan ^r	(1)
5	pET-His ₆ -SUMO-[1351]NS2 _{Jiv90} *	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2 with NADL-like Jiv90 insert; Kan ^r	(1)
6	pET-His ₆ -SUMO-[1393]NS2 _{Jiv90} *	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2 with NADL-like Jiv90 insert; Kan ^r	(1)
17	pET-His ₆ -SUMO-[1393]NS2 _{Jiv90}	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2 with NADL-like Jiv90 insert; Kan ^r	this thesis
18	pET-His ₆ -SUMO-[1393]NS2	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, NCP7 NS2; Kan ^r	this thesis
21	pCGFP-[1303]NS2 _{Jiv90}	T7 promoter/terminator; NCP7 NS2 with NADL-like Jiv90 insert, thrombin cleavage site, <i>C</i> -terminal GFP and His ₆ -tag; Amp ^r	this thesis
22	pCGFP-[1379]NS2 _{Jiv90}	T7 promoter/terminator; NCP7 NS2 with NADL-like Jiv90 insert, thrombin cleavage site, <i>C</i> -terminal GFP and His ₆ -tag; Amp ^r	this thesis
23	pNGFP-[1303]NS2 _{Jiv90}	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and GFP, thrombin cleavage site, NCP7 NS2 with NADL-like Jiv90 insert; Amp ^r	this thesis
24	pNGFP-[1379]NS2 _{Jiv90}	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and GFP, thrombin cleavage site, NCP7 NS2 with NADL-like Jiv90 insert; Amp ^r	this thesis
25	pCGFP-[1303]NS2	T7 promoter/terminator; NCP7 NS2, thrombin cleavage site, <i>C</i> -terminal GFP and His ₆ -tag; Amp ^r	this thesis
26	pCGFP-[1379]NS2	T7 promoter/terminator; NCP7 NS2, thrombin cleavage site, <i>C</i> -terminal GFP and His ₆ -tag; Amp ^r	this thesis
27	pNGFP-[1303]NS2	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and GFP, thrombin cleavage site, NCP7 NS2; Amp ^r	this thesis
28	pNGFP-[1379]NS2	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and GFP, thrombin cleavage site, NCP7 NS2; Amp ^r	this thesis
-	pET-His ₆ -SUMO-GFP	T7 promoter/terminator; <i>N</i> -terminal His ₆ -tag and SUMO, GFP; Kan ^r	this thesis

These constructs only contain the *C*-terminal part of NS2 or NS2_{Jiv90}. The number in square brackets gives the position of the *N*-terminal truncation (numbering based on the NCP7 polyprotein). In the different constructs the NS2 sequence was codon-optimized for the expression in *E. coli*. Constructs marked with asterisk contain the authentic genomic NS2 sequence. The boldface number corresponds to the serial number of the constructs used in the cloning strategies (see section 7.8). pET vector backbones are derived from pET28 (Novagen[®]), pNGFP and pCGFP derivatives are constructed from pNGFP-BC and pCGFP-BC [125].

(1) Research group of Prof. Dr. Norbert Tautz, Institute of Virology and Cell Biology, Lübeck University.

Table 7.3.: NS2 constructs for insect cell expression

No.	Vector	Features	Source
48	pFastBac1-cyt- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -mCherry	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, <i>C</i> -terminal mCherry; Amp ^r , Gent ^f	this thesis
49	pFastBac1-ER- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -mCherry	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal ER localization sequence, <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, <i>C</i> -terminal mCherry; Amp ^r , Gent ^f	this thesis
50	pFastBac1-ER- <i>StreptII</i> -NS2-NS3 _{7aa} -mCherry-KDEL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> - and <i>C</i> -terminal ER localization and retention sequences, <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , mCherry; Amp ^r , Gent ^f	this thesis
51	pFastBac1- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -mCherry-SKL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, mCherry, <i>C</i> -terminal peroxisomal localization sequence; Amp ^r , Gent ^f	this thesis
52	pFastBac1-cyt- <i>StreptII</i> -NS2-NS3 _{7aa} -mCherry	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , <i>C</i> -terminal mCherry; Amp ^r , Gent ^f	this thesis
53	pFastBac1-ER- <i>StreptII</i> -NS2-NS3 _{7aa} -mCherry	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal ER localization sequence, <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , <i>C</i> -terminal mCherry; Amp ^r , Gent ^f	this thesis
54	pFastBac1-ER- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -mCherry-KDEL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> - and <i>C</i> -terminal ER localization and retention sequences, <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, mCherry; Amp ^r , Gent ^f	this thesis
55	pFastBac1- <i>StreptII</i> -NS2-NS3 _{7aa} -mCherry-SKL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , mCherry, <i>C</i> -terminal peroxisomal localization sequence; Amp ^r , Gent ^f	this thesis

These constructs contain the full-length NS2 or NS2_{Jiv90} and the first 7 amino acids of NS3. The *C*-terminal portion of NS2 is codon-optimized for the expression in *E. coli*, while the *N*-terminal part corresponds to a humanized version of the coding sequence. The boldface number corresponds to the serial number of the constructs used in the cloning strategies (see section 7.8). The vector backbone is originally derived from pFastBacTM 1 (Thermo Scientific).

Table 7.4.: NS2 constructs for cell free expression in tobacco BY-2 cell extract

No.	Vector	Features	Source
56	pEUMod- <i>StreptII</i> -NS2-NS3 _{7aa} -EGFP	SP6 promoter/T7 terminator; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , <i>C</i> -terminal TEV protease cleavage site, EGFP, HA-tag; Amp ^r	this thesis
57	pEUMod- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -EGFP	SP6 promoter/T7 terminator; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, <i>C</i> -terminal TEV protease cleavage site, EGFP, HA-tag; Amp ^r	this thesis
58	pEUMod- <i>StreptII</i> -NS2-NS3 _{7aa} -EGFP-TMV3	SP6 promoter/T7 terminator; 3' TMV enhancer sequence; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , <i>C</i> -terminal TEV protease cleavage site, EGFP, HA-tag; Amp ^r	this thesis

(Continued on next page)

Table 7.4.: NS2 constructs for cell free expression in tobacco BY-2 cell extract (*continued*)

No.	Vector	Features	Source
59	pEUmod- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -EGFP-TMV3	SP6 promoter/T7 terminator; 3' TMV enhancer sequence; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, <i>C</i> -terminal TEV protease cleavage site, EGFP, HA-tag; Amp ^r	this thesis
62	pUC57-BY2- <i>StreptII</i> -NS2-NS3 _{7aa} -EGFP	T7 promoter/T7 terminator; 5' (Ω) and 3' TMV enhancer sequences; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} , <i>C</i> -terminal TEV protease cleavage site, EGFP, HA-tag; Kan ^r	this thesis
63	pUC57-BY2- <i>StreptII</i> -NS2 _{Jiv90} -NS3 _{7aa} -EGFP	T7 promoter/T7 terminator; 5' (Ω) and 3' TMV enhancer sequences; <i>N</i> -terminal <i>StreptII</i> -tag, NCP7 NS2-NS3 _{7aa} with NADL-like Jiv90 insert, <i>C</i> -terminal TEV protease cleavage site, EGFP, HA-tag; Kan ^r	this thesis

These constructs contain the full-length NS2 or NS2_{Jiv90} and the first 7 amino acids of NS3. The *C*-terminal portion of NS2 is codon-optimized for the expression in *E. coli*, while the *N*-terminal part corresponds to a humanized version of the coding sequence. The boldface number corresponds to the serial number of the constructs used in the cloning strategies (see section 7.8). pEU vectors are based on pEU-E01-MCS from CellFree Sciences Co., Ltd. (Yokohama, Japan). The pUC57 constructs have the vector backbone of pUC57 (Thermo Scientific).

7.3. *Sorghum bicolor* eIF-5A and DHS expression constructs

Table 7.5.: eIF-5A constructs for expression in *E. coli*

No.	Vector	Features	Source
64	pET-ZmelF5A-His ₆	T7 promoter/terminator; full-length <i>Zea mays</i> eIF-5A, <i>C</i> -terminal His ₆ -tag; Amp ^r	(¹)
65	pET-SbelF5A1-His ₆	T7 promoter/terminator; full-length <i>S. bicolor</i> eIF-5A isoform 1, <i>C</i> -terminal His ₆ -tag; Amp ^r	this thesis
66	pET-SbelF5A2-His ₆	T7 promoter/terminator; full-length <i>S. bicolor</i> eIF-5A isoform 2, <i>C</i> -terminal His ₆ -tag; Amp ^r	this thesis
67	pET-SbelF5A1-His ₆ -SA <i>StreptII</i>	T7 promoter/terminator; full-length <i>S. bicolor</i> eIF-5A isoform 1, His ₆ -tag, <i>C</i> -terminal Ser-Ala- <i>StreptII</i> -tag; Amp ^r	this thesis
68	pET-SbelF5A2-His ₆ -SA <i>StreptII</i>	T7 promoter/terminator; full-length <i>S. bicolor</i> eIF-5A isoform 2, His ₆ -tag, <i>C</i> -terminal Ser-Ala- <i>StreptII</i> -tag; Amp ^r	this thesis

The boldface number corresponds to the serial number of the constructs used in the cloning strategies (see section 7.9). pET vector backbones are derived from pET23 (Novagen[®]).

(¹) Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober at the Botanical Institute, Kiel University.

Table 7.6.: DHS constructs for expression in *E. coli*

No.	Vector	Features	Source
73	pET-SbDHS3-His ₆	T7 promoter/terminator; full-length <i>S. bicolor</i> DHS isoform 3, <i>C</i> -terminal His ₆ -tag; Amp ^r	(¹)
74	pET-SbDHS2-His ₆	T7 promoter/terminator; full-length <i>S. bicolor</i> DHS isoform 2, <i>C</i> -terminal His ₆ -tag; Amp ^r	(¹)

(Continued on next page)

Table 7.6.: DHS constructs for expression in *E. coli* (continued)

No.	Vector	Features	Source
75	pET-SbDHS1-His ₆	T7 promoter/terminator; full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag; Amp ^r	(1)
77	pET- <i>StreptII</i> -SbDHS3-His ₆	T7 promoter/terminator; N-terminal <i>StreptII</i> -tag, full-length <i>S. bicolor</i> DHS isoform 3, C-terminal His ₆ -tag; Amp ^r	this thesis
78	pET- <i>StreptII</i> -SbDHS2-His ₆	T7 promoter/terminator; N-terminal <i>StreptII</i> -tag, full-length <i>S. bicolor</i> DHS isoform 2, C-terminal His ₆ -tag; Amp ^r	this thesis
79	pET- <i>StreptII</i> -SbDHS1-His ₆	T7 promoter/terminator; N-terminal <i>StreptII</i> -tag, full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag; Amp ^r	this thesis
80	pET-SAS <i>StreptII</i> -SbDHS3-His ₆	T7 promoter/terminator; N-terminal Ser-Ala- <i>StreptII</i> -tag, full-length <i>S. bicolor</i> DHS isoform 3, C-terminal His ₆ -tag; Amp ^r	this thesis
81	pET-SAS <i>StreptII</i> -SbDHS2-His ₆	T7 promoter/terminator; N-terminal Ser-Ala- <i>StreptII</i> -tag, full-length <i>S. bicolor</i> DHS isoform 2, C-terminal His ₆ -tag; Amp ^r	this thesis
82	pET-SAS <i>StreptII</i> -SbDHS1-His ₆	T7 promoter/terminator; N-terminal Ser-Ala- <i>StreptII</i> -tag, full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag; Amp ^r	this thesis
85	pCDFDuet-SbDHS3-His ₆	T7 promoter/terminator; MCS1: full-length <i>S. bicolor</i> DHS isoform 3, C-terminal His ₆ -tag, MCS2: empty; Sm ^r	this thesis
86	pCDFDuet-SbDHS2-His ₆	T7 promoter/terminator; MCS1: full-length <i>S. bicolor</i> DHS isoform 2, C-terminal His ₆ -tag, MCS2: empty; Sm ^r	this thesis
87	pCDFDuet-SbDHS1-His ₆	T7 promoter/terminator; MCS1: full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag, MCS2: empty; Sm ^r	this thesis
88	pCDFDuet-SbDHS3- <i>StreptII</i>	T7 promoter/terminator; MCS1: empty, MCS2: full-length <i>S. bicolor</i> DHS isoform 3, C-terminal <i>StreptII</i> -tag; Sm ^r	this thesis
89	pCDFDuet-SbDHS2- <i>StreptII</i>	T7 promoter/terminator; MCS1: empty, MCS2: full-length <i>S. bicolor</i> DHS isoform 2, C-terminal <i>StreptII</i> -tag; Sm ^r	this thesis
90	pCDFDuet-SbDHS1- <i>StreptII</i>	T7 promoter/terminator; MCS1: empty, MCS2: full-length <i>S. bicolor</i> DHS isoform 1, C-terminal <i>StreptII</i> -tag; Sm ^r	this thesis
91	pCDFDuet-SbDHS1-His ₆ -SbDHS3- <i>StreptII</i>	T7 promoter/terminator; MCS1: full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag, MCS2: full-length <i>S. bicolor</i> DHS isoform 3, C-terminal <i>StreptII</i> -tag; Sm ^r	this thesis
92	pCDFDuet-SbDHS1-His ₆ -SbDHS2- <i>StreptII</i>	T7 promoter/terminator; MCS1: full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag, MCS2: full-length <i>S. bicolor</i> DHS isoform 2, C-terminal <i>StreptII</i> -tag; Sm ^r	this thesis
93	pCDFDuet-SbDHS1-His ₆ -SbDHS1- <i>StreptII</i>	T7 promoter/terminator; MCS1: full-length <i>S. bicolor</i> DHS isoform 1, C-terminal His ₆ -tag, MCS2: full-length <i>S. bicolor</i> DHS isoform 1, C-terminal <i>StreptII</i> -tag; Sm ^r	this thesis

The boldface number corresponds to the serial number of the constructs used in the cloning strategies (see section 7.9). pET vector backbones are derived from pET23 (Novagen®), while pCDFDuet constructs are based on pCDFDuetTM 1 (Novagen®).

(1) Research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober at the Botanical Institute, Kiel University.

Table 7.7.: DHS constructs for expression in insect cells

No.	Vector	Features	Source
94	pFastBac1-ER-SbDHS3	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal ER localization sequence, full-length <i>S. bicolor</i> DHS isoform 3; Amp ^r , Gent ^r	this thesis
95	pFastBac1-ER-SbDHS2	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal ER localization sequence, full-length <i>S. bicolor</i> DHS isoform 2; Amp ^r , Gent ^r	this thesis
96	pFastBac1-ER-SbDHS1	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> -terminal ER localization sequence, full-length <i>S. bicolor</i> DHS isoform 1; Amp ^r , Gent ^r	this thesis
97	pFastBac1-ER-SbDHS1-KDEL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> - and <i>C</i> -terminal ER localization and retention sequences, full-length <i>S. bicolor</i> DHS isoform 1; Amp ^r , Gent ^r	this thesis
98	pFastBac1-ER-SbDHS3-KDEL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> - and <i>C</i> -terminal ER localization and retention sequences, full-length <i>S. bicolor</i> DHS isoform 3; Amp ^r , Gent ^r	this thesis
99	pFastBac1-ER-SbDHS2-KDEL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; <i>N</i> - and <i>C</i> -terminal ER localization and retention sequences, full-length <i>S. bicolor</i> DHS isoform 2; Amp ^r , Gent ^r	this thesis
100	pFastBac1-cyt-SbDHS3	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; full-length <i>S. bicolor</i> DHS isoform 3; Amp ^r , Gent ^r	this thesis
101	pFastBac1-cyt-SbDHS2	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; full-length <i>S. bicolor</i> DHS isoform 2; Amp ^r , Gent ^r	this thesis
102	pFastBac1-cyt-SbDHS1	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; full-length <i>S. bicolor</i> DHS isoform 1; Amp ^r , Gent ^r	this thesis
103	pFastBac1-SbDHS3-SKL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; full-length <i>S. bicolor</i> DHS isoform 3, <i>C</i> -terminal peroxisomal localization sequence; Amp ^r , Gent ^r	this thesis
104	pFastBac1-SbDHS1-SKL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; full-length <i>S. bicolor</i> DHS isoform 1, <i>C</i> -terminal peroxisomal localization sequence; Amp ^r , Gent ^r	this thesis
105	pFastBac1-SbDHS2-SKL	polyhedrin promoter/SV40 late polyA terminator; Tn7R and Tn7L sequences; full-length <i>S. bicolor</i> DHS isoform 2, <i>C</i> -terminal peroxisomal localization sequence; Amp ^r , Gent ^r	this thesis

The boldface number corresponds to the serial number of the constructs used in the cloning strategies (see section 7.9). The vector backbone is derived from pFastBacTM 1 (Thermo Scientific).

7.4. Oligonucleotides

Oligonucleotides for use as PCR primers were obtained as salt-free or HPLC-purified lyophilisates from Eurofins Genomics (Ebersberg, Germany).

Table 7.8.: List of oligonucleotides used for cloning purposes

Primer name	Sequence (5'–3')
CH009for	ccgttcctcaagaataaacg
CH010rev	gccttcgaagttgccttcac
CH023for	tgcaagtacgtcccaggcgccgaaccaggcgcccagg
CH024rev	atgtctgctgcagttcatgccatggccttaggatccagcccag
CH027for	gaccctggtgctgatcctgccacgtatgaactggcgaaact
CH028rev	cattattgctctgcagttcatgccatggacgaggatccagcccag
CH029fw	cagtttgaaaaactggttccgcgtggtagcatggccggcgccgggag
CH031rv	cgcggaaccagttttcaaacctgcggatggctccacatagtatatctccttcttaaag
CH032for	cgatgaattcatggaaccaggcgcccaggggctac
CH035rev	gcataagctttacttctcgaactgagggtgggaccaaccaccaccacgcaggatccagcccagatg
CH036fw	cagtttgaaaaactggttccgcgtggtagcatgggtgccgcgcccgg
CH037fw	cagtttgaaaaactggttccgcgtggtagcatggccggcgccgggag
CH038for	ggtgcccgcgcccgggagaagg
CH039rev	gtttgctgatttagtccatgaacc
cyt-DHS3-for	tcgggcgaggatccatgggcatgggtgccgcgcccggg
cyt-DHS3-rev	ttctcgacaagcttttaggcgtttgctgatttagtccatgaacc
cyt-DHS2-for	tcgggcgaggatccatgggcatggccggcgccgggagg
cyt-DHS2-rev	ttctcgacaagcttttaggccttggagccatgaaccttgcg
cyt-DHS1-rev	ttctcgacaagcttttagccttggagctgtgaaccttacgtg
cyt- <i>StrepII</i> -NS2-rev	agggtgggaccagcccagatccgcgcccagatggtg
ER-DHS3-for	ctgtagtcgccgtaaacgcggcatgggtgccgcgcccgggag
ER-DHS3-rev	ctagtactctcgacaagcttttaggcgtttgctgatttagtccatg
ER-DHS2-for	ctgtagtcgccgtaaacgcggcatggccggcgccgggaggc
ER-DHS2-rev	gcctaaaagcttctcgagaagtactagcacgtaaggttcacagctccaaag
ER-DHS1-rev	ctagtactctcgacaagcttttagccttggagctgtgaaccttacgtg
ER- <i>StrepII</i> -NS2-rev	agggtgggaccagcccgcgtttacggcgactac
KDEL-DHS3-rev	cgacaagcttttacagttcatctttggcgtttgctgatttagtccatg
KDEL-DHS2-rev	cgacaagcttttacagttcatctttggccttggagccatgaaccttgcg
KDEL-DHS1-rev	cgacaagcttttacagttcatctttggccttggagctgtgaaccttacg
mCherry-fw	gcttatgacgtcgtgagcaaggcgcaagaag
mCherry-rev	tcttgagagctcctgtacagctcgtccatg
NS2for	gaaccagggtgccagggtctacctggaac
NS2rev	acgcaggatccagcccagatg
NS2-rev	ttcttcacacggcaggtccacgcaggatccagcccagatg
NS3-fw	agacgtcaagatagagctcgctaaagcttgcgagaagtactagag
NS3-KDEL-fw	agacgtcaagatagagctcgcaaaagatgaactgtaaaagcttgcgag
NS3-SKL-fw	agacgtcaagatagagctcgcaaaagctgtaaaagcttgcg
pCDFDuet1- <i>StrepII</i> -for	ccatccgcagtttgaaaaataaataacctaggctgctgccac
pCDFDuet1- <i>StrepII</i> -rev	ctccagtcgacggtagcagctcagcgtac
pEU- <i>Strep</i> -NS2-fw	catacgtcatgagcgcaggtcccaccctcagttcgagaag
pEU-NS2-NS3-rw	gatgtcctcgagtttcttcacacggcaggtccacgcaggatccagcccagatg
pEU-GFP-Sall-rv	cagttgattgggtcgacttaagcgtaactctggaacatcgtatgg
pUC57-BY2-del-for	atgagcgcaggtcccac
pUC57-BY2-del-rev	ggtagttgtagaatgtaaaatgtaag
RF-DHS3for	gtttaactttaataaggagatataccatgggtgccgcgcccgggag
RF-DHS3rev	cttaagcattatgcgcccgaagctttagtgggtgggtgggtggtgc
RF2-DHS2-for	gttaagtataagaaggagatatacatatggccggcgccgggag
RF2-DHS2-rev	caaacctggaggtcctcagctcagcttggagccatgaaccttgc
RF-DHS1-rev	caaacctggaggtcctcagctcagcttggagctgtgaaccttacg
<i>Sa</i> I/mt-for	cgagctcggcgccgctcagcagacagaagcttgcggccgataatgc

(Continued on next page)

Table 7.8.: List of oligonucleotides used for cloning purposes (*continued*)

Primer name	Sequence (5'–3')
<i>Sa</i> I/mut-rev	cattatgcgccgcaagcttgtctgcctgcaggcgcgcccagc
<i>SAStrepII</i> -for	gagatatacatatgtcagcatggagccatccgcagtttg
<i>SAStrepII</i> -rev	ctgcggatggctccatgctgacatgtatatctccttctaaag
<i>Sb</i> DHS3- <i>Nde</i> I-for	gaaggagatatacatatgggtg
<i>Sb</i> DHS3- <i>Xho</i> I-rev	ggtggtgctcgaggtttgc
<i>Sbe</i> IF5A-for	ggaattccatattgtcggactccgagg
<i>Sbe</i> IF5A1-rev	ccagctcgagcttggggccgatatccttcag
<i>Sbe</i> IF5A2-rev	ccagctcgaggttcttgcccccaatctcc
<i>Sbe</i> IF5A1- <i>Strep</i> F	agcgcattggagccatccgcagtttgaaaaatgagatccggctgctaac
<i>Sbe</i> IF5A1- <i>Strep</i> R	ttttcaaacctgcggatggctccatgcgctgtggtggtggtggtg
<i>Sbe</i> IF5A1-17 <i>N</i> -fw	tccaagacttatccgcagcag
<i>Sbe</i> IF5A2-17 <i>N</i> -fw	tccaagacgtaccgcagc
<i>Sbe</i> IF5A-17 <i>N</i> -rv	accatattgtatatctccttcttaaag
SKL-DHS2-for	gcgcgcggaattcatcatgggcatggccggcgggaggc
SKL-DHS2-rev	cgacaagctttacagtttgctggccttggagccatgaaaccttgctg
SKL-DHS1-rev	cgacaagctttacagtttgctggccttggagctgtgaaaccttacgtg
<i>StrepII</i> -NS2-fw	cagttcgagaagggtgaaccaggtgccagggtctac

7.5. Chemicals

Chemicals used throughout this thesis were purchased from Carl Roth (Karlsruhe, Germany) and Merck (Darmstadt, Germany; formerly known as Sigma-Aldrich[®]), unless stated otherwise.

7.6. Molecular biology

7.6.1. *In silico* design of constructs and evaluation of cloning strategies

To facilitate construct design and cloning strategy evaluation, the freely available software Serial Cloner (http://serialbasics.free.fr/Serial_Cloner/) and Unipro UGENE (<http://ugene.net/>, [193]) were applied in the current version.

7.6.2. Codon usage analysis

The codon usage of constructs applied in this thesis was analyzed with the web-based tool Graphical Codon Usage Analyser (GCUA, <http://gcu.schoedl.de/>) in the "each triplet position vs. usage table" mode in comparison to the codon usage of the desired expression organism. The codon usage tables used in the analyses were derived from the Kazusa web server (<http://www.kazusa.or.jp/codon/>) [181, 84].

7.6.3. Sequencing

The integrity of all created expression constructs was confirmed by sequencing of the expression cassettes with appropriate primers. The sequencing service was provided by Eurofins Genomics (Ebersberg, Germany) [241].

7.6.4. Oligonucleotide synthesis

Oligonucleotides, which were employed as primers for PCRs throughout this thesis, were all purchased as salt-free or HPLC-purified lyophilisates from Eurofins Genomics.

7.6.5. Analytical and preparative agarose gel electrophoresis

DNA samples for analytical or preparative purposes were separated on a self-cast agarose gel in a submerged horizontal electrophoresis cell (Mini-Sub[®] Cell GT Cell or Wide Mini-Sub[®] Cell GT Cell, Bio-Rad) at constant voltage, usually between 40 and 100 V. The gel was prepared from 1 to 1.5% agarose in **TAE buffer** (40 mM Tris, 0.114% (v/v) acetic acid, 1 mM EDTA (added as solution with pH 8.0)), depending on the fragment size range to be separated. The run time was chosen according to voltage, gel length and percentage, fragment size and desired separation efficiency.

7.6.6. Purification of DNA fragments

DNA fragments from PCRs (7.6.13) or restriction enzyme digests (7.6.11) were usually either purified directly from the reaction mixture with the PCR Clean-up protocol of the NucleoSpin[®] Gel and PCR Clean-up Kit (Macherey-Nagel), or with the Gel Clean-up protocol of the same kit, following a preparative agarose gel electrophoresis (7.6.5) and extraction of the desired bands [283].

7.6.7. Plasmid DNA preparation

6 mL LB (Luria/Miller; 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.0) in glass tubes containing the appropriate antibiotics were inoculated with a single colony or from a glycerol stock and grown at 37 °C and 230 rpm in an orbital shaker (Infors Multitron) overnight. Vector DNA was prepared from 3 mL densely grown culture with the NucleoSpin[®] Spin Plasmid Kit (Macherey-Nagel) or the Presto Mini Plasmid Kit (Geneaid) according to manufacturer's recommendations [32].

7.6.8. DNA and protein concentration determination

DNA concentrations were measured with a NanoPhotometer[®] (Implen) at 260 nm. Protein concentration was determined by 280 nm absorption measurements with the NanoPhotometer[®] using protein-specific absorption constants calculated by the web-based tool ProtParam (<http://web.expasy.org/protparam/>) based on the provided amino acid sequence.

In addition, protein concentrations were in some cases verified by Bradford protein assay [35]. The measurements were performed with the Protein Assay Dye Reagent Concentrate (Bio-Rad) according to manufacturer's recommendations. The Quick Start Bovine Serum Albumin Standard Set (Bio-Rad), which contains seven protein concentration standards ranging from 0.125 to 2.0 mg, was used to generate a standard curve for the evaluation of measurement results. Absorption of the samples was measured at 595 nm in the NanoPhotometer[®].

7.6.9. Phosphorylation of oligonucleotides

Oligonucleotides for deletion and insertion PCR based on the site-directed mutagenesis approach (7.6.14) were phosphorylated for subsequent ligation with T4 polynucleotide kinase (Thermo Scientific) according to manufacturer's recommendations.

7.6.10. Dephosphorylation of DNA fragments

DNA fragments were dephosphorylated in the course of a cloning strategy to prevent unwanted vector re-ligation. The dephosphorylation reaction was performed with either Thermosensitive Alkaline Phosphatase (TSAP, Promega) or FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) according to manufacturer's recommendations [251, 279].

7.6.11. Restriction enzyme digestion of DNA

Site specific digestion of DNA for cloning purposes was performed with restriction enzymes purchased from New England Biolabs, Promega or Thermo Scientific. The enzymes used are specified in the respective cloning strategy. Digestion conditions for single or double digests were adjusted according to manufacturer's recommendations.

7.6.12. DNA ligation

DNA ligation was performed with T4 ligase (Promega) according to manufacturer's recommendations. Insert-to-vector ratios were varied from 1:1 to 5:1, if necessary. Incubation

time and temperature were adjusted according to the type of ligation (sticky- or blunt-end, self-ligation).

7.6.13. Polymerase Chain Reaction (PCR)

PCRs for the different cloning strategies were performed based on with proofreading Phusion[®] High-Fidelity DNA Polymerase (Thermo Scientific) following manufacturer's recommendations and with the primers indicated in the respective cloning strategy [178]. Colony PCRs were performed with the 2X PCR Master Mix (Promega) according to manufacturer's recommendations. Annealing temperatures and elongation times were adjusted according to primer and template specifications. If necessary, a temperature gradient was chosen to determine the best annealing temperature empirically. For a standard or colony PCR, 30-35 PCR cycles were chosen, while for site-directed mutagenesis or restriction-free cloning approaches 18 cycles were sufficient.

7.6.14. RF-Cloning and site-directed mutagenesis

The restriction-free cloning strategies (RF-Cloning, [33]) used throughout this thesis employed two PCR rounds. In a first PCR (7.6.13), megaprimers for the desired insert were generated from a corresponding template and purified by agarose gel extraction (7.6.6). These megaprimers encompassed the complete insert and had 3' and 5' appendices derived from the original primers, which were complementary to the insert region on the final target plasmid. In a second PCR, the complete target plasmid was amplified with the previously generated megaprimers.

Two variants of a site-directed mutagenesis-based strategy (QuikChange[®], Stratagene; [197]) were used to insert or delete specific DNA sequences from plasmid DNA. For the first variant, in most parts complementary primers were designed, containing the new insert sequence or enclosing the desired deletion. A PCR with these primers was done to amplify the whole target vector. For the second site-directed mutagenesis variant, adjacent primers – inversely flanking the sequence to be deleted or containing the desired insertion in one of them – were phosphorylated (7.6.9) prior to a PCR amplifying the complete target vector. Primer phosphorylation enabled subsequent self-ligation (7.6.12) of the newly synthesized blunt-ended DNA.

The resulting PCR products were treated with *DpnI* to selectively remove methylated parental vector. In the case of the second site-directed mutagenesis variant, the PCR product was ligated prior to transformation into XL1-Blue or BL21-Gold(DE3) (7.6.15). In the other cases, the large complementary insert region enabled annealing and repair of single-strand breaks in *E. coli* without external ligation. Positive clones were identified by colony PCR (7.6.13) and a following analytical agarose gel electrophoresis (7.6.5), if possible, and the construct sequence verified by sequencing (7.6.3)

7.6.15. Transformation of competent *E. coli* with plasmid DNA

All transformations carried out as part of the different cloning strategies or as re-transformation of already existing plasmids were based on heat shock transformation of home-made CaCl_2 competent *E. coli* [26]. In the standard protocol, either *E. coli* XL1-Blue or BL21-Gold(DE3) were used for DNA preparation and cloning purposes, while for protein expression *E. coli* BL21 (DE3) or one of its derivatives was applied. *E. coli* DH10BacTM or DH10EMBacY were transformed to transpose a desired expression construct from its plasmid source into a bacmid suitable for insect cell expression.

For a heat shock transformation, 100 μl frozen home-made CaCl_2 competent *E. coli* were thawed on ice for 10 min and subsequently mixed with 1 to 50 ng purified plasmid DNA or approx. 2 μl of inactivated ligation reaction. After a 20 min incubation on ice, the mixture was heat shocked at 42 °C for 45 s. Immediately, 900 μl of **LB** (Luria/Miller; 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, pH 7.0) or **SOC medium** (2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 0.05 % (w/v) NaCl, 2.5 mM KCl, 10 mM MgCl_2 , 20 mM glucose, pH 7.0) were added and the mixture was incubated at 37 °C for 1 h (standard protocol) or overnight (bacmid transposition protocol) while agitated in a ThermoMixer[®] (Eppendorf) at 800 rpm. Afterwards, the mixture was centrifuged at 3000 rcf for 3 min, the pellet resuspended in 100 μl supernatant and streaked in different dilutions onto **LB agar** plates (Luria/Miller; 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, 1.5 % (w/v) agar, pH 7.0) containing appropriate antibiotics. The agar plates were incubated at 37 °C overnight.

7.7. Protein-related methods

7.7.1. SDS-PAGE

Standard reducing SDS-PAGE was performed using 0.75 mm thick self-cast discontinuous mini gels with 15 wells in the Mini-PROTEAN[®] Tetra Vertical Electrophoresis Cell (Bio-Rad) [228]. The percentage of acrylamide/bisacrylamide in the **separating gel** (400 mM Tris-HCl, 0.1 % (w/v) SDS, 0.1 % (w/v) APS, 0.08 % (v/v) TEMED, pH 8.8 at 22 °C) was chosen to suit the molecular weight range of the samples, while the **stacking gel** (130 mM Tris-HCl, 0.05 % (w/v) SDS, 0.2 % (w/v) APS, 0.2 % (v/v) TEMED, pH 6.8 at 22 °C) always had 5 % acrylamide/bisacrylamide. Sample volume ranged from 1 to 10 μL .

12 % SDS-PAGEs as control of the purification progress during protein purification were run for 35 min at 250 V. Gels for western blot were run for 30 min at 60 V until the sample reached the upper edge of the separating gel and then at 140 V until the running front reached the lower border of the gel.

The samples for a reducing SDS-PAGE were prepared in **Laemmli buffer** (62.5 mM Tris-HCl, 2 % (w/v) SDS, 10 % (v/v) glycerol, 100 mM DTT, 0.01 % (w/v) bromophenol blue, pH 6.8 at 22 °C [141]) or a **variation** of that buffer (100 mM disodium phosphate, 2 % (w/v) SDS, 10 % (v/v) glycerol, 5 % (v/v) β -mercaptoethanol, 0.01 % (w/v) bromophenol blue, pH 7.5) and heated for 5 min at 95 °C. The samples for in-gel fluorescence were prepared in the Laemmli buffer variant and heated for 30 min at 37 °C to preserve the integrity and thereby fluorescence of the fluorescing fusion proteins.

7.7.2. SDS-PAGE staining

SDS-PAGEs were stained with **Coomassie Brilliant Blue G-250** (0.1 % (w/v) CBB G-250, 30 % (v/v) ethanol, 10 % (v/v) acetic acid) for several hours and destained with 30 % (v/v) ethanol, 10 % (v/v) acetic acid, or stained with **Colloidal Coomassie Brilliant Blue G-250** (0.02 % (w/v) CBB G-250, 5 % (w/v) aluminum sulfate-(14-18)-hydrate, 10 % (v/v) ethanol, 2 % (v/v) orthophosphoric acid) overnight and destained with distilled water, if necessary [77].

7.7.3. Western blot of protein samples from SDS-PAGE

Western blots were done as tank blots in a Mini Trans-Blot[®] Electrophoretic Transfer Cell (Bio-Rad) at constant 300 mA for 50 min in **transfer buffer** (25 mM Tris, 192 mM glycine, 20 % (v/v) methanol, pH 8.3). The proteins were blotted onto methanol-activated PVDF membrane (Roth). After transfer, the membrane was optionally activated again with methanol and stained for whole protein with **Ponceau S** (0.1 % (w/v) Ponceau S, 5 % (v/v) acetic acid) for several minutes and destained in distilled water. The destained blot was then probed for the desired target with the respective agents [231, 43].

7.7.4. Chromogenic and fluorescence detection of His₆- and StrepII-tagged proteins on western blot

For the detection of His₆-tagged proteins on western blot, the blot was washed 2 times for 10 min with 20 mL **TBS** (10 mM Tris-HCl, 150 mM NaCl, pH 7.5 at 22 °C) and blocked for 60 min, slowly shaking in **blocking solution** (3 % (w/v) bovine serum albumin (BSA) in TBS). Afterwards, the blot was washed again twice for 10 min in 20 mL **TBSTT** (20 mM Tris-HCl, 500 mM NaCl, 0.2 % (v/v) Triton[™] X-100, 0.05 % (v/v) Tween[®] 20, pH 7.5 at 22 °C) and once in TBS. Subsequently, the blot was incubated overnight at 4 °C with a 1:1000 dilution of the primary Mouse Anti-His-Tag Monoclonal Antibody (Novagene) in blocking solution. On the next day, the blot was washed twice for 10 min in 20 mL TBSTT and once in TBS. The secondary antibody, the Goat Anti-Mouse IgG alkaline phosphatase (AP) Conjugate (Novagene), was applied to the blot in a 1:5000 dilution in blocking solution. The blot was subsequently incubated for 60 min on a rotator (neoLab) at room temperature and afterwards washed 5 times for 10 min with 20 mL TBSTT.

For the chromogenic detection of AP-conjugated Goat Anti-Mouse IgG, 10 µL BCIP (42 % (w/v) in DMF) and 10 µL NBT (83 % (w/v), 70 % DMF) were added to 2.5 mL of AP-Buffer (Novagen) and applied onto the blot. The reaction was stopped in distilled water when the desired signal intensity was reached.

For the detection of StrepII-tagged proteins on western blot, the blot was washed 2 times for 10 min with 20 mL TBSTT and blocked overnight at 4 °C, slowly shaking in SuperBlock[™] (PBS, Pierce). Afterwards, the membrane was washed 3 times for 5 min with 20 mL TBSTT, transferred to a 50 mL tube with 5 mL TBSTT and 5 µL Biotin Blocking Buffer (iba) and incubated for 10 min on a rotator (neoLab) at room temperature.

For the chromogenic detection with AP, 1.25 µL of AP-conjugated Strep-Tactin[®] (iba) were added and the blot was incubated for another 60 min on the rotator. The blot was

then washed 2 times for 1 min with 20 mL TBSTT and subsequently 2 times for 1 min with 20 mL TBS. The chromogenic detection of AP-conjugated *Strep*-Tactin[®] was performed as described for AP-conjugated Goat Anti-Mouse IgG (see above).

For the antibody-based fluorescence detection of *Strep*II-tagged proteins, 1.5 µL of 1:10 diluted horseradish peroxidase (HRP)-conjugated *Strep*MAB-Classic primary antibody (iba) were added to the biotin blocking solution and the blot incubated for another 60 min on the rotator. Afterwards, the blot was washed 2 times for 5 min with TBSTT. 12.5 µL Alexa Fluor[®] 532-conjugated Goat anti-Mouse IgG secondary antibody (Thermo Scientific) were added in 5 mL TBSTT to the blot. Subsequently, the blot was incubated in the dark for 90 min on the rotator. Then, the blot was washed twice for 1 min in TBSTT and afterwards twice in TBS. The fluorescence of Alexa Fluor[®] 532 was detected on a Blue/Green LED Transilluminator (Nippon Genetics) equipped with an Amber filter.

7.8. Cloning of BVDV NS2 constructs

In the following cloning strategies, already existing and newly generated vectors are numbered in the order of their appearance and afterwards referred to by these numbers to avoid too many long construct names throughout the text.

Vectors pET-His₆-SUMO-[1286]NS2* (1), pET-His₆-SUMO-[1351]NS2* (2), pET-His₆-SUMO-[1393]NS2* (3) and the respective constructs with a Jiv90 insertion at glycine 1536 (4, 5 and 6) and pCITE-BVDV-CP7-E2-p7-hNS2-NS3-NS4a (7) were kindly provided by the work group of Prof. Dr. Norbert Tautz at the Institute of Virology and Cell Biology of Lübeck University. They are based on a pET28 or pCITE[®]-2 vector (Novagen[®]), respectively.

Vectors pFastBac1-cyt (8), pFastBac1-ER (9), pFastBac1-ER-KDEL (10) and pFastBac1-SKL (11) for the baculovirus-mediated expression in insect cells were a kind gift of the work group Structural Infection Biology applying new Radiation Sources of Prof. Dr. Lars Redecke from Hamburg and Lübeck University. The vectors are constructed from pFast BacTM 1 (Thermo Scientific).

pEUMod-His₆-hTREK2b-EGFP (12) as basis for constructs for expression in the cell-free BY2 tobacco cell system was kindly provided by Marianne Musinszki of the work group Ion Channels of Prof. Dr. Thomas Baukrowitz at the Institute of Physiology at Kiel University. The vector is based on the pEU-E01-MCS vector of CellFree Sciences Co., Ltd. (Yokohama, Japan), which was developed for the wheat germ cell-free protein synthesis system.

pUC57_Kan_BY2_CFS1 (13) and pEUMod-His-GFP-b5-His-TMV3 (14), two other basic vectors for expression in the cell-free BY2 tobacco cell system, were designed by Dr. Sebastian Krossa of the work group Structural Biology of Prof. Dr. Axel Scheidig from Kiel University.

7.8.1. Codon-optimized [1303]NS2_{Jiv90} and [1303]NS2 for expression in *E. coli*

For the expression of codon-optimized NS2 constructs in *E. coli*, a construct encompassing the C-terminal part of NCP7 NS2 (beginning at threonine 1303) with the NADL-like Jiv90 insertion after glycine 1536 was designed. The construct was ordered from GenScript, where it was codon-optimized and in addition provided with suitable restriction enzyme recognition sites.

It was synthesized *de novo* and delivered in a pUC57 vector (termed pUC57-[1303]NS2_{Jiv90}, **15**). pUC57-[1303]NS2 (**16**) without the Jiv90 insert was generated by PCR-based deletion of Jiv90 (7.6.14) with phosphorylated primers CH009for and CH010rev (7.6.9).

7.8.2. Codon-optimized NS2 constructs with N-terminal His₆-tagged SUMO fusion

To improve *E. coli* expression, the NS2 portion in **6** was replaced by the one from **15**. Therefore, **6** and **15** were digested with restriction enzymes *Bam*HI and *Kpn*I (7.6.11) and the target pET-His₆-SUMO vector dephosphorylated (7.6.10). After PCR clean-up of the restriction fragments (7.6.6), they were ligated (7.6.12) to yield pET-His₆-SUMO-[1393]NS2_{Jiv90} (**17**).

pET-His₆-SUMO-[1393]NS2 (**18**) was derived from **17** by a deletion PCR based on site-directed mutagenesis (7.6.14) with phosphorylated primers CH009for and CH010rev (7.6.9) inversely flanking the Jiv90 insertion.

7.8.3. NS2 constructs with N- or C-terminal His₆-tagged GFP fusion

The bacterial expression vectors pCGFP-BC (**19**) and pNGFP-BC (**20**) as well as **17** were all digested with either 1) *Eco*RI and *Xho*I for [1303]NS2_{Jiv90} constructs, or 2) *Nco*I and *Xho*I for [1379]NS2_{Jiv90} constructs (7.6.11). The desired DNA fragments were purified by agarose gel extraction (7.6.6), ligated and transformed into *E. coli* XL1-Blue (7.6.15).

The resulting vectors pCGFP-[1303]NS2_{Jiv90} (**21**), pCGFP-[1379]NS2_{Jiv90} (**22**), pNGFP-[1303]NS2_{Jiv90} (**23**) and pNGFP-[1379]NS2_{Jiv90} (**24**) were subsequently submitted to a site-directed mutagenesis approach (7.6.14) to delete the Jiv90 insert from the constructs. Therefore, both vectors were amplified with phosphorylated primers CH009for and CH010rev (7.6.9) inversely flanking the Jiv90 insert to yield pCGFP-[1303]NS2 (**25**), pCGFP-[1379]NS2 (**26**), pNGFP-[1303]NS2 (**27**) and pNGFP-[1379]NS2 (**28**).

7.8.4. NS2 constructs for baculovirus-mediated expression in Sf21 and High FiveTM insect cells

For the generation of pETM11-hNS2_{Jiv90}-Avitag (**30**), originally intended for the generation of an NS2 truncation library, empty pETM11-Avitag vector (**29**) and full-length hNS2 (optimized for expression in human cell lines), from a PCR (7.6.13) with primers CH023for and CH024rev on **7**, were digested with restriction enzymes *Pst*I and *Aat*II (7.6.11). The desired fragments were purified by preparative agarose gel electrophoresis and subsequent extraction (7.6.5, 7.6.6), and ligated (7.6.12) to yield **30**.

pETM11-NS2-Avitag (**31**) and pETM11-NS2_{Jiv90}-Avitag (**32**) were constructed from **30** by replacing the C-terminal part of hNS2 (starting at threonine 1303) with the respective sequence from **16** or **15** with an RF-Cloning strategy (7.6.14). Megaprimers therefore were generated with primers CH027for and CH028rev on **16** or **15**, respectively.

To generate *Strept*II-tagged NS2 and NS2_{Jiv90} for insertion into pACEBac1 (**33**), a vector for transposition into a bacmid for insect cell expression, a PCR with primers CH032 and CH035 containing the coding sequence for the *Strept*II-tag was performed on the respective

vector, either **31** or **32**. **33** and the purified PCR products were digested with *EcoRI* and *HindIII*, and ligated to yield pACEBac1-NS2-*StreptII* (**34**) and pACEBac1-NS2_{Jiv90}-*StreptII* (**35**).

A full-length NS2 or NS2_{Jiv90} fragment was generated by PCR with primers NS2for and NS2rev on vector **34** or **35**, respectively. To insert the NS2 fragment into **10** and the NS2_{Jiv90} fragment into **9** or **11**, the vectors were separately digested with *SfoI* and in the same reaction simultaneously ligated to the corresponding blunt-end NS2 or NS2_{Jiv90} PCR product to yield pFastBac1-ER-NS2-KDEL (**36**), pFastBac1-ER-NS2_{Jiv90} (**37**) and pFastBac1-NS2_{Jiv90}-SKL (**38**).

pFastBac1-cyt-*StreptII*-NS2_{Jiv90} (**39**) and pFastBac1-ER-*StreptII*-NS2_{Jiv90} (**40**) were derived from **37** by PCR amplification of the whole vectors with phosphorylated primers cyt-*StreptII*-NS2-rev/*StreptII*-NS2-fw or ER-*StreptII*-NS2-rev/*StreptII*-NS2-fw, respectively. pFastBac1-ER-*StreptII*-NS2-KDEL (**41**) was derived from **36** by PCR with primers ER-*StreptII*-NS2-rev/*StreptII*-NS2-fw. pFastBac1-*StreptII*-NS2_{Jiv90}-SKL (**42**) was derived from **38** with primers cyt-*StreptII*-NS2-rev/*StreptII*-NS2-fw analogously. Subsequently, the PCR products were treated with *DpnI* and self-ligated.

The C-terminal 7 aa of NS3 were added by insertion PCR based on the site-directed mutagenesis approach. a) **39** and **40**, b) **41** as well as c) **42** were completely amplified with phosphorylated primer pairs a) NS2-rev/NS3-fw, b) NS2-rev/NS3-KDEL-fw or c) NS2-rev/NS3-SKL-fw, respectively, which all contain the coding sequence for the first 7 NS3 amino acids. The PCR products were digested with *DpnI* prior to self-ligation.

To introduce a C-terminal mCherry fusion into pFastBac1-cyt-*StreptII*-NS2_{Jiv90}-NS3_{7aa} (**43**), pFastBac1-ER-*StreptII*-NS2_{Jiv90}-NS3_{7aa} (**44**), pFastBac1-ER-*StreptII*-NS2-NS3_{7aa}-KDEL (**45**) and pFastBac1-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-SKL (**46**), a PCR product was generated from pET21a(+)-His-tag-mCherry (**47**) with the two primers mCherry-fw and mCherry-rev and subsequently digested with restriction enzymes *AatII* and *SacI*. The desired fragment was ligated into **43**, **44**, **45** or **46** prepared with the same enzymes to yield pFastBac1-cyt-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-mCherry (**48**), pFastBac1-ER-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-mCherry (**49**), pFastBac1-ER-*StreptII*-NS2-NS3_{7aa}-mCherry (**50**) and pFastBac1-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-mCherry-SKL (**51**).

pFastBac1-cyt-*StreptII*-NS2-NS3_{7aa}-mCherry (**52**) and pFastBac1-ER-*StreptII*-NS2-NS3_{7aa}-mCherry (**53**) were derived from **48** and **49**, and pFastBac1-*StreptII*-NS2-NS3_{7aa}-mCherry-SKL (**55**) from **51** by PCR-based deletion of the Jiv90 insert with primers CH009for and CH010rev. After digestion with *DpnI* to remove parental vector, the PCR products were self-ligated.

To create pFastBac1-ER-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-mCherry-KDEL (**54**), **49** and **50** were digested with restriction enzymes *AatII* and *KpnI*. The desired fragments were purified by agarose gel extraction and ligated, yielding **54**.

7.8.5. NS2 constructs for expression in the cell-free BY2 tobacco cell system

pEUMod-*StreptII*-NS2-NS3_{7aa}-EGFP (**56**) and pEUMod-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-EGFP (**57**) were derived from **12** by a classical cloning strategy. The full-length NS2 or NS2_{Jiv90} insert was amplified by PCR (7.6.13) from **34** or **35** (7.8.4) with primers pEU-*Strep*-NS2-fw and

pEU-NS2-NS3-rv. The PCR product was purified by agarose gel extraction (7.6.5, 7.6.6) and subsequently digested with restriction enzymes *BspHI* and *XhoI*. **12** was digested with *NcoI* and *XhoI* and the desired vector backbone fragment purified by agarose gel extraction. Vector backbone and corresponding NS2-NS3_{7aa} or NS2_{Jiv90}-NS3_{7aa} insert were ligated.

pEUmod-*StreptII*-NS2-NS3_{7aa}-EGFP-TMV3 (**58**) and the Jiv90 variant (**59**) were generated from **56** and **57** by PCR-amplification with primers pEU-Strep-NS2-fw and pEU-GFP-Sall-rv. The purified PCR product was digested with restriction enzymes *BspHI* and *Sall* and ligated into the target vector **14**, which was digested with *NcoI* and *XhoI*, dephosphorylated with FastAP (Thermo Scientific; 7.6.10) and purified by agarose gel extraction.

To construct pUC57-BY2-His₆-PreSc-*StreptII*-NS2_{Jiv90}-NS3_{7aa}-EGFP (**60**), the required insert was first generated from **57** by PCR with primers pEU-Strep-NS2-fw and pEU-GFP-Sall-rv. The PCR product was purified and digested with restriction enzymes *BspHI* and *Sall*. The target vector **13** was digested with *NcoI* and *XhoI*, dephosphorylated with FastAP (Thermo Scientific) and purified by agarose gel extraction. Subsequently insert and vector were ligated to yield **60**.

For the generation of pUC57-BY2-His₆-PreSc-*StreptII*-NS2-NS3_{7aa}-EGFP (**61**), the Jiv90 coding sequence was removed from **60** by PCR-based deletion (7.6.14) with phosphorylated primers CH009for and CH010rev (7.6.9).

pUC57-BY2-*StreptII*-NS2-NS3_{7aa}-EGFP (**62**) as well as the Jiv90 variant (**63**) were derived from **60** or **61** by deletion PCR with phosphorylated primers pUC57-BY2-del-for and pUC57-BY2-del-rev.

7.9. Cloning of *Sorghum bicolor* DHS and eIF-5A isoform constructs

pET-*Zm*eIF5A-His₆ (**64**), pET-*Sb*DHS1-His₆ (**75**), pET-*Sb*DHS2-His₆ (**74**), pET-*Sb*DHS3-His₆ (**73**) and pCDFDuetTM 1 (**76**; Novagen[®]) were kindly provided by the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober at the Botanical Institute of Kiel University. pET vector backbones are originally derived from pET23 (Novagen[®]).

7.9.1. Cloning of *Sb*eIF-5A-His₆ isoform constructs

Prior to use, all reusable material that would come into contact with RNA containing sample was treated with RNase AWAY[®] (Molecular BioProducts).

A fresh leaf from *S. bicolor* (Botanical Garden, Kiel University) was flash frozen in liquid N₂ and ground to a fine powder in a ceramic mortar. For RNA extraction from leaf powder and reverse transcription of the extracted RNA into cDNA, the innuPREP Plant RNA Kit (analytikjena) and the First Strand cDNA Synthesis Kit (Thermo Scientific) were applied according to manufacturer's recommendations.

To insert the coding DNA sequence for both eIF-5A isoforms from *S. bicolor* into a bacterial expression vector, an RF-Cloning strategy (7.6.14) was applied, using primers *Sb*eIF5A-for and *Sb*eIF5A1-rev (*Sb*eIF-5A1) or *Sb*eIF5A2-rev (*Sb*eIF-5A2), respectively. With this technique, the sequence encoding *Zea mays* eIF-5A was substituted for one of the *Sb*eIF-5A isoform

sequences in the vector pET-Zmelf5A-His₆ (**64**), which already proved to be a successful expression construct. The resulting vectors were termed pET-Sbelf5A1-His₆ (**65**) and pET-Sbelf5A-His₆ (**66**).

7.9.2. Cloning of *Sbelf5A-His₆-SAStrepII* isoform constructs

For the addition of a C-terminal serine-alanine-*Strep*-tag II (*SAStrep*-tag II) to both pET-Sbelf5A-His₆ isoform constructs (**65**, **66**), a PCR-based insertion strategy was chosen (7.6.14). Both vectors were amplified with primers *Sbelf5A1-StrepF* and *Sbelf5A1-StrepR*. Parental vector was digested with *DpnI*. 2 µL of digested, heat inactivated PCR products were transformed into *E. coli* XL1-Blue (7.6.15) to yield pET-Sbelf5A1-His₆-*SAStrepII* (**67**) and pET-Sbelf5A2-His₆-*SAStrepII* (**68**).

7.9.3. Cloning of *StrepII-SbDHS-His₆* isoform constructs

For the generation of pET-*StrepII-SbDHS1-His₆* (**79**), pET-*StrepII-SbDHS2-His₆* (**78**) and pET-*StrepII-SbDHS3-His₆* (**77**), an N-terminal *StrepII*-tag and a thrombin cleavage site were introduced into **75**, **74** and **73** by a site-directed mutagenesis approach (7.6.14). The respective pET-*SbDHS-His₆* vector was PCR-amplified with 5'-overlapping primers CH037fw and CH031rv (DHS1), CH029fw and CH031rv (DHS2) or CH036fw and CH031rv (DHS3).

7.9.4. Cloning of *SAStrepII-SbDHS-His₆* isoform constructs

pET-*SAStrepII-SbDHS1-His₆* (**82**), pET-*SAStrepII-SbDHS2-His₆* (**81**) and pET-*SAStrepII-SbDHS3-His₆* (**80**) were obtained from **79**, **78** and **77** with a site-directed mutagenesis approach (7.6.14) using primers *SAStrepII-for* and *SAStrepII-rev*.

7.9.5. Cloning of *SbDHS* coexpression constructs

For the cloning of *SbDHS* coexpression constructs, the commercially available pCDFDuet-1 vector (**76**) was first modified to suit the cloning strategy. Therefore, **76** was PCR-amplified based on a site-directed mutagenesis approach (7.6.14) with phosphorylated primers (7.6.9) *pCDFDuet1-StrepII-for* and *pCDFDuet1-StrepII-rev*, which contained the *Strep*-tag II sequence to replace the original S-tag sequence in multiple cloning site2 (MCS2). *pCDFDuet-His₆-StrepII* (**83**) was further optimized by deleting the *SaI* site in MCS1 with primers *SaI*mut-for and *SaI*mut-rev. Vector DNA of **83** and optimized *pCDFDuet-His₆-StrepII*[mod] (**84**) was each prepared from a single clone and sequenced to verify the integrity of both vectors before the next cloning steps.

Vectors containing N-terminally His₆-tagged *SbDHS2* or *SbDHS1* in MCS1 (**86**, **87**) were created by a classical cloning strategy. **83** as well as **74** or **75** were digested with *HindIII* and *NcoI* and the desired DNA fragments purified by agarose gel extraction (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel; 7.6.6). Vector and corresponding insert were ligated (7.6.12). *pCDFDuet-SbDHS3-His₆-StrepII* (**85**) could be derived from **86** by substituting the *SbDHS2* for *SbDHS3* with an RF-Cloning approach (7.6.14) using primers RF-DHS3for/RF-DHS3rev on **73** to generate the necessary megaprimers.

To create pCDFDuet-His₆-*SbDHS3-StrepII* (**88**), a PCR product was generated using primers *SbDHS3-NdeI*-for and *SbDHS3-XhoI*-rev on the template plasmid **73**. The PCR product was subsequently digested with *NdeI* and *XhoI* and ligated to **84** digested with *NdeI* and *Sall*. To create pCDFDuet-His₆-*SbDHS2-StrepII* (**89**) and pCDFDuet-His₆-*SbDHS1-StrepII* (**90**), an RF-Cloning strategy was chosen, in which the DHS3 of **88** was replaced by either DHS2 or DHS1. The megaprimers for this approach were generated with primers RF2-DHS2-for and RF2-DHS2-rev on **74** or with primers RF2-DHS2-for and RF2-DHS1-rev on **75**, respectively.

Coexpression vector pCDFDuet-*SbDHS1-His6-SbDHS3-StrepII* (**91**) was obtained with a classical cloning strategy. The purified product of a PCR (7.6.13) with primers *SbDHS3-NdeI*-for and *SbDHS3-XhoI*-rev on **73** was digested with *NdeI* and *XhoI* and subsequently ligated to **87** digested with *NdeI* and *Sall*. pCDFDuet-*SbDHS1-His6-SbDHS2-StrepII* (**92**) was created with an RF-Cloning approach using the megaprimers from pCDFDuet-His₆-*SbDHS2-StrepII* cloning on vector **91**. To obtain coexpression vector pCDFDuet-*SbDHS1-His6-SbDHS1-StrepII* (**93**), vectors **87** and **90** were digested with *HindIII* and *XbaI*. The desired fragments of both digestions were purified by agarose gel extraction (NucleoSpin[®] Gel and PCR Clean-up, Macherey-Nagel) and ligated.

7.9.6. Cloning of *SbDHS* constructs for baculovirus-mediated expression in Sf21 and High Five[™] insect cells

pFastBac1-ER-*SbDHS1* (**96**), pFastBac1-ER-*SbDHS2* (**95**) and pFastBac1-ER-*SbDHS3* (**94**) were constructed by an RF-Cloning strategy (7.6.14). Therefore, megaprimers were generated on **75** with primers ER-DHS2-for/ER-DHS1-rev. With these megaprimers, the coding sequence for *SbDHS1* was inserted into **9**. In a next step, the *SbDHS1* sequence was substituted in **96** by the ones for *SbDHS2* or *SbDHS3*. Megaprimers containing these sequences were generated with primers ER-DHS2-for/ER-DHS2-rev or ER-DHS3-for/ER-DHS3-rev on **74** or **73**, respectively.

pFastBac1-ER-*SbDHS1-KDEL* (**97**) was derived from **96** with an RF-Cloning strategy. Megaprimers therefore were generated with primers ER-DHS2-for/KDEL-DHS1-rev on **75**. For the construction of pFastBac1-ER-*SbDHS3-KDEL* (**98**) and pFastBac1-ER-*SbDHS2-KDEL* (**99**), the coding sequence for *SbDHS1* was substituted in **97** by the ones for *SbDHS3* or *SbDHS2* with megaprimers that were generated with primers ER-DHS3-for/KDEL-DHS3-rev or ER-DHS2-for/KDEL-DHS2-rev on **73** or **74**, respectively.

For the generation of pFastBac1-cyt-*SbDHS1* (**102**), pFastBac1-cyt-*SbDHS2* (**101**), pFastBac1-cyt-*SbDHS3* (**100**) and pFastBac1-*SbDHS3-SKL* (**103**), the desired sequences were PCR-amplified (7.6.13) with phosphorylated primers (7.6.9) cyt-DHS2-for/cyt-DHS1-rev, cyt-DHS2-for/cyt-DHS2-rev, cyt-DHS3-for/cyt-DHS3-rev or CH038for/CH039rev on **97**, **98**, **99** or **77**, respectively. The PCR products were purified (7.6.6) and subsequently ligated (7.6.12) into *SfoI*-digested, dephosphorylated (7.6.10) pFastBac1-cyt (**8**).

pFastBac1-*SbDHS1-SKL* (**104**) was constructed with an RF-Cloning strategy. Therefore, megaprimers were generated with primers SKL-DHS2-for/SKL-DHS1-rev on **75** and applied to insert the *SbDHS1* coding sequence into pFastBac1-SKL (**11**). With megaprimers that

were generated with primers SKL-DHS2-for/SKL-DHS2-rev on **74**, pFastBac1-SbDHS2-SKL (**105**) was derived from **104**.

7.10. Culture of *E. coli* for protein expression

If not stated otherwise, *E. coli* cultures throughout this thesis were performed in **LB medium** (Luria/Miller; 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.0) in three different scales: mini (5 to 6 mL medium in 160x15 mm glass tubes), midi (baffled and plain 100, 250 or 500 mL Erlenmeyer flasks, filled with medium between 1/5 and 2/5 of the flask's volume) and maxi (baffled and plain 1, 2 or 5 L Erlenmeyer flasks, filled with medium between 1/5 and 2/5 of the flask's volume). The cultures were shaken in a an orbital shaker (Infors Multitron) to supply the growing cultures with oxygen. Mini cultures were shaken in tilted racks at 230 rpm. For midi and maxi cultures, the shaking speed was adjusted according to flask size, form and filling volume, ranging from about 90 rpm for 2 L medium in baffled 5 L flasks to about 180 rpm for plain 100 mL flasks. The cultures were normally grown at 37 °C.

E. coli expression cultures of midi or maxi size were inoculated with pelleted bacteria from small scale overnight cultures – mini or midi size – to an OD₆₀₀ of about 0.1. For expressions at 30 to 37 °C for up to 6 h, the cultures were grown to an OD₆₀₀ of 0.6 before induction with IPTG, while the expression at 10 or 20 to 30 °C for 16 to 24 h was induced at an OD₆₀₀ of about 1 or higher.

For expression trials, different media (LB (Luria/Miller) and LB with supplements like glucose to reduce basal expression; terrific broth (TB, [267]); super optimal broth (SOB, [99]) and SOB with additional glucose (SOC); ZYP-AIM [265]), expression temperatures (10, 18, 20, 22, 25, 30, 33 and 37 °C) and times (3 to 24 h), as well as IPTG concentrations in the range of 0.05 to 1 mM were applied.

7.11. Expression and Purification of *Sb*elF-5A isoform constructs

7.11.1. Expression of *Sb*elF-5A-His₆ isoforms and the SA*Stre*ptII-tagged variants

E. coli BL21(DE3) or BL21-Gold(DE3) harboring one of the two *Sb*elF-5A-His₆ isoform constructs or the SA*Stre*ptII-tagged versions were pre-cultured overnight at 37 °C in 250 mL **LB medium** (Luria/Miller; 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.0) containing 100 µg mL⁻¹ ampicillin and 12.5 µg mL⁻¹ tetracycline (only BL21-Gold(DE3)). For the His₆-tagged isoforms, 2 times 2 L LB medium (Luria/Miller) containing 100 µg mL⁻¹ ampicillin and several drops of Antifoam A in baffled 5 L Erlenmeyer flasks were prepared, while the SA*Stre*ptII-tagged versions were expressed in 3 times 1 L LB medium in baffled 5 L Erlenmeyer flasks.

The main cultures were inoculated from the precultures to an OD₆₀₀ of 0.1. The cultures were incubated at 37 °C and 100 rpm until the OD₆₀₀ reached 1.5 for the His₆-tagged or 0.6 to 0.7 for the SA*Stre*ptII-tagged versions. The expression was induced with 0.5 mM IPTG and maintained for 22 to 26 h at 22 °C for the His₆-tagged isoforms. The SA*Stre*ptII-tagged versions were expressed for 4 h at 37 °C.

The cells were harvested in several rounds of centrifugation for 10 min at 10 000 rcf and 4 °C. The pellet was subsequently resuspended in **PBS** (10 mM NaH₂PO₄, 150 mM NaCl, pH 7.4 at 22 °C) and centrifuged for another 20 min at 4600 rpm and 4 °C. The pellet was flash frozen in liquid N₂ and stored at –80 °C.

7.11.2. Purification of *Sbelf-5A-His₆* and the *SASTrepII*-tagged isoforms

10 to 14 g pellet from the expression cultures were thawed in ice water for approx. 1 h and afterwards resuspended in 100 mL cold **lysis buffer** (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 1 mM DTT, 0.1 mM PMSF, pH 7.5 at 22 °C). The suspension was kept on ice. The cells were disrupted by 2 passages through the EmulsiFlex-C3 Cell Homogenizer (Avestin) at 1200 bar in the cold room. To remove cell debris and to prepare the cleared lysate for application onto the ÄKTA chromatography system (GE Healthcare), the lysate was first centrifuged at 75 600 rcf, 10 °C for 1 h and subsequently vacuum filtrated through a 0.45 µm cellulose acetate membrane.

The cleared lysate was applied onto a 5 mL HisTrap HP (GE Healthcare), which was pre-equilibrated with **binding buffer** (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 1 mM DTT, pH 7.5 at 22 °C). After sample application, the column was washed with binding buffer until baseline and the bound protein was eluted in one step with 50 % **elution buffer** (50 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole, 1 mM DTT, pH 7.5 at 22 °C). Fractions containing the His-tagged *Sbelf-5A* were pooled according to SDS-PAGE results and concentrated in an Amicon Ultra 4 (3 kDa MWCO) centrifugal filter (Merck-Millipore) to between 10 and 17 mg mL⁻¹.

In the case of the *Sbelf-5A-His₆* isoforms, the concentrated samples were applied onto a Superdex 75 HiLoad 16/60 (GE Healthcare) size exclusion chromatography (SEC) column, which was pre-equilibrated in **running buffer** (50 mM Tris-HCl, 150 mM NaCl, pH 7.5 at 22 °C). Fractions containing the *Sbelf-5A* were pooled according to retention time and SDS-PAGE results and concentrated in an Amicon Ultra 4 (3 kDa MWCO) centrifugal filter (Merck-Millipore) to 15 to 20 mg mL⁻¹.

The *Sbelf-5A-His₆-SASTrepII* isoforms were further purified by *Strep-Tactin*[®] instead of SEC. The concentrated protein was applied to a 2 mL *Strep-Tactin*[®] gravity flow column pre-equilibrated with **Buffer W** (iba; 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0) and washed with 5 times 1 CV Buffer W. The protein was eluted with 6 times 0.5 CV **Buffer E** (iba; 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, pH 8.0). Since the *Strep-Tactin*[®] gravity flow column was heavily overloaded, the procedure was repeated 3 times with the flow through. The eluted protein was pooled according to SDS-PAGE results. The buffer was exchanged to SEC running buffer with a HiTrap Desalting column (GE Healthcare) and the protein afterwards concentrated in an Amicon Ultra 4 (3 kDa MWCO) centrifugal filter (Merck-Millipore) to 10 to 12 mg mL⁻¹.

All concentrated protein samples were aliquoted, flash frozen in liquid N₂ and stored at –80 °C.

7.12. Expression and purification of *SbDHS/HSS* constructs

7.12.1. Expression of *SbDHS/HSS* constructs

The expression of *SbDHS1-His₆*, *SbDHS2-His₆* and *SbDHS3-His₆* as well as *SASTrepII-SbDHS1-His₆*, *SASTrepII-SbDHS2-His₆* and *SASTrepII-SbDHS3-His₆* was performed in *E. coli* BL21 (DE3) or BL21-Gold(DE3). *StreptII-SbDHS1-His₆*, *StreptII-SbDHS2-His₆* and *StreptII-SbDHS3-His₆* were expressed in RosettaTM 2(DE3).

The main expression cultures of 2 L **LB medium** (Luria/Miller; 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, pH 7.0) with appropriate antibiotics in baffled 5 L Erlenmeyer flasks were inoculated to an OD₆₀₀ of about 0.1 from midi scale overnight cultures (7.10) started from a glycerol stock and grown at 37 °C. The main cultures were supplied with 5–10 drops of Antifoam A and shaken at 37 °C and 110 rpm in an orbital shaker (Infors Multitron) until the OD₆₀₀ reached 0.7–0.8 for His₆- and *StreptII*-tagged variants or 1.0–1.3 for *SASTrepII*-tagged versions. The expression was induced with 0.5 mM IPTG and maintained at 22 or 25 °C for 20–24 h.

The cells were harvested in several rounds of centrifugation for 10 min at 10 000 rcf and 4 °C. The pellets were subsequently resuspended in **PBS** (10 mM NaH₂PO₄, 150 mM NaCl, pH 7.4 at 22 °C) and centrifuged for another 20 min at 4600 rpm and 4 °C. The pellets were flash frozen in liquid N₂ and stored at –80 °C.

7.12.2. Purification of *SbDHS/HSS* constructs

IMAC of His₆-tagged *SbDHS/HSS*

The frozen cell pellets from the expression cultures (7.12.1) were thawed in ice water and resuspended in a lysis buffer volume of at least 10 times of the pellet wet weight (**lysis buffer**: 100 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 2–5 mM NAD⁺, 0.5–1 mM PMSF, pH 8.5 at 22 °C). The cells were lysed by two passages at 1200 bar through a cell homogenizer (EmulsiFlex-C3, Avestin) in the cold room. After 1 h centrifugation at 75 600 rcf and 10 °C, and filtration through a 0.45 µm cellulose acetate filter, the cleared lysate was applied to a 1 or 5 mL HisTrap HP column (GE Healthcare) equilibrated with **binding buffer** (100 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 1–2 mM NAD⁺, pH 8.5 at 22 °C) at an ÄKTA chromatography system (GE Healthcare). The column was washed with binding buffer for 5–15 CV until the 280 nm baseline was stable and subsequently washed with 5–10 CV of a high salt **washing buffer** (100 mM Tris-HCl, 1 M NaCl, 20 mM imidazole, 1–2 mM NAD⁺, pH 8.5 at 22 °C) followed by further 3–5 CV of binding buffer. The bound sample was eluted with a gradient from the initial 20 to 260 mM imidazole over 20 CV. The elution fractions were pooled according to chromatogram and SDS-PAGE results.

Affinity purification of *StreptII*-tagged *SbDHS/HSS*

For the purifications with *Strep-Tactin*[®], the frozen cell pellets from the expression cultures (7.12.1) were thawed in ice water and resuspended in a lysis buffer volume of at least 10 times of the pellet wet weight (**lysis buffer**: 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 2 mM

NAD⁺, 0.5 mM PMSF, pH 8.0 at 22 °C). Lysis and clarification were done in the same way as for His₆-tagged proteins (see subsection above). Cleared lysate was applied to a *Strep-Tactin*[®] gravity flow column equilibrated with **Buffer W** (iba; 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0 at 22 °C). The column was washed with 8 times 1 CV Buffer W and the bound proteins eluted with 6 times 0.5 CV **Buffer E** (iba; 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, pH 8.0 at 22 °C).

The elution fractions of *StreptII*-tag purified proteins were pooled according to SDS-PAGE results.

SEC of affinity-purified *SbDHS/HSS* fusion proteins

The concentrated samples from His₆- or *StreptII*-tag-based affinity purifications were applied onto a Superdex 200 HiLoad 16/60 SEC column (GE Healthcare) equilibrated with **running buffer** (100 mM glycine-NaOH, 150 mM NaCl, pH 9.0 at 22 °C). Fractions were again pooled according to retention time and SDS-PAGE results and concentrated in an Amicon Ultra centrifugal filter device with a MWCO of 30 kDa (Merck-Millipore) to about 10 g L⁻¹.

The concentrated, purified protein solution was supplied with 2 mM NAD⁺, and centrifuged at 21 000 rcf for at least 10 min prior to aliquotation. The aliquots were then flash frozen in liquid N₂ for long-term storage at -80 °C.

7.12.3. Purification of *SbDHS/HSS* coexpression and control *SbDHS/HSS* for the detection of heteromers

SbDHS1-His₆ and the coexpressed *SbDHS1*-His₆/*SbDHS1*-*StreptII*, *SbDHS1*-His₆/*SbDHS2*-*StreptII* and *SbDHS1*-His₆/*SbDHS3*-*StreptII* from their respective pCDFDuet vectors were first expressed in BL21-Gold(DE3) as described in section 7.12.1 and subsequently purified by an initial immobilized metal affinity chromatography (IMAC) using the *N*-terminal His₆-tag of the recombinant proteins (see section 7.12.2). In contrast, *SbDHS1*-*StreptII*, *SbDHS2*-*StreptII* and *SbDHS3*-*StreptII* were purified by an affinity chromatography (AC) step with *Strep-Tactin*[®] (iba; 7.12.2). Afterwards, all recombinant proteins were polished by a size exclusion chromatography (SEC) step (see section 7.12.2).

To ensure that the detected interaction between paralogs in the following western blot was indeed due to heteromer formation, the pre-purified protein from coexpressed *SbDHS1*-His₆/*SbDHS1*-*StreptII* (positive control), *SbDHS1*-His₆/*SbDHS2*-*StreptII*, *SbDHS1*-His₆/*SbDHS3*-*StreptII* and separately expressed and purified and then mixed *SbDHS1*-His₆ and *SbDHS1*-*StreptII*, *SbDHS2*-*StreptII* or *SbDHS3*-*StreptII* (negative controls) was further purified in a second small-scale His₆-tag purification round.

For the negative control, approx. 100 µg of *SbDHS1*-His₆ and *SbDHS1*-*StreptII*, *SbDHS2*-*StreptII* or *SbDHS3*-*StreptII* were mixed together, diluted to 200 µL in **binding buffer** (100 mM Tris-HCl, 300 mM NaCl, 1 mM NAD⁺, pH 8.5 at 22 °C) and incubated overnight at 4 °C. 100 µg of each other protein solution was diluted to 200 µL in binding buffer directly prior to the purification. The samples were loaded onto 100 µL Ni-NTA resin (Cube Biotech) in an 0.8 mL centrifuge column (Pierce[®]) equilibrated with binding buffer. The bound tetramers were rigorously washed with 8 times 4 CV **washing buffer** (100 mM Tris-HCl, 1 M NaCl,

1 mM NAD⁺, 2% (v/v) Tween[®] 20, pH 8.5 at 22 °C) and 2 times 4 CV binding buffer. The proteins were eluted with 6 times 1 CV **elution buffer** (100 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole, 1 mM NAD⁺, pH 8.5 at 22 °C).

7.13. Insect cell culture techniques

Sf21 insect cell stocks for the initiation of a continuous culture were a kind gift from the Eukaryotic Expression Facility at the European Molecular Biology Laboratory (EMBL) Grenoble. High Five[™] insect cell stocks were kindly provided by the work group Structural Infection Biology applying new Radiation Sources of Prof. Dr. Lars Redecke from Hamburg and Lübeck University.

Continuous suspension cultures of Sf21 and High Five[™] insect cells were grown shaking at 105 rpm in upright tissue culture flasks (T25 or T75 Cell+, Sarstedt) and split to 0.6-0.8x10⁶ cells/mL (Sf21) or 0.3-0.4x10⁶ cells/mL (High Five[™]) in ESF 921 Insect Cell Culture Medium (Expression Systems) every two or three days. After 30 to 40 passages, a fresh culture was started from a frozen cell stock.

Cells for transfection, baculovirus stock amplification or *in vivo* crystallization in 6-well plates (Cell+, Sarstedt) were usually seeded from these suspension cultures to a density of 0.6x10⁶ cells per well in ESF 921 Insect Cell Culture Medium. They were allowed to attach to the bottom of the well for at least 30 min prior to the following transfection or infection procedure.

Suspension cultures as well as adherent cultures were always grown at 27 °C.

7.13.1. Transposition of pFastBac1-cyt, -ER, -KDEL and -SKL constructs into baculovirus shuttle-vectors bMON14272 and EMBacY

For the transposition of a pFastBac1 construct into a baculovirus shuttle vector, a simplified Tn7 transposase system is used. Therefore, the pFastBac1 construct is transformed via heat shock (7.6.15, bacmid transposition protocol) into competent *E. coli* DH10Bac or DH10EMBacY. These *E. coli* strains harbor one of the baculovirus vectors bMON14272 or EMBacY, which contain a mini-Tn7 transposon, and a helper plasmid pMON7124, which provides the necessary transposase genes. In the pFastBac1 vectors, the constructs to be transposed are flanked by sequences which enable the insertion of these DNA sequences into the baculovirus shuttle vector – thus substituting the coding sequence for the lacZ α fragment (complements the lacZ Δ M15 mutation of the β -galactosidase), which is located in the mini-Tn7 transposon. The transformed bacteria are streaked onto **LB agar** plates (Luria/Miller; 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar, pH 7.0) containing 50 μ g mL⁻¹ kanamycin for the selection of the baculovirus shuttle vector, 12.5 μ g mL⁻¹ tetracycline to maintain the helper plasmid and 7 μ g mL⁻¹ gentamycin for the expression cassette with the desired construct. To enable blue-white screening, the plates also contained 100 μ g mL⁻¹ 5-Bromo-4-chloro-3-indolyl β -D-galactopyranoside (X-gal) and Isopropyl β -D-1-thiogalactopyranoside (IPTG).

After a successful transposition, the coding sequence for the lacZ α fragment is substituted with the desired construct. The bacteria, which now only express the inactive lacZ Δ M15

mutant of the β -galactosidase, are not able to convert X-gal to the intense blue 5,5'-dibromo-4,4'-dichloro-indigo product anymore and therefore will appear white. In contrast, bacteria harboring the pFastBac1 vector, which are resistant to gentamycin but have not undergone transposition yet, will express the lacZ α fragment due to IPTG induction. This fragment activates the lacZ Δ M15 mutant and the colony turns blue due to β -galactosidase activity. To ensure that the clone of choice contained only transposed bacmid, a white clone was subsequently re-streaked twice onto fresh agar plates supplemented with the necessary antibiotics, X-gal and IPTG and finally tested for the expression cassette in the transposon site by a colony PCR (7.6.13) with primers M13for and M13rev.

7.13.2. Transfection of Sf21 insect cells with baculovirus shuttle vectors

The transfection procedure described here essentially follows the protocol from "The MultiBac BEVS for producing proteins and their complexes (Prot54)" of the MultiBac Protein Complex Production Platform at the EMBL Grenoble [29, 25].

Prior to the transfection procedure, Sf21 cells were seeded into 6-well plates at a density of 0.6×10^6 cells per well in ESF 921 Insect Cell Culture Medium. The cells were allowed to adhere for at least 30 min.

DH10Bac or DH10EMBaCY mini cultures with the desired expression construct were grown overnight at 37 °C and 230 rpm. Cells from 2 mL of this culture were pelleted by centrifugation, resuspended in 250 μ L buffer A1 from the NucleoSpin[®] Spin Plasmid Kit (Macherey-Nagel) and subsequently lysed by the addition of 250 μ L buffer A2. After neutralization with 300 μ L buffer A3 the lysate was centrifuged at 21 000 rcf for 10 min. The supernatant was transferred into a fresh tube and the bacmid DNA precipitated with 0.7 volumes of 2-propanol. The solution was centrifuged at 21 000 rcf for another 10 min and the resulting bacmid pellet washed with 200 μ L 70 % (v/v) ethanol followed by another centrifugation step. The pellet was transferred to a sterile safety cabinet and allowed to dry for several minutes. Afterwards it was resuspended in 200 μ L ESF 921 Insect Cell Culture Medium.

Meanwhile, 10 μ L X-tremeGENE[™] HP DNA Transfection Reagent (Roche) was mixed with 100 μ L ESF 921 Insect Cell Culture Medium for each construct to be transfected. 100 μ L of this mixture was added to the bacmid solution. 2x 150 μ L from this mixture were applied drop-wise to two different wells, thus creating duplicates for each construct.

Virus of the first generation (P1) was harvested between 48 to 72 h after the transfection, the supernatants of the duplicates were combined at this point. The cells were harvested 7 d after the transfection to check for protein expression by SDS-PAGE. During the 7 d incubation, the cells transfected with an EMBaCY construct could be checked visually for the expression of EYFP with a Blue/Green LED Transilluminator (Nippon Genetics) and the corresponding Amber filter. The EYFP gene in the baculovirus is under the control of the same promoter as the expression construct, so that the amount of expressed EYFP should correlate with the amount of expressed protein of interest. Thereby the EYFP fluorescence is a valuable tool to assess the quality of the transfection and later on of the infection with different baculovirus stock generations.

7.13.3. Baculovirus amplification

To obtain high titer baculovirus stocks for protein expression and *in vivo* crystallization, Sf21 insect cells were seeded into 6-well plates and allowed to adhere for at least 30 min. Usually, three wells per construct were infected with three different volumes (300, 600 and 900 μL) of P1 virus stock from the transfection. Second generation baculovirus (P2) was harvested 48 to 72 h after the infection. Freshly seeded Sf21 insect cells at a density of 0.6×10^6 cells per well were infected with 75-100 μL of P2 and third generation baculovirus (P3) could be obtained 48 to 72 h after the infection.

During baculovirus amplification cells infected with an EMBacY-derived baculovirus could be checked for EYFP expression on a Blue/Green LED Transilluminator (Nippon Genetics) equipped with the corresponding Amber filter.

7.13.4. *In vivo* crystallization in Sf21 or High FiveTM insect cells

Sf21 or High FiveTM insect cells were seeded into 6-well plates and allowed to adhere for at least 30 min. The cells were infected with 20 to 40 μL of P3 baculovirus stock and checked daily for signs of intracellular crystal growth with an inverted microscope (Olympus CKX41 equipped with an Olympus C-5060 camera). Usually first crystals appeared after 5 d.

7.14. Biophysical techniques

7.14.1. Fluorescence microscopy of Sf21 or High FiveTM insect cells

Fluorescence images of Sf21 or High FiveTM insect cells were obtained with an inverted fluorescence microscope (AxioVert 200 M, Zeiss) equipped with filters 38HE, 43HE and 49 (Zeiss). Images were taken at 10x, 20x or 40x magnification and exposure times were adjusted to match the requirements of the corresponding samples. Access to the microscope at the Institute of Anatomy of Kiel University was kindly provided by PD. Dr. Kirsten Hattermann-Koch. Technical support with microscope setup and imaging was contributed by Judith Becker.

7.14.2. Immunofluorescence staining of NS2-expressing High FiveTM

High FiveTM insect cells were seeded onto sterile poly-D-lysine coated glass cover slips in 6-well plates or directly into the wells at a density of 0.6×10^6 cells per well in ESF 921 Insect Cell Culture Medium (Expression Systems). The cells were infected with 50 μL P3 virus stock of the respective constructs and incubated at 27 °C for several days until *in vivo* crystals appeared.

The cells grown on coated cover slips, were washed twice with 2 mL PBS (50 mM $\text{Na}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, 150 mM NaCl, pH 7.5) after aspiration of the medium. The cells were fixed in freshly prepared 4% (w/v) paraformaldehyde in PBS for 15 min and afterwards washed 3 times with 2 mL PBS, followed by a permeabilization step with 0.2% (v/v) TritonTM X-100 in PBS for 15 min. The cells were washed again 3 times in 2 mL PBS and blocked with 3% (w/v) BSA in PBS for 1 h. After washing the cells 3 times with 2 mL PBS, the cover slips were transferred into a humidified box. The cell-coated surface was overlaid with 60 μL of a 1:200 dilution

of a Chromeo™ 488-conjugated StrepMAB-Classic antibody (iba) in PBS with 0.2% (v/v) Triton™ X-100. The cover slips were incubated for 1 h in the dark. Afterwards, the cover slips were washed 3 times in 2 mL PBS and overlaid with 300 nM 4',6-diamidino-2-phenylindole (DAPI) in PBS and kept for 5 min in the dark to stain the nucleus. After another 3 times of washing in PBS, the cover slips were carefully patted dry and mounted cell-side-down onto a drop of ProLong™ Diamond Antifade Mountant (Thermo Scientific) on a microscope slide.

The cells directly grown in 6-well plates were carefully resuspended in 1 mL PBS and afterwards fixed in 1 mL 4% (w/v) paraformaldehyde in PBS for 15 min. The cells were then washed twice in 1 mL PBS and permeabilized in 1% (v/v) Triton™ X-100 in PBS for 15 min. The cells were washed again and blocked in 0.1% (w/v) BSA and 0.2% (w/v) glycine in PBS (**blocking solution**) for 1 h. Thereafter, the cells were stained with 100 µL of a 1:200 dilution of a Chromeo™ 488-conjugated StrepMAB-Classic antibody (iba) in blocking solution at 800 rpm in a ThermoMixer® (Eppendorf) over night in the dark. The cells were subsequently washed again twice in 1 mL PBS and then stained for 5 min in 100 µL of 300 nM DAPI in PBS. After another 3 times of washing in PBS, the cells were carefully sedimented by low-force centrifugation, and after aspiration of the supernatant gently mixed with a drop of ProLong™ Diamond Antifade Mountant (Thermo Scientific). The mixture was applied dropwise onto a microscope slide and covered with a glass cover slip.

After curing for 24 h in the dark, the cells were examined with an inverted fluorescence microscope (see section 7.14.1).

7.14.3. Transmission electron microscopy (TEM) of insect cells

High Five™ insect cells were seeded onto sterile poly-D-lysine coated glass cover slips in 6-well plates at a density of 0.6×10^6 cells per well in ESF 921 Insect Cell Culture Medium. The cells were infected with 20-50 µl P3 virus stock of the respective constructs and incubated at 27 °C for several days until *in vivo* crystals appeared.

After aspiration of the medium, the cells were washed three times with 2 mL PBS (100 mM Na/KH₂PO₄, 150 mM NaCl, pH 7.5) and fixed with 500 µl 3% (w/v) glutaraldehyde in **phosphate buffer** (100 mM Na/KH₂PO₄ pH 7.5) for 1-2 h at room temperature. Afterwards the cells were washed twice again with PBS and kept in PBS.

Final preparation of insect cell samples for TEM was kindly performed by Bettina Facompré from the Institute of Anatomy of Kiel University.

Insect cells prepared for TEM were imaged with a JEM-1400 Plus transmission electron microscope (JEOL) at 100 kV. Access to the microscope was kindly provided by the Institute of Anatomy of Kiel University. The microscope was operated by Frank Lichte.

7.14.4. Serial crystallography/X-ray diffraction experiments with *in vivo* crystals

High Five™ cells infected with ER and ER-KDEL NS2 constructs (with and without Jiv90 insert) were grown in 6-well plates for about 7 days to allow for *in vivo* crystallization. The cells were gently resuspended in the medium and sedimented by gravity in an Eppendorf reaction tube. The sedimented cells were applied onto a MicroMesh™ (700 µm diameter, 25 µm openings), quickly – but carefully – dipped into a drop of fresh ESF 921 Insect Cell Culture

Medium containing 37.5 % polyethylene glycol 400 (PEG 400) and immediately mounted into the cryo stream of the beamline. The mesh was scanned for diffraction from *in vivo* protein crystals in grids with a focused synchrotron-generated X-ray beam and an exposure time of 400 ms (or longer) per image.

The diffraction experiments were performed at beamline P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany) and supported by the work group Structural Infection Biology applying new Radiation Sources of Prof. Dr. Lars Redecke from Hamburg and Lübeck University. Technical assistance at the beamline was provided by Gleb Bourenkov from EMBL Hamburg.

7.14.5. Differential scanning fluorimetry (DSF)

Differential scanning fluorimetry (DSF) – also referred to as thermofluor assay – is a thermal shift assay (TSA) that exploits the fluorescence characteristics of a fluorescent dye like SYPRO Orange[®] to monitor protein unfolding at gradually increasing temperatures. The dye exhibits strong fluorescence when bound to hydrophobic surfaces, whereas the fluorescence is quenched by water molecules [252].

In the course of a DSF experiment, the protein sample is heated stepwise in the presence of SYPRO Orange[®] and the fluorescence of the dye is measured. With increasing temperature, the protein will start to unfold – or melt – and expose hydrophobic areas, which are usually hidden in the protein core when properly folded. The dye binds to these hydrophobic surfaces, leading to a continuous increase in fluorescence signal until all hydrophobic areas are unfolded and saturated with dye molecules. At higher temperatures, the protein will start to aggregate and the dye molecules are displaced, which will lead to a decrease in fluorescence signal. The fluorescence signal obtained in such an experiment ideally exhibits a sigmoidal shape due to the unfolding process, followed by a signal decrease arising from protein aggregation. The approximate melting temperature $T_{m(DSF)}$ of the protein can be obtained from the sigmoidal part of the fluorescence intensity curve by midpoint analysis, derivative analysis or a Boltzmann fit.

In this thesis, DSF was performed on the affinity- and SEC-purified *StreptII*-tagged DHS and His₆-tagged eIF-5A isoforms of *Sorghum bicolor*. The experiments were set up with a sample volume of 20 μ L in a 96-well plate sealed with optically clear, adhesive RT-PCR sealing foil. The SYPRO Orange[®] fluorescence was measured in the Applied Biosystem 7300 Real-Time PCR system using the TAMRA filter set. Typically, the temperature was slowly ramped up in 1 K steps from 21 to 95 °C.

Optimal protein and dye concentrations were determined for the different proteins in an initial experiment. Overall, a final concentration of approx. 0.2 mg mL⁻¹ protein and 10x SYPRO Orange[®] (supplied as 5000x concentrate in DMSO by Sigma-Aldrich) gave the best results, balancing an adequate signal-to-noise ratio with protein consumption. All further experiments were set up with these concentrations. The particular buffer compositions employed in these experiments are denoted in the context of their results..

Data analysis was performed with Frank Niesen's Microsoft Excel-based DSF Analysis script (version 3.0.2, <ftp://ftp.sgc.ox.ac.uk/pub/biophysics/>), enhanced by a script for Boltzmann fitting of the melting curves. For curves with a non-ideal sigmoidal shape,

e.g. representing a two-step unfolding mechanism, the inflection point of the curve sections were determined in their first derivative, and defined as melting point. For sigmoidal curves, a Boltzmann fit was performed to determine the melting temperature.

7.14.6. Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS; also: photon correlation spectroscopy (PCS) or quasi-elastic light scattering (QELS)) is a biophysical method to determine the apparent hydrodynamic radius of small particles and the distribution of particle sizes in a given solution or suspension. A monochromatic coherent laser beam is used to illuminate the sample and the electromagnetic wave induces an oscillating dipole moment in particles that are small in relation to the applied wavelength (typically with a diameter less than $\lambda/10$). Due to the oscillating dipole moment the particles will emanate electromagnetic waves – scattered light – of their own into all directions. This scattering mode is referred to as quasi-elastic Rayleigh scattering. The scattered light from different particles will interfere constructively or destructively, thereby generating an interference pattern. Since particles in solution or suspension move randomly due to Brownian motion, their position in relation to each other and therefore the resulting interference pattern will change over time. This constant change causes fluctuation in the overall intensity of the scattered light measured at a fixed angle.

The velocity of a particle in solution is described by its diffusion coefficient D and is influenced by the temperature and viscosity of the solution. It also inversely correlates to the particle's hydrodynamic radius. This relation is represented by the Stokes-Einstein equation:

$$d_H = \frac{k_B T}{3\pi\eta D}, \quad (7.1)$$

with the hydrodynamic radius d_H , the Boltzmann constant k_B , the absolute temperature T , the viscosity η , and the diffusion coefficient D . As a result, small particles will move faster than larger ones and their interference pattern will fluctuate more rapidly than that of larger particles.

To determine the apparent hydrodynamic radius of particles in the sample solution or suspension, the intensity of the scattered light is recorded over time. A normalized second-order autocorrelation function $g^{(2)}(q; \tau)$ is then derived from the recorded intensity signal by comparing the signal from one time point to all other time points of the measurement:

$$g^{(2)}(q; \tau) = \frac{\langle I(t)I(t + \tau) \rangle}{\langle I(t) \rangle^2}, \quad (7.2)$$

with the wave vector q (difference vector of the wave vectors of incident and scattered light), delay time τ (interval between compared time points) and intensity I .

If the interval between compared time points is very small, the signals are likely to be very similar: They correlate strongly. With increasing intervals, the correlation between signals decreases until no correlation can be detected anymore. For a monomodal, monodisperse sample, this results in an exponentially decaying autocorrelation function, from which the particle characteristics can be deduced. The slower fluctuating intensity signal of larger particles will lead to a better signal correlation over longer time intervals, which is reflected

by a later onset of the function decay compared to smaller particles. The steeper this decay, the more monodisperse the particles are.

The DLS measurements in this thesis were performed in the Zetasizer Nano S (Malvern Instruments). The sample volume was 100 μL in a UV-microcuvette (Roth) and the protein concentration in the range of 1–5 mg mL^{-1} .

DLS is not only a valuable method for the determination of particle sizes in solution or suspension, but also allows the performance of a protein melting point analysis. Therefore, the sample is heated stepwise and for each temperature step the z-average diameter (intensity weighted harmonic mean size) is calculated. When the protein unfolds, its apparent hydrodynamic radius increases noticeably. The onset of this increase is defined as the melting temperature $T_{m(\text{DLS})}$ of the protein.

To determine the $T_{m(\text{DLS})}$ of the three different *SbDHS* isoforms, 100 μL sample with a concentration of 2 mg mL^{-1} diluted in **measurement buffer** (100 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.5, 150 mM NaCl) was heated in 1 K steps from 20 to 70 $^\circ\text{C}$.

7.14.7. Size exclusion chromatography with multi-angle light scattering (SEC-MALS)

Size exclusion chromatography (SEC) followed by an online multi-angle light scattering analysis (MALS) of the eluate is a powerful biophysical method to determine the absolute molar mass and root mean square (rms) radius of molecules or molecule assemblies in solution without the need for standards. This is especially valuable for the characterization of proteins and their complexes in solution.

The protein sample to be analyzed is applied onto a size exclusion chromatography column, which separates the present molecule species based on their hydrodynamic volume in a porous column bed. If the SEC column is calibrated with a range of proteins with a known molar mass, the relative molar mass of a sample protein can be estimated from its retention time in relation to the retention times of the standard proteins on the same column. Since the molecular volume is not only depending on the molar mass, but also on the shape of the particle, this estimation only works for similarly shaped particles.

To determine the absolute molar mass and rms of the sample, the SEC-separated proteins or assemblies can be further characterized by multi-angle light scattering (MALS). The sample eluting from the SEC column is therefore illuminated with a monochromatic collimated and polarized laser beam and the particles in solution will scatter the light in all directions, similar to the DLS setup explained above. Instead of measuring the intensity fluctuation of the scattered light from a single angle, several detectors are placed at different angles, measuring the overall intensity of the scattered light under these angles. The intensity of the scattered light is directly correlated to the molecular weight M_w of the scattering molecule. From the scattering intensity measured at different angles, the root mean square (rms) radius – the molecule's size weighted by the mass distribution about its center of mass – can be calculated.

The affinity- and SEC-purified *SAS**treplI-SbDHS* and *eIF-5A* isoform samples for the analysis of their molar masses to determine their oligomeric state were diluted in **PBS** buffer (12 mM NaH_2PO_4 , 137 mM NaCl, 2.7 mM KCl, pH 7.4) to 100 μL sample volume with a concentration of 2 g L^{-1} and 1 g L^{-1} , respectively. The SEC was performed on a Superdex 200

10/30 GL SEC column (GE Healthcare) connected to an Agilent Technologies 1100 Series HPLC system (Quat Pump G1311A, Man. Inj. G1328B, COLCOM G1316A, DAD G1315B, FLD G1321A, RID G1362A, ChemStation software version B.01.03; Degasser S8515, Schambeck SFD GmbH) with PBS buffer as the mobile phase. The SEC was run at 0.5 mL min^{-1} . For the MALS analysis, a miniDAWN[®] TREOS[®] (Wyatt Technology) was applied. MALS datasets were analyzed with the ASTRA V 5.3.4.10 software (Wyatt Technology).

7.14.8. Bio-layer interferometry (BLI)

Bio-layer interferometry (BLI) is an analytical biophysical method, which enables highly sensitive measurements of the interaction of biological (macro)molecules at the surface of a sensor tip [253]. Therefore, white light is optically conducted into the sensor tip and reflected at an internal reference layer. A portion of the white light travels further through the reference layer and is reflected at the surface of the sensor tip. This second surface is made up from immobilized biological molecules (ligands) and their binding partners (analytes) from solution. The light reflected at reference layer and bio-layer exhibits a wavelength shift $\Delta\lambda$ and thereby a distinct interference pattern, which depends on the distance of the two reflecting surfaces. This distance changes, when molecules associate or dissociate from the sensor tip – resulting in a real-time variation of the measured interference pattern. Changes at the biomolecule layer are sensitively reported with this technique and allow the measurement of association and dissociation constants, binding specificity and concentrations in real-time.

The measurements are not impaired by the composition or the refractive index of the analyte solution, since only molecules binding to or dissociating from the sensor tip can influence the interference pattern significantly. In addition, the examined biomolecules do not need to be labeled – however, at least one of the binding partners has to be immobilized on the surface of the sensor tip. This can be achieved for example by highly specific antibody-antigen interactions, by affinity tags or covalent coupling via reactive amine groups.

For the experiments performed in the course of this thesis, the Octet Red96e system (PALL *forté*BIO) was applied. Access to the system was kindly provided by the work group Pharmaceutical and Medicinal Chemistry of Prof. Dr. Eric Beitz at the Pharmaceutical Institute of Kiel University. The Octet Red96e system enables measurements in a 96-well format with 8 sensor tips operated in parallel.

The measurements for the determination of *Sorghum bicolor* DHS/eIF-5A binding kinetics (see 5.5) utilized the highly specific and stable binding of the serine-alanine-*StrepII*-tag (SA*StrepII*-tag) to the mouse IgG1 antibody *StrepMAB*-Immo (iba; K_D : approx. 1 pM). The mouse IgG1 antibody itself is immobilized to the Dip and Read[™] Anti-Mouse IgG Fc Capture Biosensors (AMC; PALL *forté*BIO) by the interaction of the Fc portion of the *StrepMAB*-Immo to an anti-mouse Fc-specific antibody, which is covalently attached to the biosensor (see section 5.5 figure 5.18 for a schematic overview).

The sample volume for the all measurements was 200 μL per well in black polypropylene flat-bottom 96-well plates (Greiner Bio-One). During the association and dissociation measurements and all loading and washing steps, the plates were shaken at 1000 rpm.

The *StrepMAB*-Immo antibody was diluted to 10 mgL^{-1} in a **modified PBS buffer** (10 mM NaH_2PO_4 , 140 mM NaCl, 0.02 % (v/v) Tween[®] 20, 1 % (w/v) BSA, pH 7.5).

The purified, SA*StreptII*-tagged ligand (*SbDHS1*, -2 or -3 or *SbelF-5A1* or -2) was diluted in the respective measurement buffer – either the modified PBS buffer or a **glycine buffer** (100 mM glycine, 150 mM NaCl, 1 mM NAD⁺, pH9.0) – at a concentration of 10 mg L⁻¹ and loaded onto the immobilized *StrepMAB*-Immo antibody for 50 s.

With the completely loaded biosensors, the association and dissociation experiments with the corresponding purified analytes (*SbelF-5A1* or -2 or *SbDHS1*, -2 or -3) were performed. Therefore, the analytes were prepared in either modified PBS buffer or glycine buffer at varying concentrations and with different supplements (see section 5.5 for details of the individual experiments).

Regeneration of biosensors was performed in modified PBS buffer supplied with 5 mM NAD⁺ and 5 mM GC7. Therewith, the bound DHS analyte was effectively removed from the immobilized eIF-5A ligand (see also section 5.5.4). To prevent inactivation effects, the biosensors were washed 12x times in 12 wells with fresh PBS buffer prior to the next measurement. For about 6 regeneration cycles the biosensors yielded comparable results.

A. Appendix

A.1. Codon usage analysis of NS2 sequences

The codon usage analyses shown in figures A.1, A.2, A.3, A.4 and A.5 were performed with the web-based tool Graphical Codon Usage Analyser (GCUA) in the "each triplet position vs. usage table" mode in comparison to the codon usage of the desired expression organism. The codon usage tables of *E. coli* B, *S. frugiperda* or *N. tabacum* were obtained from Kazusa (<http://www.kazusa.or.jp/codon/>) [181, 84].

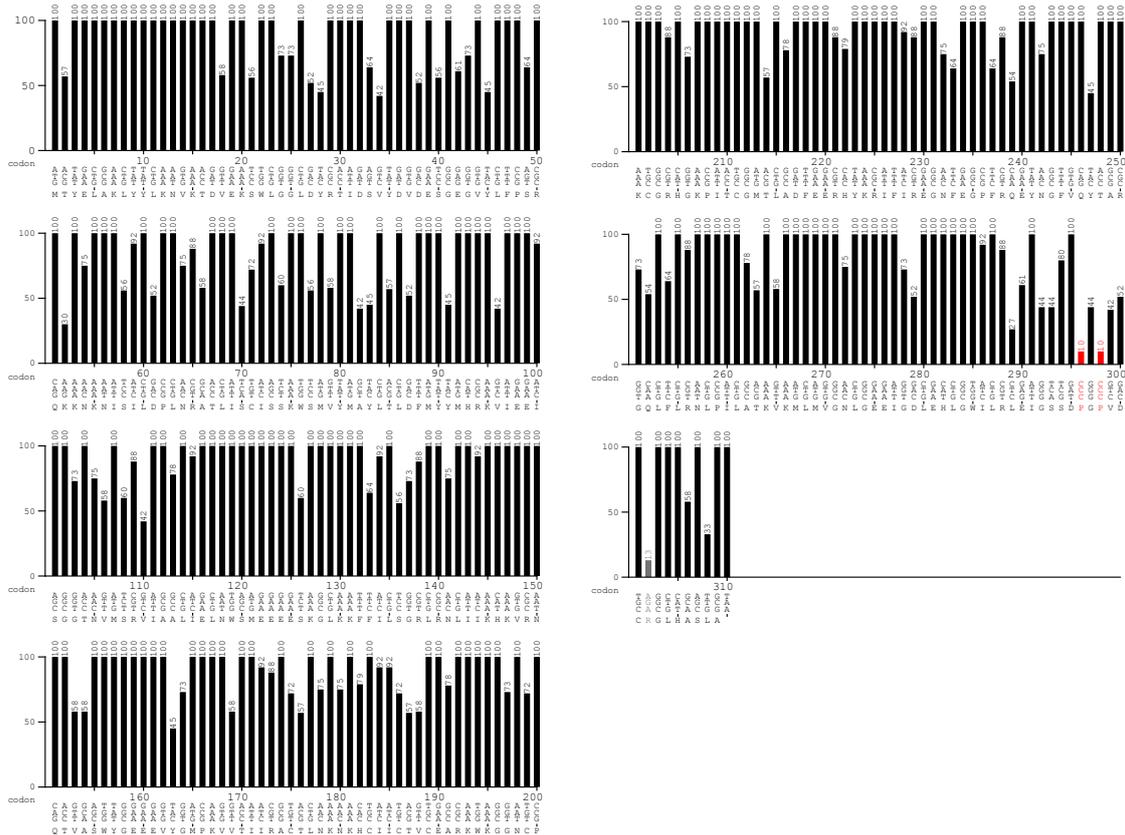


Figure A.1.: Codon usage analysis of BVDV NCP7 [1303]NS2 optimized for bacterial expression compared to *E. coli* B

The NS2 sequence was the codon-optimized sequence from construct pUC57-[1303]NS2. Ordinate: relative adaptiveness in percent; gray bars: relative adaptiveness below 20%; red bars: relative adaptiveness below 10%.

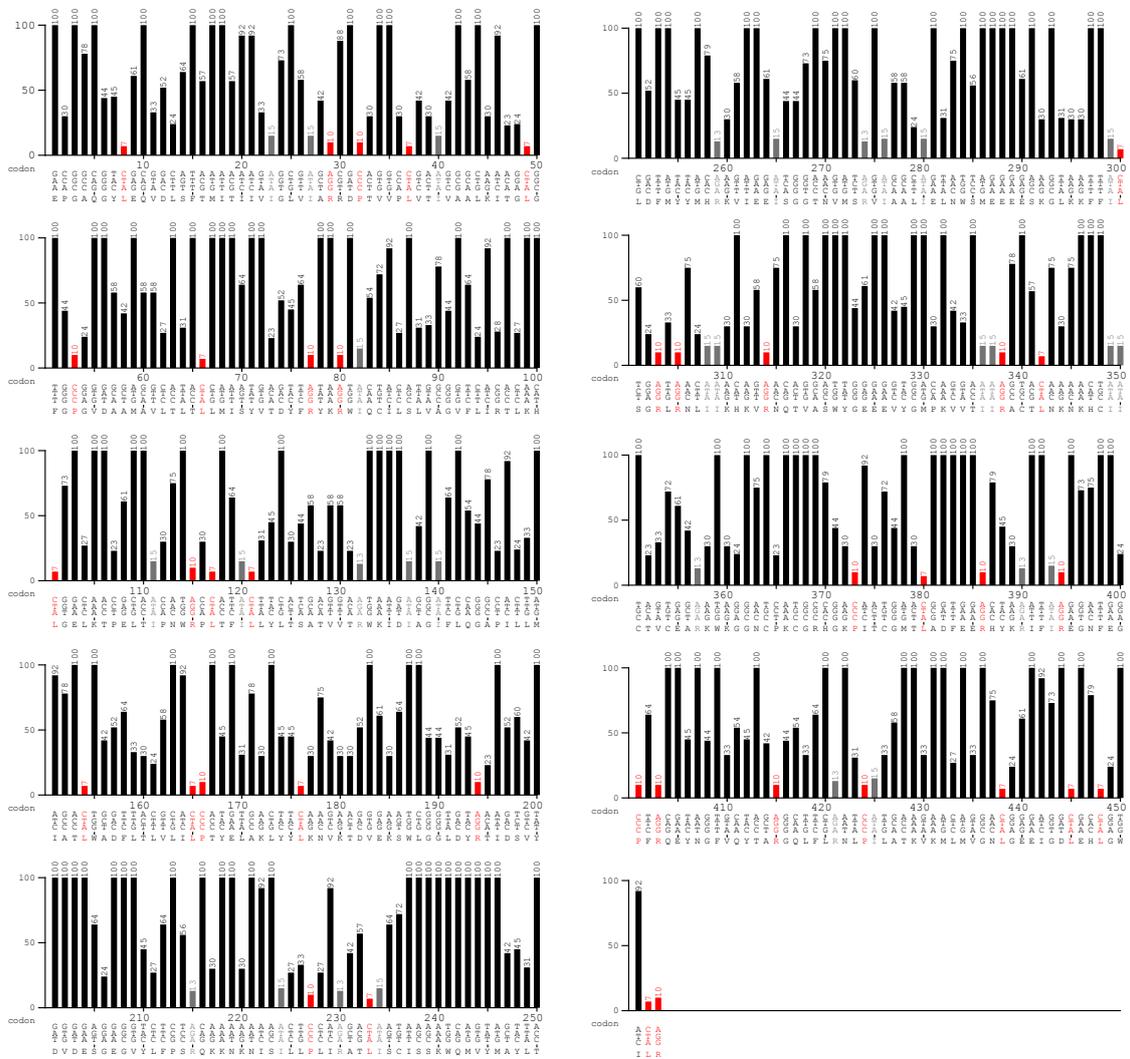


Figure A.2.: Codon usage analysis of BVDV NS2* of strain NCP7 compared to *E. coli* B

The NS2* sequence was the authentic sequence from strain NCP7 (GenBank accession no.: U63513.1). Ordinate: relative adaptiveness in percent; gray bars: relative adaptiveness below 20%; red bars: relative adaptiveness below 10%.

A.2. Amino acid sequence alignments of Jiv-containing proteins

The alignments were generated with the T-Coffee webserver (<http://tcoffee.crg.cat/>, [186, 286, 176, 65]) and displayed using `TEXshade` for `LATEX` [21].

A



B

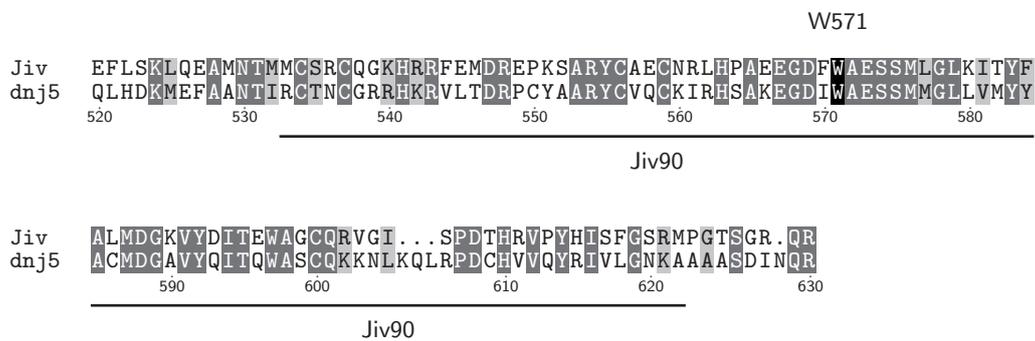


Figure A.6.: Alignment of Jiv (*Bos taurus*) and dnj-5 (*Spodoptera litura*) amino acid sequences

For the alignment the amino acid sequences of Jiv (*B. taurus*, UniProt Acc. No.: Q95J56.1) and dnj-5 (*S. litura*, NCBI Ref. Seq.: XP_022816320.1) were used. **A:** Fingerprint of the whole alignment, black lines represent identical residues. **B:** Detail of the alignment around the Jiv90 region. Dark gray residues are identical, light gray residues are similar, black residue indicates the essential W571 in Jiv90. Residue numbering according to Jiv.

A.2. Amino acid sequence alignments of Jiv-containing proteins

A



B

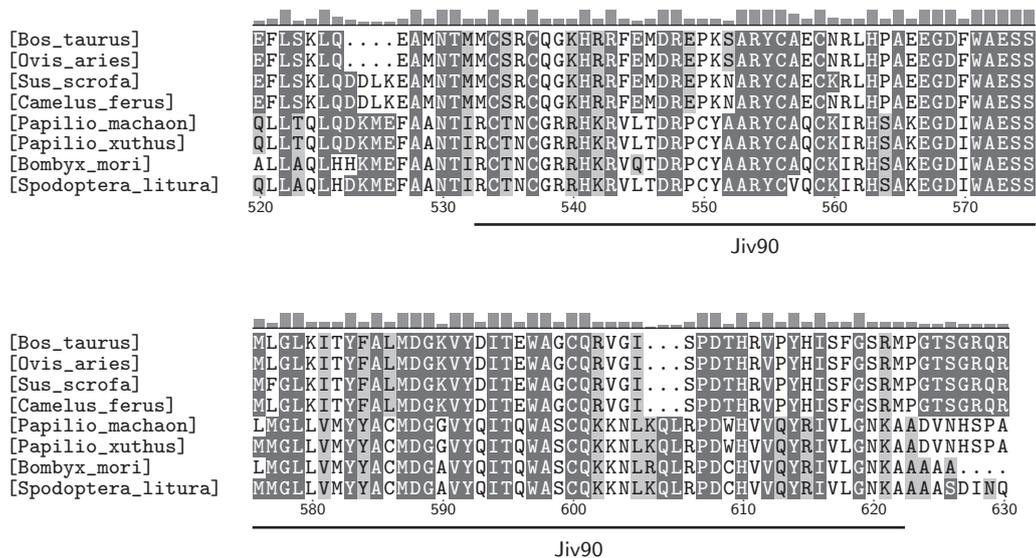


Figure A.7.: Alignment of Jiv90-containing protein sequences from *Artiodactyla* and *Lepidoptera* species

Amino acid sequences of Jiv90-containing proteins from the *Artiodactyla* *B. taurus* (DnaJ C14, UniProt Acc. No.: Q95J56.1), *O. aries* (DnaJ C14, UniProt Acc. No.: W5PN96.1), *S. scrofa* (DnaJ C14, UniProt Acc. No.: F1SPJ3.1) and *C. ferus* (uncharacterized protein, UniProt Acc. No.: T0MEX3.1), and the *Lepidoptera* *P. machaon* (DnaJ C14, UniProt Acc. No.: A0A0N1IQ13.1), *P. xuthus* (DnaJ C14, UniProt Acc. No.: A0A194Q7V4.1), *B. mori* (uncharacterized protein, UniProt Acc. No.: H9JT24.1) and *S. litura* (dnaj-5, NCBI Ref. Seq.: XP_022816320.1) were used to generate the alignment. **A:** Fingerprint of the whole alignment, black lines represent identical residues. **B:** Detail of the alignment around the Jiv90 region. Dark gray residues are conserved in more than 50% of the sequences, light gray residues are similar. Residue numbering according to Jiv. Residue conservation is depicted as bar graph above the alignment.

A.3. TEM image of an insect cell infected with *Baculovirus heliothis*

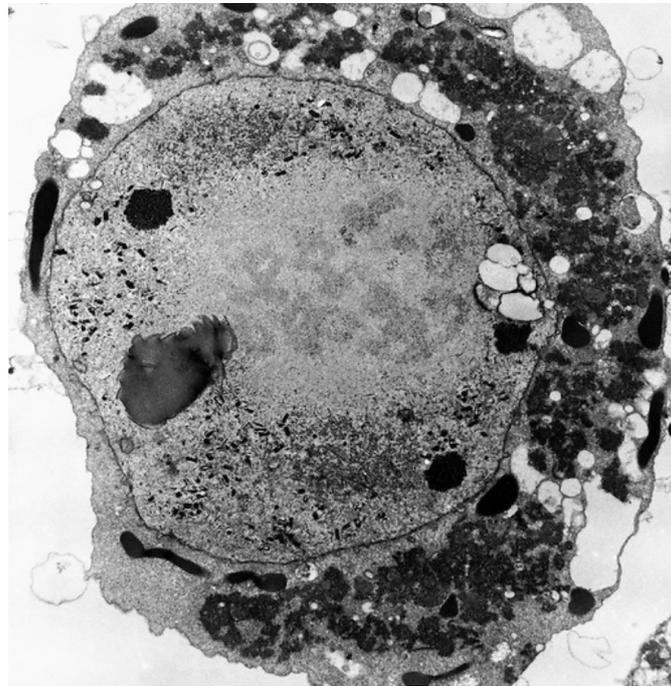


Figure A.8.: TEM image of an insect cell infected with *Baculovirus heliothis*

The infected insect cell shows baculovirus cross-sections and "furry" structures of putatively baculoviral origin in the cell's nucleus. Copyright: Dr. David Phillips.

A.4. Codon usage analysis of *SbDHS* and *eIF-5A* sequences

The codon usage analyses shown in figures A.11, A.10 and A.9 were performed with the web-based tool Graphical Codon Usage Analyser (GCUA) in the "each triplet position vs. usage table" mode in comparison to the codon usage of the expression organism. The codon usage tables of *E. coli* B was obtained from Kazusa (<http://www.kazusa.or.jp/codon/>) [181, 84].

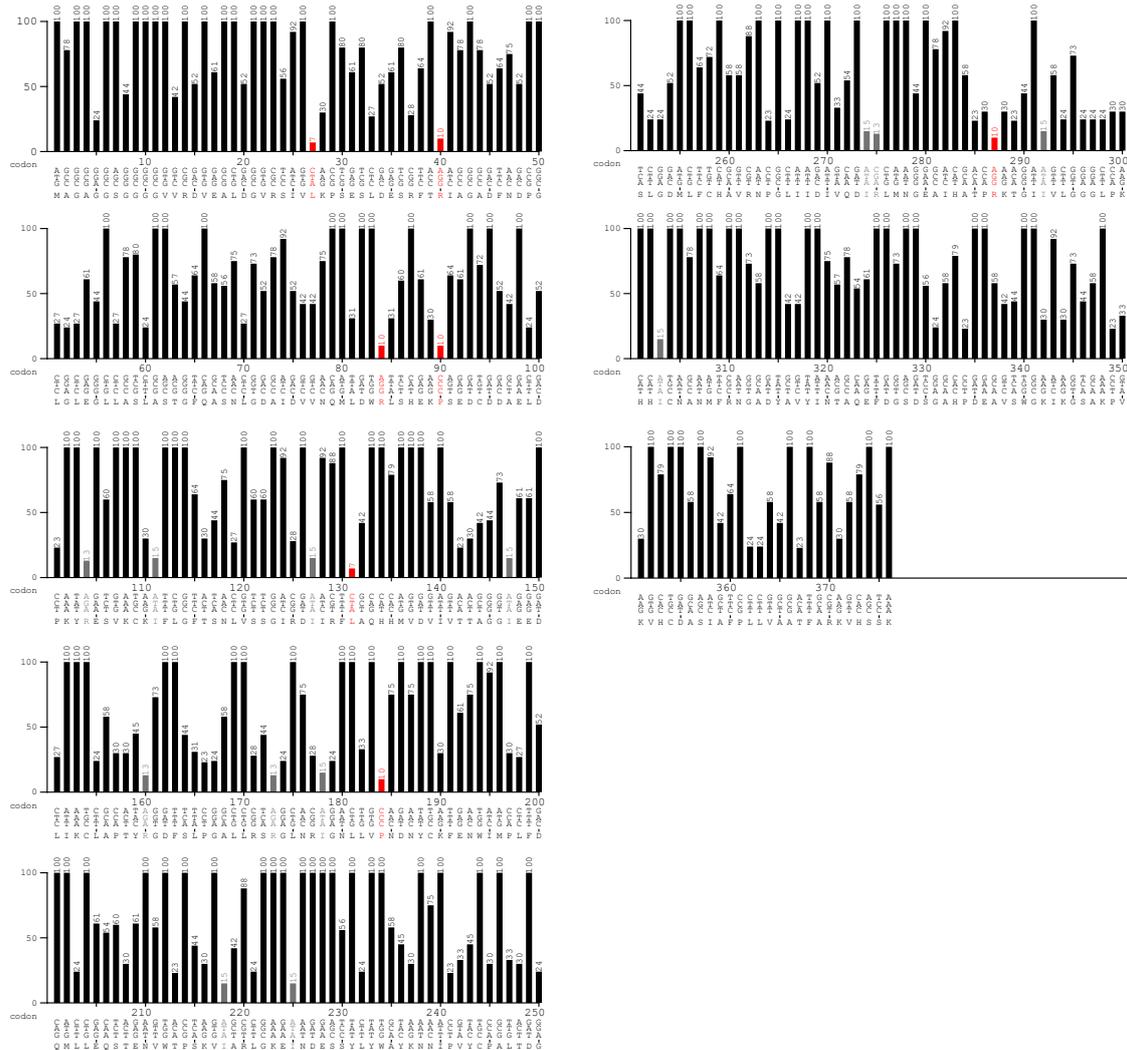


Figure A.9.: Codon usage analysis of *Sorghum bicolor* DHS1 compared to *E. coli* B

Ordinate: relative adaptiveness in percent; gray bars: relative adaptiveness below 20 %; red bars: relative adaptiveness below 10 %.

A.4. Codon usage analysis of *SbDHS* and *eIF-5A* sequences

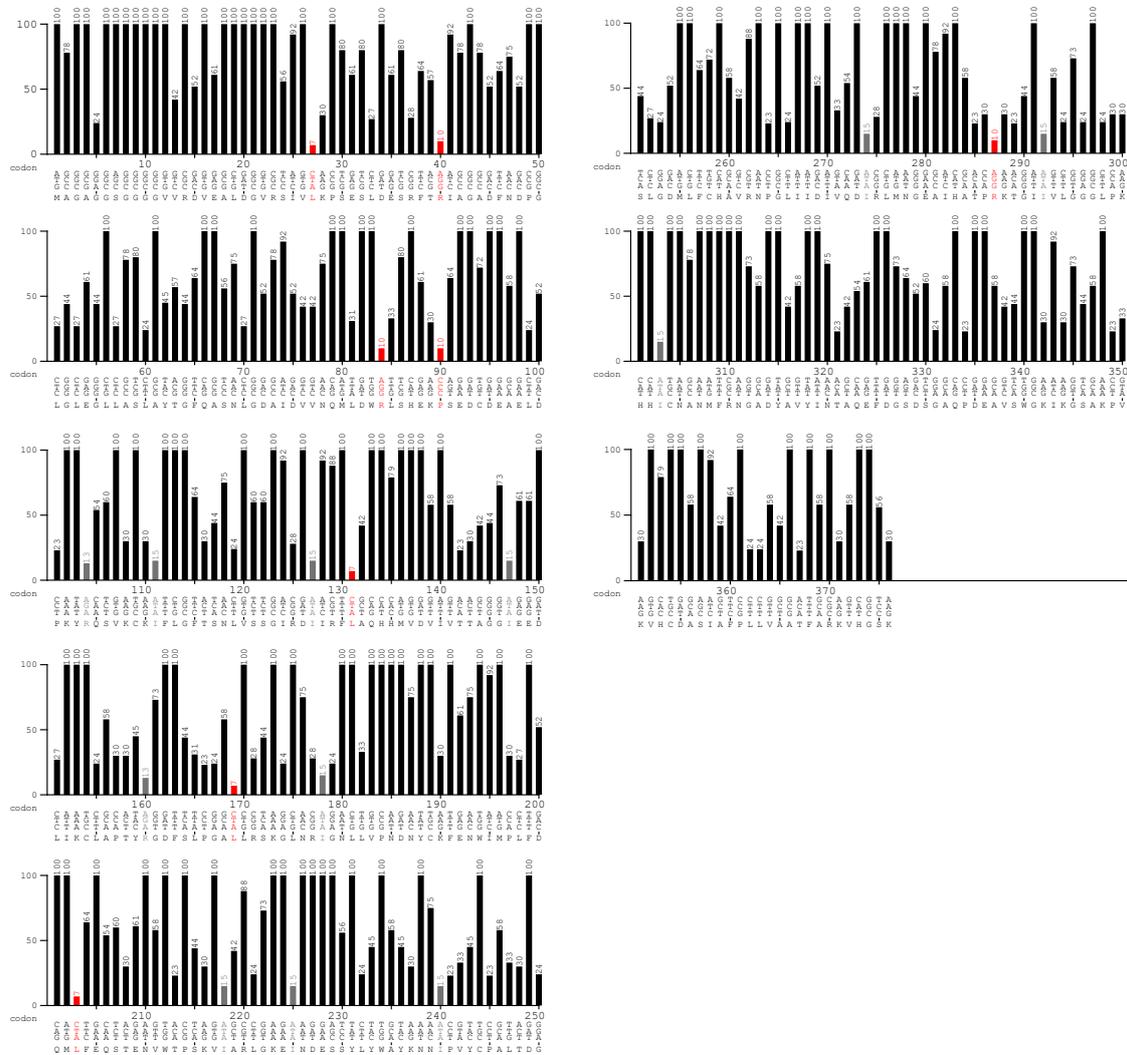


Figure A.10.: Codon usage analysis of *Sorghum bicolor* DHS2 compared to *E. coli* B

Ordinate: relative adaptiveness in percent; gray bars: relative adaptiveness below 20%; red bars: relative adaptiveness below 10%.

A.4. Codon usage analysis of *SbDHS* and *eIF-5A* sequences

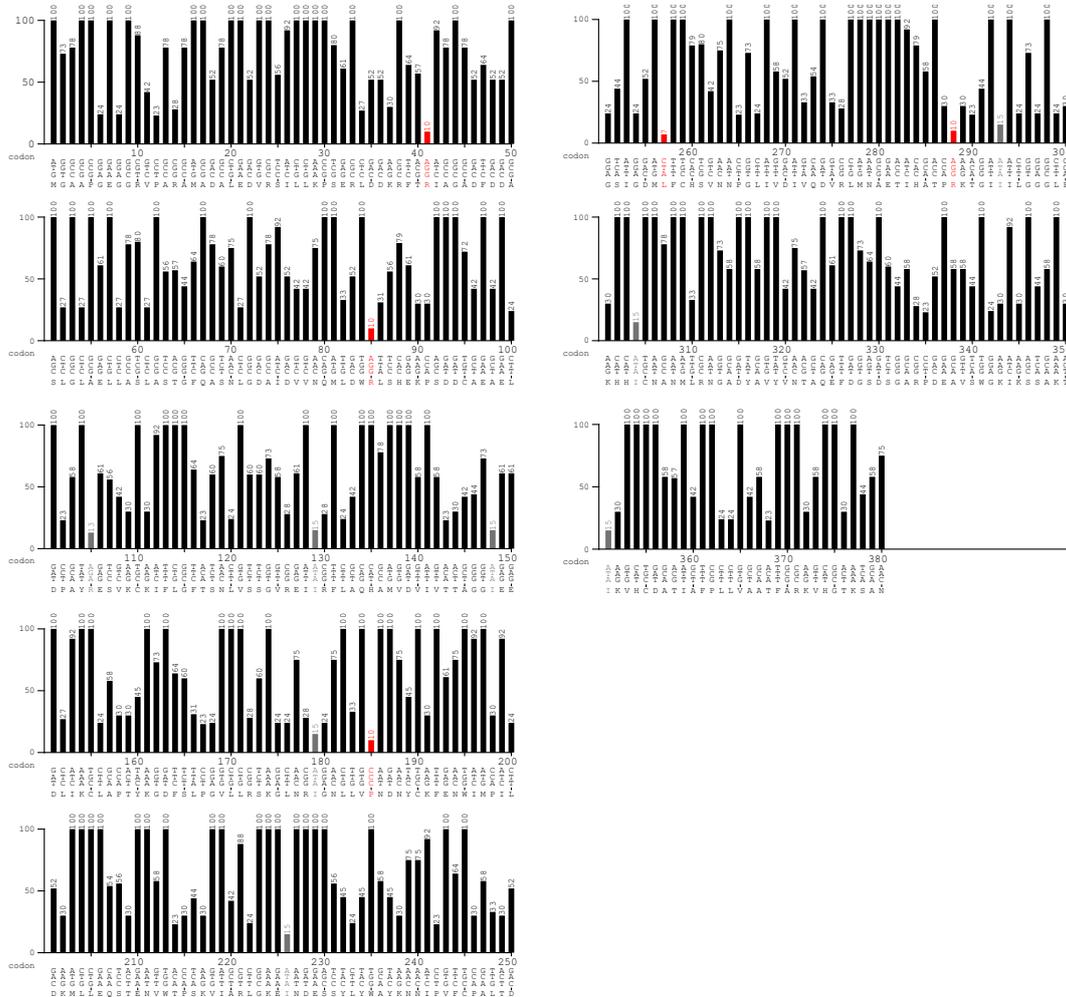


Figure A.11.: Codon usage analysis of *Sorghum bicolor* DHS3 compared to *E. coli* B

Ordinate: relative adaptiveness in percent; gray bars: relative adaptiveness below 20%; red bars: relative adaptiveness below 10%.

A.4. Codon usage analysis of *SbDHS* and *eIF-5A* sequences

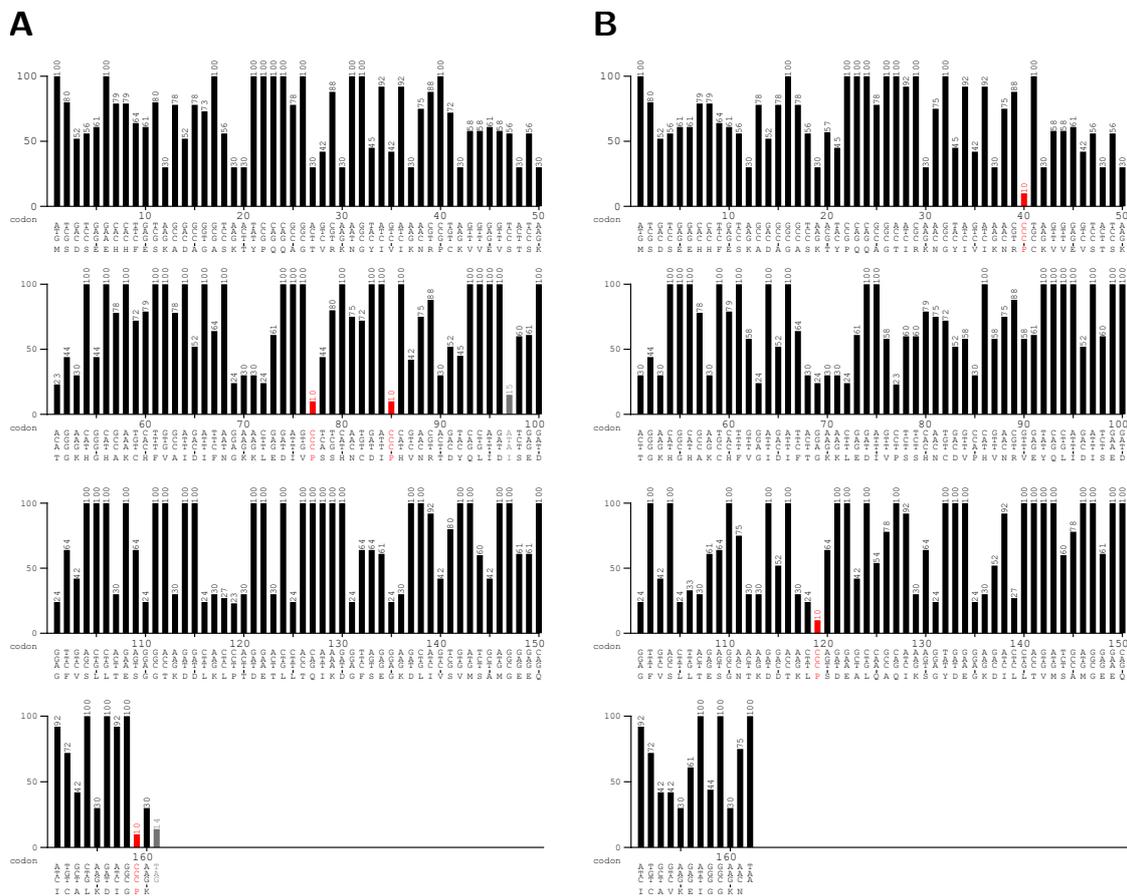


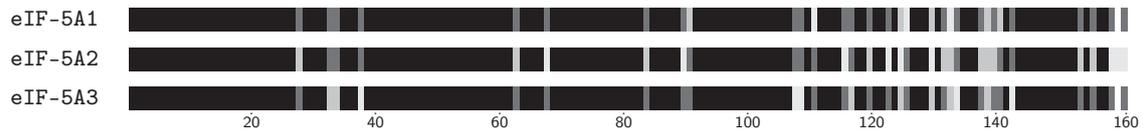
Figure A.12.: Codon usage analysis of *Sorghum bicolor* eIF-5A1 and eIF-5A2 compared to *E. coli* B

A: *SbIF-5A1*. B: *SbIF-5A1*. Ordinate: relative adaptiveness in percent; gray bars: relative adaptiveness below 20%; red bars: relative adaptiveness below 10%.

A.5. Alignments of e/aIF-5A and DHS homo- and paralogs

The alignments were generated with the T-Coffee webserver (<http://tcoffee.crg.cat/>, [186, 286, 176, 65]) and displayed using `TEXshade` for `LATEX` [21].

A



B

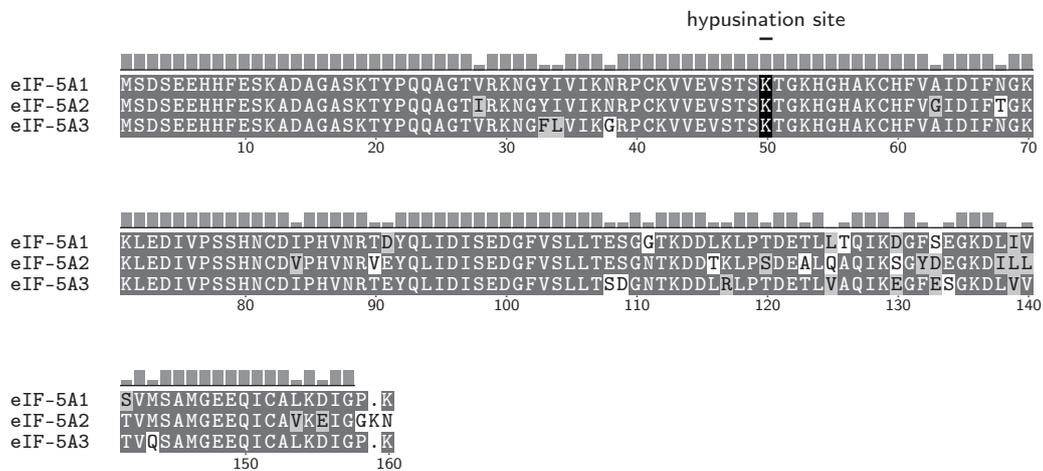


Figure A.13.: Alignment of the amino acid sequences of initiation factor 5A variants from *Sorghum bicolor*

For the alignment of eIF-5A variants from *Sorghum bicolor*, the amino acid sequences of eIF-5A1 (GenBank Acc. No.: KXG34270.1), eIF-5A2 (GenBank Acc. No.: KXG37504.1) and eIF-5A3 (GenBank Acc. No.: KXG36613.1) were used. **A:** Fingerprint of the whole alignment. Black regions represent identical residues, dark gray regions are conserved and medium gray regions are similar in more than 50% of the aligned sequences, light gray areas are not conserved. **B:** Detailed alignment. The conserved hypusinated lysine residue is shaded in black. Dark gray residues are conserved, light gray residues are similar in more than 50% of the aligned sequences. Residue numbering according to human *Sb*eIF-5A1. Residue conservation is depicted as bar graph above the alignment.

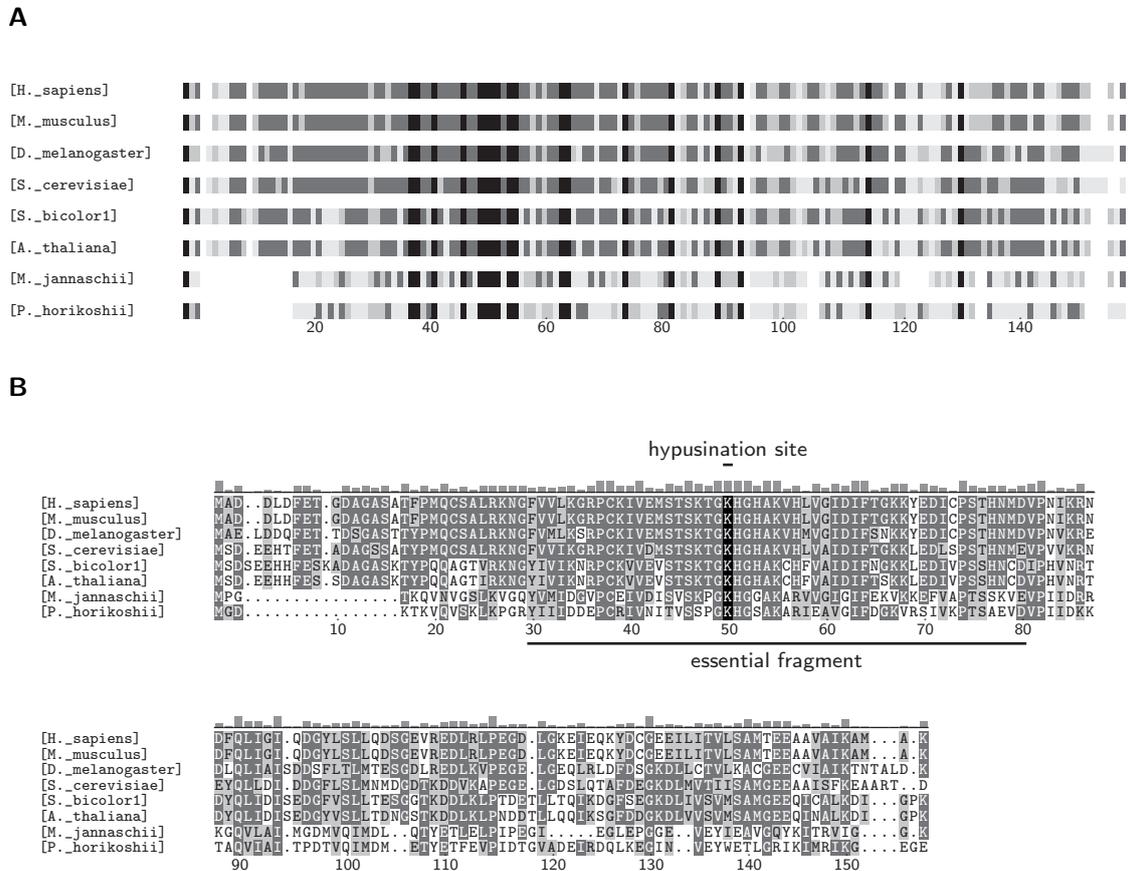


Figure A.14.: Alignment of the amino acid sequences of initiation factor 5A homologs from eukaryotes and archae

For the alignment of e/aIF-5A homologs, the amino acid sequences of eIF-5A-1 from *Homo sapiens* (UniProt Acc. No.: P63241.2), *Mus musculus* (NCBI Ref. Seq.: NP_001160066.1) and *Arabidopsis thaliana* (GenBank Acc. No.: AEE29088.1), eIF-5A-A from *Drosophila melanogaster* (NCBI Ref. Seq.: NP_611878.1), eIF-5A-like 1 from *Sorghum bicolor* (NCBI Ref. Seq.: XP_021306769.1), Hyp2p from *Saccharomyces cerevisiae* (GenBank Acc. No.: AJU52647.1) and aIF-5A from *Methanocaldococcus jannaschii* and *Pyrococcus horikoshii* (UniProt Acc. Nos.: Q58625.2 and O50089.1) were used. **A:** Fingerprint of the whole alignment. Black regions represent identical residues, dark gray regions are conserved and medium gray regions are similar in more than 50% of the aligned sequences, light gray areas are not conserved. **B:** Detailed alignment. The conserved hypusinated lysine residue is shaded in black. Dark gray residues are conserved, light gray residues are similar in more than 50% of the aligned sequences. Residue numbering according to human eIF-5A-1. Residue conservation is depicted as bar graph above the alignment.

A.5. Alignments of e/aIF-5A and DHS homo- and paralogs

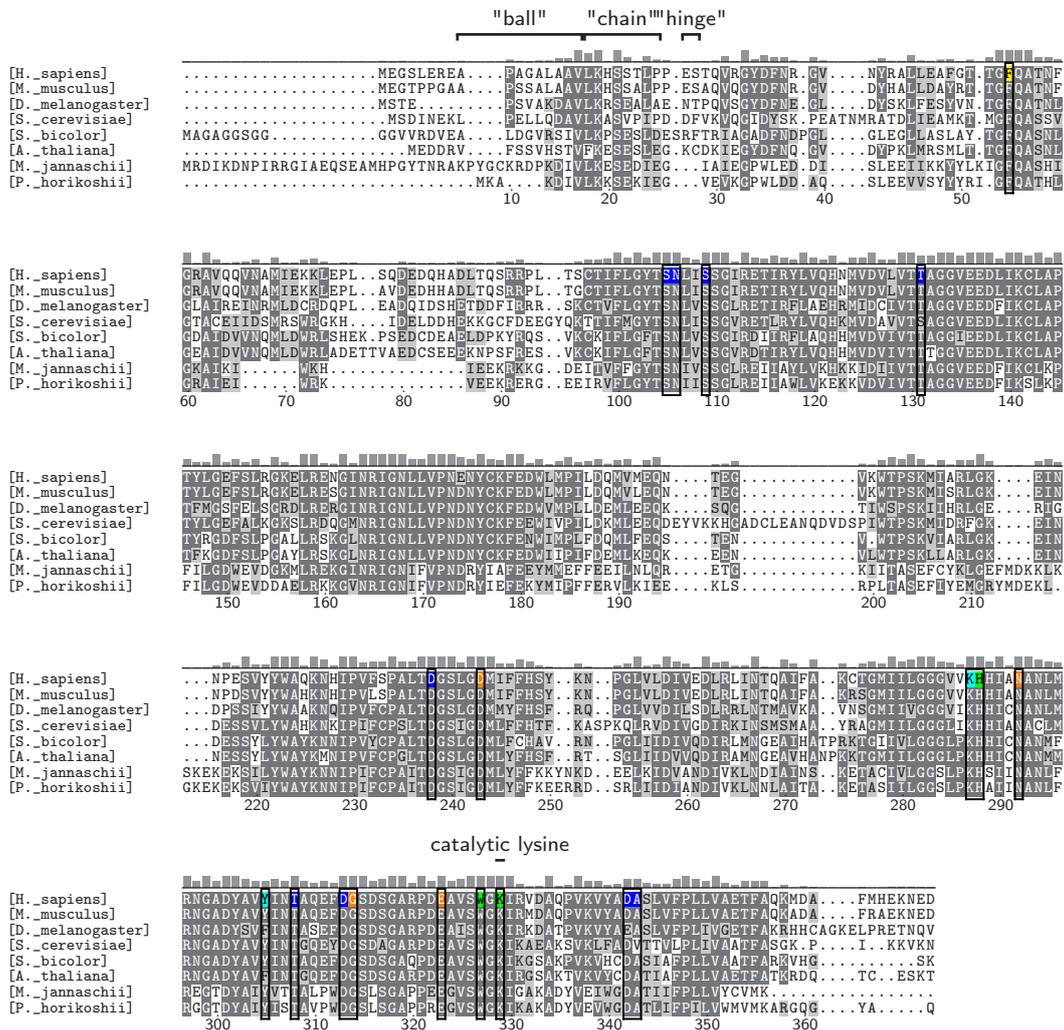


Figure A.15.: Alignment of the amino acid sequences of DHS homologs from eukaryotes and archae

For the alignment of DHS homologs, the amino acid sequences of DHS from *Homo sapiens* (PDB Acc. No.: 1RQD), *Mus musculus* (UniProt Acc. No.: Q3TXU5.2), *Drosophila melanogaster* (UniProt Acc. No.: Q9VSF4.2), *Saccharomyces cerevisiae* (UniProt Acc. No.: P38791.1), *Sorghum bicolor* (DHS2, isolate from the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober at the Botanical Institute, Kiel University), *Arabidopsis thaliana* (UniProt Acc. No.: Q9FI94.1), *Methanocaldococcus jannaschii* (GenBank Acc. No.: AAB98813.1) and *Pyrococcus horikoshii* (PDB Acc. No.: 4P63) were used. □: non-matching residues; ■: similar residues in >50% of the aligned sequences; ■: conserved residues. Specific residues are further shaded based on structural information from human DHS (PDB Acc. No.: 1RQD) as follows: ■: involved in subunit communication; ■: direct interaction with NAD; ■: direct interaction with GC7; ■: side pocket; ■: involved in catalysis. Residue numbering according to human DHS. Residue conservation is depicted as bar graph above the alignment.

A.5. Alignments of e/aIF-5A and DHS homo- and paralogues

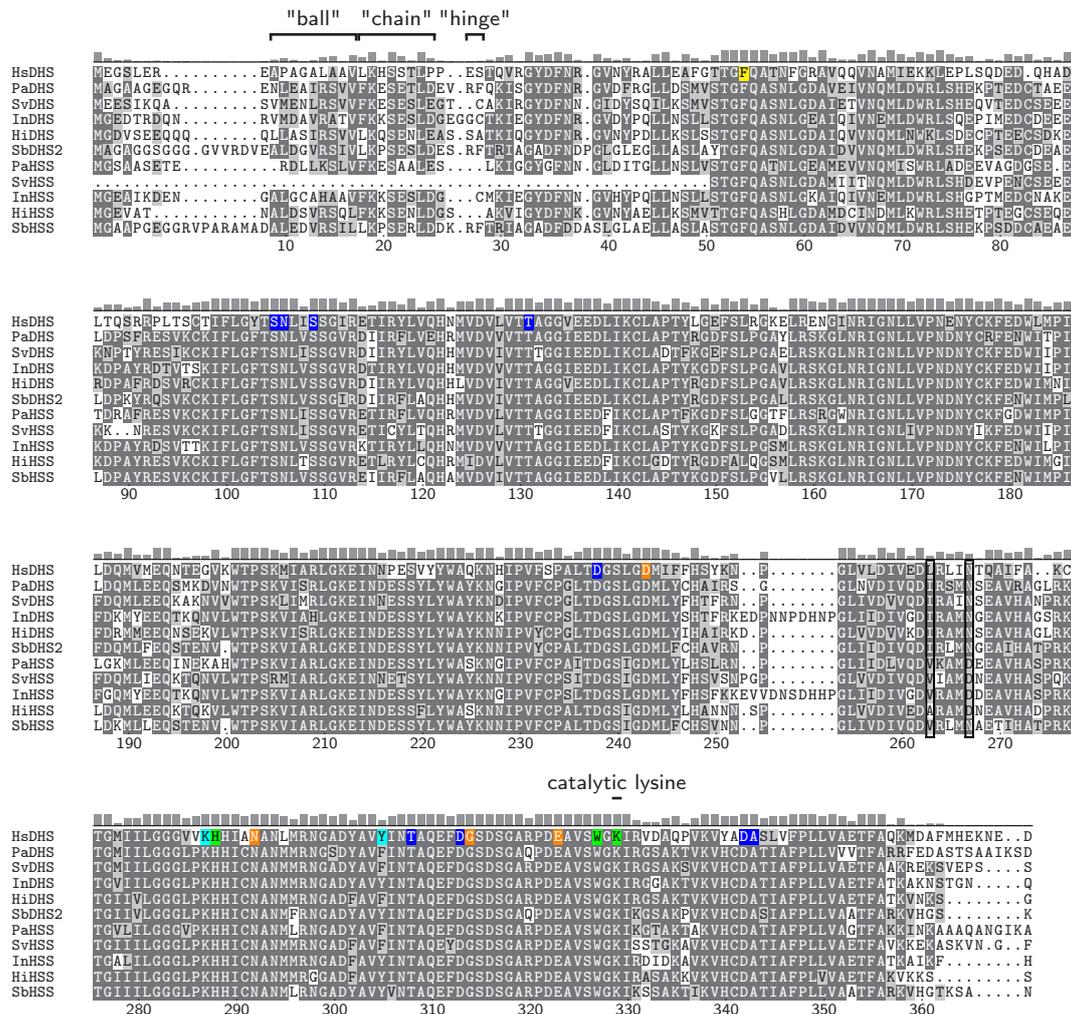


Figure A.16.: Alignment of the amino acid sequences of angiosperm DHS and HSS

For the alignment of angiosperm DHS homologs and their HSS paralogues, the amino acid sequences of DHS and HSS from the Orchid *Phalaenopsis amabilis* (GenBank Acc. Nos.: CAR31338.1 and CAL69908.1), the Asteraceae *Senecio vernalis* (GenBank Acc. No.: CAD21436.1; UniProt Acc. No.: Q9SC13.1), the Convolvaceae *Ipomoea neei* (GenBank Acc. Nos.: CCU83297.1, CCU83287.1), the Boraginaceae *Heliotropium indicum* (GenBank Acc. Nos.: CAG28993.1, CAG28992.1) and the Poaceae *Sorghum bicolor* (HSS and DHS2 isolates from the research group Biochemical Ecology and Molecular Evolution of Prof. Dr. Dietrich Ober at the Botanical Institute, Kiel University) were used. The sequences were chosen based on Livshultz et al. 2018 [155]. To allow comparisons to other DHS alignments in this thesis, the DHS of *Homo sapiens* (PDB Acc. No.: 1RQD) is included in the alignment. □: non-matching residues; ■: similar residues in >50% of the aligned sequences; ■: conserved residues. Specific residues are further shaded based on structural information from human DHS (PDB Acc. No.: 1RQD) as follows: ■: involved in subunit communication; ■: direct interaction with NAD; ■: direct interaction with GC7; ■: side pocket; ■: involved in catalysis. Residue numbering according to human DHS. Residue conservation is depicted as bar graph above the alignment. Boxed residues emphasize putative telltale residues for DHS or HSS functionality [120, 155].

Bibliography

- [1] A. Abbruzzese, M. H. Park, and J. E. Folk. "Deoxyhypusine hydroxylase from rat testis. Partial purification and characterization." In: *The Journal of biological chemistry* 261.7 (1986), pp. 3085–3089.
- [2] H. Agaisse and D. Lereclus. "How does *Bacillus thuringiensis* produce so much insecticidal crystal protein?" In: *Journal of bacteriology* 177.21 (1995), pp. 6027–6032.
- [3] E. V. Agapov et al. "Uncleaved NS2-3 is required for production of infectious bovine viral diarrhea virus." In: *Journal of Virology* 78.5 (2004), pp. 2414–2425.
- [4] M. Aksu, S. Trakhanov, and D. Görlich. "Structure of the exportin Xpo4 in complex with RanGTP and the hypusine-containing translation factor eIF5A." In: *Nature communications* 7 (2016), p. 11952.
- [5] Al-Amoudi et al. "Cryo-electron microscopy of vitreous sections." In: *The EMBO journal* 23.18 (2004), pp. 3583–3588.
- [6] S. Anke et al. "Polyphyletic origin of pyrrolizidine alkaloids within the Asteraceae. Evidence from differential tissue expression of homospermidine synthase." In: *Plant physiology* 136.4 (2004), pp. 4037–4047.
- [7] S. Anke et al. "Pyrrolizidine alkaloid biosynthesis in *Phalaenopsis* orchids: Developmental expression of alkaloid-specific homospermidine synthase in root tips and young flower buds." In: *Plant physiology* 148.2 (2008), pp. 751–760.
- [8] H. J. Arnett and K. M. Smith. "An ultrastructural study of the development of a granulosis virus in the cells of the moth *Plodia interpunctella* (Hbn.)." In: *Journal of ultrastructure research* 21.3 (1967), pp. 251–268.
- [9] H. Ashkenazy et al. "ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids." In: *Nucleic acids research* 38.Web Server issue (2010), W529–33.
- [10] H. Ashkenazy et al. "ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules." In: *Nucleic acids research* 44.W1 (2016), W344–50.
- [11] J. G. Baillon et al. "Fluorinated analogues of spermidine as substrates of spermine synthase." In: *European journal of biochemistry* 176.2 (1988), pp. 237–242.
- [12] C. R. Baker, V. Hanson-Smith, and A. D. Johnson. "Following gene duplication, paralog interference constrains transcriptional circuit evolution." In: *Science (New York, N.Y.)* 342.6154 (2013), pp. 104–108.
- [13] J. C. Baker. "Bovine viral diarrhea virus: A review." In: *Journal of the American Veterinary Medical Association* 190.11 (1987), pp. 1449–1458.
- [14] D. Baltimore. "Expression of animal virus genomes." In: *Bacteriological reviews* 35.3 (1971), pp. 235–241.
- [15] M. Bannwarth and G. E. Schulz. "The expression of outer membrane proteins for crystallization." In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1610.1 (2003), pp. 37–45.
- [16] D. Bartig and F. Klink. "Determination of the unusual amino acid hypusine at the lower picomole level by derivatization with 4-dimethylaminoazobenzene-4'-sulphonyl chloride and reversed-phase high-performance or medium-pressure liquid chromatography." In: *Journal of chromatography* 606.1 (1992), pp. 43–48.
- [17] D. Bartig, H. Schümann, and F. Klink. "The Unique Posttranslational Modification Leading to Deoxyhypusine or Hypusine is a General Feature of the Archaeobacterial Kingdom." In: *Systematic and Applied Microbiology* 13.2 (1990), pp. 112–116.
- [18] P. Becher et al. "Genetic and antigenic characterization of novel pestivirus genotypes: Implications for classification." In: *Virology* 311.1 (2003), pp. 96–104.
- [19] P. Becher et al. "Genetic diversity of pestiviruses: Identification of novel groups and implications for classification." In: *Virology* 262.1 (1999), pp. 64–71.
- [20] S. E. Behrens et al. "Characterization of an autonomous subgenomic pestivirus RNA replicon." In: *Journal of Virology* 72.3 (1998), pp. 2364–2372.
- [21] E. Beitz. "TEXshade: Shading and labeling of multiple sequence alignments using LATEX2 epsilon." In: *Bioinformatics (Oxford, England)* 16.2 (2000), pp. 135–139.

- [22] E. Beitz. "TEXtopo: Shaded membrane protein topology plots in LATEX2." In: *Bioinformatics (Oxford, England)* 16.11 (2000), pp. 1050–1051.
- [23] R. Benne, M. L. Brown-Luedi, and J. W. Hershey. "Purification and characterization of protein synthesis initiation factors eIF-1, eIF-4C, eIF-4D, and eIF-5 from rabbit reticulocytes." In: *The Journal of biological chemistry* 253.9 (1978), pp. 3070–3077.
- [24] R. Benne and J. W. Hershey. "The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes." In: *The Journal of biological chemistry* 253.9 (1978), pp. 3078–3087.
- [25] I. Berger et al. "The multiBac protein complex production platform at the EMBL." In: *Journal of visualized experiments : JoVE* 77 (2013), e50159.
- [26] H. E. Bergmans, I. M. van Die, and W. P. Hoekstra. "Transformation in *Escherichia coli*: Stages in the process." In: *Journal of bacteriology* 146.2 (1981), pp. 564–570.
- [27] J. C. Bermak et al. "Regulation of transport of the dopamine D1 receptor by a new membrane-associated ER protein." In: *Nature cell biology* 3.5 (2001), pp. 492–498.
- [28] D. Bevec et al. "Inhibition of HIV-1 Replication in Lymphocytes by Mutants of the Rev Cofactor eIF-5A." In: *Science* 271.5257 (1996), pp. 1858–1860.
- [29] C. Bieniossek, T. J. Richmond, and I. Berger. "MultiBac: Multigene baculovirus-based eukaryotic protein complex production." In: *Current protocols in protein science* Chapter 5 (2008), Unit 5.20.
- [30] I. Bintintan and G. Meyers. "A new type of signal peptidase cleavage site identified in an RNA virus polyprotein." In: *The Journal of biological chemistry* 285.12 (2010), pp. 8572–8584.
- [31] A. V. Birk et al. "Cytoplasmic vacuolization responses to cytopathic bovine viral diarrhoea virus." In: *Virus research* 132.1-2 (2008), pp. 76–85.
- [32] H. C. Birnboim and J. Doly. "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." In: *Nucleic acids research* 7.6 (1979), pp. 1513–1523.
- [33] S. R. Bond and C. C. Naus. "RF-Cloning.org: An online tool for the design of restriction-free cloning projects." In: *Nucleic acids research* 40.Web Server issue (2012), W209–13.
- [34] M. Boudes et al. "A pipeline for structure determination of in vivo-grown crystals using in cellulo diffraction." In: *Acta crystallographica. Section D, Structural biology* 72.Pt 4 (2016), pp. 576–585.
- [35] M. M. Bradford. "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." In: *Analytical biochemistry* 72 (1976), pp. 248–254.
- [36] A. K. Brödel, A. Sonnabend, and S. Kubick. "Cell-free protein expression based on extracts from CHO cells." In: *Biotechnology and bioengineering* 111.1 (2014), pp. 25–36.
- [37] J. Broecker, B. T. Eger, and O. P. Ernst. "Crystallogensis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs." In: *Structure* 25.2 (2017), pp. 384–392.
- [38] J. Brownlie. "The pathogenesis of bovine virus diarrhoea virus infections." In: *Revue scientifique et technique (International Office of Epizootics)* 9.1 (1990), pp. 43–59.
- [39] J. Brownlie et al. "Pathogenesis and epidemiology of bovine virus diarrhoea virus infection of cattle." In: *Annales de recherches vétérinaires. Annals of veterinary research* 18.2 (1987), pp. 157–166.
- [40] V. E. Buckwold, B. E. Beer, and R. O. Donis. "Bovine viral diarrhoea virus as a surrogate model of hepatitis C virus for the evaluation of antiviral agents." In: *Antiviral research* 60.1 (2003), pp. 1–15.
- [41] T. J. Bullwinkle et al. "(R)- β -lysine-modified elongation factor P functions in translation elongation." In: *The Journal of biological chemistry* 288.6 (2013), pp. 4416–4423.
- [42] M. Buntru et al. "Tobacco BY-2 cell-free lysate: An alternative and highly-productive plant-based in vitro translation system." In: *BMC biotechnology* 14 (2014), p. 37.
- [43] W. N. Burnette. "Western blotting: Electrophoretic transfer of proteins from sodium dodecyl sulfate polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A." In: *Analytical biochemistry* 112.2 (1981), pp. 195–203.
- [44] T. R. Butt et al. "SUMO fusion technology for difficult-to-express proteins." In: *Protein expression and purification* 43.1 (2005), pp. 1–9.
- [45] S. Cabantous et al. "New molecular reporters for rapid protein folding assays." In: *PLoS one* 3.6 (2008), e2387.

- [46] M. Caffrey. "A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes." In: *Acta crystallographica. Section F, Structural biology communications* 71.Pt 1 (2015), pp. 3–18.
- [47] N. Callens et al. "Morphology and Molecular Composition of Purified Bovine Viral Diarrhea Virus Envelope." In: *PLoS pathogens* 12.3 (2016), e1005476.
- [48] V. S. P. Cano et al. "Mutational analyses of human eIF5A-1—identification of amino acid residues critical for eIF5A activity and hypusine modification." In: *The FEBS journal* 275.1 (2008), pp. 44–58.
- [49] M. Carroni and H. R. Saibil. "Cryo electron microscopy to determine the structure of macromolecular complexes." In: *Methods (San Diego, Calif.)* 95 (2016), pp. 78–85.
- [50] G. Celniker et al. "ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function." In: *Israel Journal of Chemistry* 53.3-4 (2013), pp. 199–206.
- [51] D. Chamot and C. Kuhlemeier. "Differential expression of genes encoding the hypusine-containing translation initiation factor, eIF-5A, in tobacco." In: *Nucleic acids research* 20.4 (1992), pp. 665–669.
- [52] K. Y. Chen. "An 18000-dalton protein metabolically labeled by polyamines in various mammalian cell lines." In: *Biochimica et biophysica acta* 756.3 (1983), pp. 395–402.
- [53] K. Y. Chen and Q. P. Dou. "NAD⁺ stimulated the spermidine-dependent hypusine formation on the 18 kDa protein in cytosolic lysates derived from NB-15 mouse neuroblastoma cells." In: *FEBS letters* 229.2 (1988), pp. 325–328.
- [54] Y. Chen and H. P. Erickson. "FtsZ filament dynamics at steady state: Subunit exchange with and without nucleotide hydrolysis." In: *Biochemistry* 48.28 (2009), pp. 6664–6673.
- [55] S. I. Chung et al. "Eukaryotic initiation factor 5A: The molecular form of the hypusine-containing protein from human erythrocytes." In: *Biochimica et biophysica acta* 1076.3 (1991), pp. 448–451.
- [56] P. M. J. Clement et al. "Identification and characterization of eukaryotic initiation factor 5A-2." In: *European journal of biochemistry* 270.21 (2003), pp. 4254–4263.
- [57] M. S. Collett et al. "Proteins encoded by bovine viral diarrhea virus: The genomic organization of a pestivirus." In: *Virology* 165.1 (1988), pp. 200–208.
- [58] P. M. Colman, E. Suzuki, and A. van Donkelaar. "The structure of cucurbitin: Subunit symmetry and organization in situ." In: *European journal of biochemistry* 103.3 (1980), pp. 585–588.
- [59] H. L. Cooper, M. H. Park, and J. E. Folk. "Posttranslational formation of hypusine in a single major protein occurs generally in growing cells and is associated with activation of lymphocyte growth." In: *Cell* 29.3 (1982), pp. 791–797.
- [60] H. L. Cooper et al. "Identification of the hypusine-containing protein hy⁺ as translation initiation factor eIF-4D." In: *Proceedings of the National Academy of Sciences of the United States of America* 80.7 (1983), pp. 1854–1857.
- [61] W. V. Corapi, R. O. Donis, and E. J. Dubovi. "Monoclonal antibody analyses of cytopathic and noncytopathic viruses from fatal bovine viral diarrhea virus infections." In: *Journal of Virology* 62.8 (1988), pp. 2823–2827.
- [62] M.-P. Courageot, A. Catteau, and P. Desprès. "Mechanisms of dengue virus-induced cell death." In: *Advances in virus research* 60 (2003), pp. 157–186.
- [63] R. Deng and K. V. Brock. "Molecular cloning and nucleotide sequence of a pestivirus genome, noncytopathic bovine viral diarrhea virus strain SD-1." In: *Virology* 191.2 (1992), pp. 867–869.
- [64] M. L. Di Martino et al. "Polyamines: Emerging players in bacteria-host interactions." In: *International journal of medical microbiology : IJMM* 303.8 (2013), pp. 484–491.
- [65] P. Di Tommaso et al. "T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension." In: *Nucleic acids research* 39.Web Server issue (2011), W13–7.
- [66] C. A. O. Dias et al. "eIF5A dimerizes not only in vitro but also in vivo and its molecular envelope is similar to the EF-P monomer." In: *Amino acids* 44.2 (2013), pp. 631–644.
- [67] C. A. O. Dias et al. "Structural modeling and mutational analysis of yeast eukaryotic translation initiation factor 5A reveal new critical residues and reinforce its involvement in protein synthesis." In: *The FEBS journal* 275.8 (2008), pp. 1874–1888.
- [68] G. Dodson and D. Steiner. "The role of assembly in insulin's biosynthesis." In: *Current opinion in structural biology* 8.2 (1998), pp. 189–194.

- [69] L. K. Doerfel et al. "EF-P is essential for rapid synthesis of proteins containing consecutive proline residues." In: *Science (New York, N.Y.)* 339.6115 (2013), pp. 85–88.
- [70] L. K. Doerfel et al. "Entropic Contribution of Elongation Factor P to Proline Positioning at the Catalytic Center of the Ribosome." In: *Journal of the American Chemical Society* 137.40 (2015), pp. 12997–13006.
- [71] T. J. Dolinsky et al. "PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations." In: *Nucleic acids research* 32.Web Server issue (2004), W665–7.
- [72] Q. P. Dou and K. Y. Chen. "Characterization and reconstitution of a cell free system for NAD(+)-dependent deoxyhypusine formation on the 18 kDa eIF-4D precursor." In: *Biochimica et biophysica acta* 1036.2 (1990), pp. 128–137.
- [73] D. A. Dougherty. "Cation- π interactions in chemistry and biology: A new view of benzene, Phe, Tyr, and Trp." In: *Science (New York, N.Y.)* 271.5246 (1996), pp. 163–168.
- [74] J. Doye and W. Poon. "Protein crystallization in vivo." In: *Current Opinion in Colloid & Interface Science* 11.1 (2006), pp. 40–46.
- [75] R. F. Duncan and J. W. Hershey. "Changes in eIF-4D hypusine modification or abundance are not correlated with translational repression in HeLa cells." In: *The Journal of biological chemistry* 261.27 (1986), pp. 12903–12906.
- [76] M. Duzsenko et al. "In vivo protein crystallization in combination with highly brilliant radiation sources offers novel opportunities for the structural analysis of post-translationally modified eukaryotic proteins." In: *Acta crystallographica. Section F, Structural biology communications* 71.Pt 8 (2015), pp. 929–937.
- [77] N. Dyballa and S. Metzger. "Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels." In: *Journal of visualized experiments : JoVE* 30 (2009).
- [78] K. Elbers et al. "Processing in the pestivirus E2-NS2 region: Identification of proteins p7 and E2p7." In: *Journal of Virology* 70.6 (1996), pp. 4131–4135.
- [79] G. Y. Fan et al. "In vivo calcineurin crystals formed using the baculovirus expression system." In: *Microscopy research and technique* 34.1 (1996), pp. 77–86.
- [80] C. Firth et al. "Detection of zoonotic pathogens and characterization of novel viruses carried by commensal *Rattus norvegicus* in New York City." In: *mBio* 5.5 (2014), e01933–14.
- [81] V. G. Francis, M. A. Majeed, and S. N. Gummadi. "Recovery of functionally active recombinant human phospholipid scramblase 1 from inclusion bodies using N-lauroyl sarcosine." In: *Journal of industrial microbiology & biotechnology* 39.7 (2012), pp. 1041–1048.
- [82] P. A. Frey. In: *Pyridine Nucleotide Coenzymes, Chemical, Biochemical and Medical Aspects: Part B, Vol. II*. Ed. by D. Dolphin, R. Poulson, and O. Avramovic. New York: John Wiley & Sons, 1987, pp. 461–511.
- [83] J. Fritzemeier et al. "The development of early vs. late onset mucosal disease is a consequence of two different pathogenic mechanisms." In: *Archives of virology* 142.7 (1997), pp. 1335–1350.
- [84] M. Fuhrmann et al. "Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene." In: *Plant Molecular Biology* 55.6 (2004), pp. 869–881.
- [85] C. Gati et al. "Serial crystallography on in vivo grown microcrystals using synchrotron radiation." In: *IUCr* 1.Pt 2 (2014), pp. 87–94.
- [86] P. M. Gentz, G. L. Blatch, and R. A. Dorrington. "Dimerization of the yeast eukaryotic translation initiation factor 5A requires hypusine and is RNA dependent." In: *The FEBS journal* 276.3 (2009), pp. 695–706.
- [87] F. Glaser et al. "ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information." In: *Bioinformatics (Oxford, England)* 19.1 (2003), pp. 163–164.
- [88] A. E. Gorbalenya and E. J. Snijder. "Viral cysteine proteinases." In: *Perspectives in Drug Discovery and Design* 6.1 (1996), pp. 64–86.
- [89] E. D. Gordon et al. "Hypusine formation in eukaryotic initiation factor 4D is not reversed when rates or specificity of protein synthesis is altered." In: *The Journal of biological chemistry* 262.34 (1987), pp. 16590–16595.
- [90] A. Grakoui et al. "A second hepatitis C virus-encoded proteinase." In: *Proceedings of the National Academy of Sciences of the United States of America* 90.22 (1993), pp. 10583–10587.
- [91] C. W. Grassmann, O. Isken, and S. E. Behrens. "Assignment of the multifunctional NS3 protein of bovine viral diarrhoea virus during RNA replication: An in vivo and in vitro study." In: *Journal of Virology* 73.11 (1999), pp. 9196–9205.

- [92] H. C. T. Groom, E. C. Anderson, and A. M. L. Lever. "Rev: Beyond nuclear export." In: *The Journal of general virology* 90.Pt 6 (2009), pp. 1303–1318.
- [93] B. Grummer et al. "Localization of viral proteins in cells infected with bovine viral diarrhoea virus." In: *The Journal of general virology* 82.Pt 11 (2001), pp. 2597–2605.
- [94] B. Gu et al. "The RNA helicase and nucleotide triphosphatase activities of the bovine viral diarrhoea virus NS3 protein are essential for viral replication." In: *Journal of Virology* 74.4 (2000), pp. 1794–1800.
- [95] R. Gupta and S. Brunak. "Prediction of glycosylation across the human proteome and the correlation to protein function." In: *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* (2002), pp. 310–322.
- [96] R. Gupta, E. Jung, and S. Brunak. *Prediction of N-glycosylation sites in human proteins*. Vol. 46. 2004.
- [97] E. Gutierrez et al. "eIF5A promotes translation of polyproline motifs." In: *Molecular cell* 51.1 (2013), pp. 35–45.
- [98] K. Hamana et al. "Cellular polyamine catalogues of the five classes of the phylum Proteobacteria: Distributions of homospermidine within the class Alphaproteobacteria, hydroxyputrescine within the class Betaproteobacteria, norspermidine within the class Gammaproteobacteria, and spermidine within the classes Deltaproteobacteria and Epsilonproteobacteria." In: *Ann Gunma Health Sci* 27 (2007).
- [99] D. Hanahan. "Studies on transformation of *Escherichia coli* with plasmids." In: *Journal of molecular biology* 166.4 (1983), pp. 557–580.
- [100] H. M. Hanauske-Abel et al. "Inhibition of the G1-S transition of the cell cycle by inhibitors of deoxyhypusine hydroxylation." In: *Biochimica et biophysica acta* 1221.2 (1994), pp. 115–124.
- [101] T. Harada, N. Tautz, and H. J. Thiel. "E2-p7 region of the bovine viral diarrhoea virus polyprotein: Processing and functional studies." In: *Journal of Virology* 74.20 (2000), pp. 9498–9506.
- [102] C. Haupt et al. "Combining Chemical Cross-linking and Mass Spectrometry of Intact Protein Complexes to Study the Architecture of Multi-subunit Protein Assemblies." In: *Journal of visualized experiments : JoVE* 129 (2017).
- [103] A. J. R. Heck. "Native mass spectrometry: A bridge between interactomics and structural biology." In: *Nature methods* 5.11 (2008), pp. 927–933.
- [104] M. Hijikata et al. "Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus." In: *Journal of Virology* 67.8 (1993), pp. 4665–4675.
- [105] C. E. Hodgman and M. C. Jewett. "Optimized extract preparation methods and reaction conditions for improved yeast cell-free protein synthesis." In: *Biotechnology and bioengineering* 110.10 (2013), pp. 2643–2654.
- [106] M. Hoque et al. "Inhibition of HIV-1 gene expression by Cyclopirox and Deferiprone, drugs that prevent hypusination of eukaryotic initiation factor 5A." In: *Retrovirology* 6 (2009), p. 90.
- [107] H. Houe. "Economic impact of BVDV infection in dairies." In: *Biologicals : journal of the International Association of Biological Standardization* 31.2 (2003), pp. 137–143.
- [108] Z. Huang et al. "Active inclusion bodies of acid phosphatase PhoC: Aggregation induced by GFP fusion and activities modulated by linker flexibility." In: *Microbial cell factories* 12 (2013), p. 25.
- [109] C. Hulo et al. "ViralZone: A knowledge resource to understand virus diversity." In: *Nucleic acids research* 39.Database issue (2011), pp. D576–82.
- [110] P. Huter et al. "Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P." In: *Molecular cell* 68.3 (2017), 515–527.e6.
- [111] N. Imaoka and T. Nakajima. "Hypusine, N6-(4-amino-2-hydroxybutyl)-2,6-diaminohexanoic acid, in tissue proteins of mammals." In: *Biochimica et Biophysica Acta (BBA) - General Subjects* 320.1 (1973), pp. 97–103.
- [112] M. Iqbal, H. Flick-Smith, and J. W. McCauley. "Interactions of bovine viral diarrhoea virus glycoprotein E(rns) with cell surface glycosaminoglycans." In: *The Journal of general virology* 81.Pt 2 (2000), pp. 451–459.
- [113] A. J. Jakobi et al. "In cellulose serial crystallography of alcohol oxidase crystals inside yeast cells." In: *IUCrJ* 3.Pt 2 (2016), pp. 88–95.
- [114] J. Jakus et al. "Features of the spermidine-binding site of deoxyhypusine synthase as derived from inhibition studies. Effective inhibition by bis- and mono-guanylated diamines and polyamines." In: *The Journal of biological chemistry* 268.18 (1993), pp. 13151–13159.
- [115] X. Ji et al. "How baculovirus polyhedra fit square pegs into round holes to robustly package viruses." In: *The EMBO journal* 29.2 (2010), pp. 505–514.

- [116] B.-F. Jin et al. "The effect of eIF-5A on the G1-S in cell cycle regulation." In: *Zhongguo shi yan xue ye xue za zhi* 11.4 (2003), pp. 325–328.
- [117] Y. A. Joe and M. H. Park. "Structural features of the eIF-5A precursor required for posttranslational synthesis of deoxyhypusine." In: *Journal of Biological Chemistry* 269.41 (1994), pp. 25916–25921.
- [118] Y. A. Joe, E. C. Wolff, and M. H. Park. "Cloning and expression of human deoxyhypusine synthase cDNA. Structure-function studies with the recombinant enzyme and mutant proteins." In: *The Journal of biological chemistry* 270.38 (1995), pp. 22386–22392.
- [119] E. Jurrus et al. "Improvements to the APBS biomolecular solvation software suite." In: *Protein science : a publication of the Protein Society* 27.1 (2018), pp. 112–128.
- [120] E. Kaltenecker, E. Eich, and D. Ober. "Evolution of homospermidine synthase in the convolvulaceae: A story of gene duplication, gene loss, and periods of various selection pressures." In: *The Plant cell* 25.4 (2013), pp. 1213–1227.
- [121] E. Kaltenecker and D. Ober. "Paralogue Interference Affects the Dynamics after Gene Duplication." In: *Trends in plant science* 20.12 (2015), pp. 814–821.
- [122] H. A. Kang and J. W. Hershey. "Effect of initiation factor eIF-5A depletion on protein synthesis and proliferation of *Saccharomyces cerevisiae*." In: *The Journal of biological chemistry* 269.6 (1994), pp. 3934–3940.
- [123] K. R. Kang and S. I. Chung. "Characterization of yeast deoxyhypusine synthase: PKC-dependent phosphorylation in vitro and functional domain identification." In: *Experimental & molecular medicine* 31.4 (1999), pp. 210–216.
- [124] K. R. Kang et al. "Identification of YHR068w in *Saccharomyces cerevisiae* chromosome VIII as a gene for deoxyhypusine synthase. Expression and characterization of the enzyme." In: *The Journal of biological chemistry* 270.31 (1995), pp. 18408–18412.
- [125] T. Kawate and E. Gouaux. "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins." In: *Structure (London, England : 1993)* 14.4 (2006), pp. 673–681.
- [126] K. K. Kim et al. "Crystal structures of eukaryotic translation initiation factor 5A from *Methanococcus jannaschii* at 1.8 Å resolution." In: *Proceedings of the National Academy of Sciences of the United States of America* 95.18 (1998), pp. 10419–10424.
- [127] K. K. Kim et al. *Eukaryotic translation initiation factor 5A from Methanococcus jannaschii*. 1999.
- [128] P. D. Kirkland et al. "Bungowannah virus—a probable new species of pestivirus—what have we found in the last 10 years?" In: *Animal health research reviews* 16.1 (2015), pp. 60–63.
- [129] P. D. Kirkland et al. "Replication of bovine viral diarrhoea virus in the bovine reproductive tract and excretion of virus in semen during acute and chronic infections." In: *The Veterinary record* 128.25 (1991), pp. 587–590.
- [130] M. Klinge. "Strukturelle Charakterisierung von Proteinen onkolytischer Viren." Dissertation. Hamburg: Universität Hamburg, 2015.
- [131] R. Koopmann et al. "In vivo protein crystallization opens new routes in structural biology." In: *Nature methods* 9.3 (2012), pp. 259–262.
- [132] T. Krey, H. J. Thiel, and T. Rümenapf. "Acid-resistant bovine pestivirus requires activation for pH-triggered fusion during entry." In: *Journal of Virology* 79.7 (2005), pp. 4191–4200.
- [133] S. Krossa et al. "Comprehensive Structural Characterization of the Bacterial Homospermidine Synthase—an Essential Enzyme of the Polyamine Metabolism." In: *Scientific reports* 6 (2016), p. 19501.
- [134] L. H. Kruse et al. "Identification of a Second Site of Pyrrolizidine Alkaloid Biosynthesis in Comfrey to Boost Plant Defense in Floral Stage." In: *Plant physiology* 174.1 (2017), pp. 47–55.
- [135] E. Kubovicová et al. "Alteration in ultrastructural morphology of bovine embryos following subzonal microinjection of bovine viral diarrhoea virus (BVDV)." In: *Zygote (Cambridge, England)* 16.3 (2008), pp. 187–193.
- [136] N. A. Kulak et al. "Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells." In: *Nature methods* 11.3 (2014), pp. 319–324.
- [137] B. M. Kümmerer, D. Stoll, and G. Meyers. "Bovine viral diarrhoea virus strain Oregon: A novel mechanism for processing of NS2-3 based on point mutations." In: *Journal of Virology* 72.5 (1998), pp. 4127–4138.
- [138] T. Lackner, H. J. Thiel, and N. Tautz. "Dissection of a viral autoprotease elucidates a function of a cellular chaperone in proteolysis." In: *Proceedings of the National Academy of Sciences of the United States of America* 103.5 (2006), pp. 1510–1515.

- [139] T. Lackner et al. "Persistence of bovine viral diarrhea virus is determined by a cellular cofactor of a viral autoprotease." In: *Journal of Virology* 79.15 (2005), pp. 9746–9755.
- [140] T. Lackner et al. "Temporal modulation of an autoprotease is crucial for replication and pathogenicity of an RNA virus." In: *Journal of Virology* 78.19 (2004), pp. 10765–10775.
- [141] U. K. Laemmli. "Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4." In: *Nature* 227.5259 (1970), pp. 680–685.
- [142] A. A. Laminet et al. "The Escherichia coli heat shock proteins GroEL and GroES modulate the folding of the beta-lactamase precursor." In: *The EMBO journal* 9.7 (1990), pp. 2315–2319.
- [143] M. Landau et al. "ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures." In: *Nucleic acids research* 33.Web Server issue (2005), W299–302.
- [144] M. Lang and V. Orgogozo. "Identification of homologous gene sequences by PCR with degenerate primers." In: *Methods in molecular biology (Clifton, N.J.)* 772 (2011), pp. 245–256.
- [145] S. R. Lanyon et al. "Bovine viral diarrhoea: Pathogenesis and diagnosis." In: *Veterinary journal (London, England : 1997)* 199.2 (2014), pp. 201–209.
- [146] J. Lassak et al. "Arginine-rhamnosylation as new strategy to activate translation elongation factor P." In: *Nature chemical biology* 11.4 (2015), pp. 266–270.
- [147] S. Lecot et al. "Bovine viral diarrhea virus entry is dependent on clathrin-mediated endocytosis." In: *Journal of Virology* 79.16 (2005), pp. 10826–10829.
- [148] C. H. Lee and M. H. Park. "Human deoxyhypusine synthase: Interrelationship between binding of NAD and substrates." In: *The Biochemical journal* 352.3 (2000), pp. 851–857.
- [149] C. H. Lee, P. Y. Um, and M. H. Park. "Structure-function studies of human deoxyhypusine synthase: Identification of amino acid residues critical for the binding of spermidine and NAD." In: *The Biochemical journal* 355.Pt 3 (2001), pp. 841–849.
- [150] J. Lee et al. "An alternative polyamine biosynthetic pathway is widespread in bacteria and essential for biofilm formation in *Vibrio cholerae*." In: *The Journal of biological chemistry* 284.15 (2009), pp. 9899–9907.
- [151] Y. B. Lee et al. "Complex formation between deoxyhypusine synthase and its protein substrate, the eukaryotic translation initiation factor 5A (eIF5A) precursor." In: *The Biochemical journal* 340.1 (1999), pp. 273–281.
- [152] D. D. Leonidas et al. "Crystal structure of human Charcot–Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins." In: *Structure (London, England : 1993)* 3.12 (1995), pp. 1379–1393.
- [153] D. I. Liao et al. "Crystal structure of the NAD complex of human deoxyhypusine synthase: An enzyme with a ball-and-chain mechanism for blocking the active site." In: *Structure (London, England : 1993)* 6.1 (1998), pp. 23–32.
- [154] Y. P. Liu et al. "Interaction of eukaryotic initiation factor 5A with the human immunodeficiency virus type 1 Rev response element RNA and U6 snRNA requires deoxyhypusine or hypusine modification." In: *Biological signals* 6.3 (1997), pp. 166–174.
- [155] T. Livshultz et al. "Evolution of pyrrolizidine alkaloid biosynthesis in Apocynaceae: Revisiting the defence de-escalation hypothesis." In: *The New phytologist* (2018).
- [156] I. C. Lorenz et al. "Structure of the catalytic domain of the hepatitis C virus NS2-3 protease." In: *Nature* 442.7104 (2006), pp. 831–835.
- [157] M. Macel. "Attract and deter: A dual role for pyrrolizidine alkaloids in plant-insect interactions." In: *Phytochemistry Reviews* 10.1 (2011), pp. 75–82.
- [158] A. Macovei et al. "Brefeldin A inhibits pestivirus release from infected cells, without affecting its assembly and infectivity." In: *Biochemical and biophysical research communications* 346.3 (2006), pp. 1083–1090.
- [159] V. Magdolen et al. "The function of the hypusine-containing proteins of yeast and other eukaryotes is well conserved." In: *Molecular & general genetics : MGG* 244.6 (1994), pp. 646–652.
- [160] B. Maier, S. A. Tersey, and R. G. Mirmira. "Hypusine: A new target for therapeutic intervention in diabetic inflammation." In: *Discovery medicine* 10.50 (2010), pp. 18–23.
- [161] B. Maier et al. "The unique hypusine modification of eIF5A promotes islet beta cell inflammation and dysfunction in mice." In: *The Journal of clinical investigation* 120.6 (2010), pp. 2156–2170.
- [162] A. Mandal, S. Mandal, and M. H. Park. "Genome-wide analyses and functional classification of proline repeat-rich proteins: Potential role of eIF5A in eukaryotic evolution." In: *PloS one* 9.11 (2014), e111800.

- [163] J. G. Marblestone et al. "Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO." In: *Protein science : a publication of the Protein Society* 15.1 (2006), pp. 182–189.
- [164] M. B. Mathews and J. W. B. Hershey. "The translation factor eIF5A and human cancer." In: *Biochimica et biophysica acta* 1849.7 (2015), pp. 836–844.
- [165] K. Maurer et al. "CD46 Is a Cellular Receptor for Bovine Viral Diarrhea Virus." In: *Journal of Virology* 78.4 (2004), pp. 1792–1799.
- [166] S. Melnikov et al. "Crystal Structure of Hypusine-Containing Translation Factor eIF5A Bound to a Rotated Eukaryotic Ribosome." In: *Journal of molecular biology* 428.18 (2016), pp. 3570–3576.
- [167] E. Mendez et al. "Infectious bovine viral diarrhoea virus (strain NADL) RNA from stable cDNA clones: A cellular insert determines NS3 production and viral cytopathogenicity." In: *Journal of Virology* 72.6 (1998), pp. 4737–4745.
- [168] C. Meyer et al. "Recovery of virulent and RNase-negative attenuated type 2 bovine viral diarrhoea viruses from infectious cDNA clones." In: *Journal of Virology* 76.16 (2002), pp. 8494–8503.
- [169] G. Meyers and H. J. Thiel. "Molecular characterization of pestiviruses." In: *Advances in virus research* 47 (1996), pp. 53–118.
- [170] G. Meyers et al. "Recovery of cytopathogenic and noncytopathogenic bovine viral diarrhoea viruses from cDNA constructs." In: *Journal of Virology* 70.12 (1996), pp. 8606–8613.
- [171] A. Meyling and A. M. Jensen. "Transmission of bovine virus diarrhoea virus (BVDV) by artificial insemination (AI) with semen from a persistently-infected bull." In: *Veterinary microbiology* 17.2 (1988), pp. 97–105.
- [172] T. Miyake et al. "XPO1/CRM1 Inhibition Causes Antitumor Effects by Mitochondrial Accumulation of eIF5A." In: *Clinical cancer research : an official journal of the American Association for Cancer Research* 21.14 (2015), pp. 3286–3297.
- [173] V. Moennig and P. Becher. "Pestivirus control programs: How far have we come and where are we going?" In: *Animal health research reviews* 16.1 (2015), pp. 83–87.
- [174] V. Moennig and B. Liess. "Pathogenesis of intrauterine infections with bovine viral diarrhoea virus." In: *The Veterinary clinics of North America. Food animal practice* 11.3 (1995), pp. 477–487.
- [175] S. Moll et al. "Cell-specific expression of homospermidine synthase, the entry enzyme of the pyrrolizidine alkaloid pathway in *Senecio vernalis*, in comparison with its ancestor, deoxyhypusine synthase." In: *Plant physiology* 130.1 (2002), pp. 47–57.
- [176] S. Moretti et al. "The M-Coffee web server: A meta-method for computing multiple sequence alignments by combining alternative alignment methods." In: *Nucleic acids research* 35.Web Server issue (2007), W645–8.
- [177] A. Müller et al. "Cell-derived sequences in the N-terminal region of the polyprotein of a cytopathogenic pestivirus." In: *Journal of Virology* 77.19 (2003), pp. 10663–10669.
- [178] K. Mullis et al. "Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction." In: *Cold Spring Harbor symposia on quantitative biology* 51 Pt 1 (1986), pp. 263–273.
- [179] R. J. Murphey and E. W. Gerner. "Hypusine formation in protein by a two-step process in cell lysates." In: *The Journal of biological chemistry* 262.31 (1987), pp. 15033–15036.
- [180] T. Nakajima et al. "Distribution of hypusine, N6(-4-amino-2-hydroxybutyl)-2, 6-diaminohexanoic acid, in mammalian organs." In: *Biochimica et Biophysica Acta (BBA) - General Subjects* 252.1 (1971), pp. 92–97.
- [181] Y. Nakamura, T. Gojobori, and T. Ikemura. "Codon usage tabulated from international DNA sequence databases: Status for the year 2000." In: *Nucleic acids research* 28.1 (2000), p. 292.
- [182] S. Nakanishi and J. L. Cleveland. "Targeting the polyamine-hypusine circuit for the prevention and treatment of cancer." In: *Amino acids* 48.10 (2016), pp. 2353–2362.
- [183] G. K. Nasrallah et al. "Legionella pneumophila requires polyamines for optimal intracellular growth." In: *Journal of bacteriology* 193.17 (2011), pp. 4346–4360.
- [184] D. Niemüller, A. Reimann, and D. Ober. "Distinct cell-specific expression of homospermidine synthase involved in pyrrolizidine alkaloid biosynthesis in three species of the boraginales." In: *Plant physiology* 159.3 (2012), pp. 920–929.
- [185] M. Nikolaev et al. "Integral Membrane Proteins Can Be Crystallized Directly from Nanodiscs." In: *Crystal Growth & Design* 17.3 (2017), pp. 945–948.

- [186] C. Notredame, D. G. Higgins, and J. Heringa. "T-Coffee: A novel method for fast and accurate multiple sequence alignment." In: *Journal of molecular biology* 302.1 (2000), pp. 205–217.
- [187] N. Nurhayati, D. Gondé, and D. Ober. "Evolution of pyrrolizidine alkaloids in Phalaenopsis orchids and other monocotyledons: Identification of deoxyhypusine synthase, homospermidine synthase and related pseudogenes." In: *Phytochemistry* 70.4 (2009), pp. 508–516.
- [188] N. Nurhayati and D. Ober. "Recruitment of alkaloid-specific homospermidine synthase (HSS) from ubiquitous deoxyhypusine synthase: Does *Crotalaria* possess a functional HSS that still has DHS activity?" In: *Phytochemistry* 66.11 (2005), pp. 1346–1357.
- [189] D. Ober and T. Hartmann. "Deoxyhypusine synthase from tobacco. cDNA isolation, characterization, and bacterial expression of an enzyme with extended substrate specificity." In: *The Journal of biological chemistry* 274.45 (1999), pp. 32040–32047.
- [190] D. Ober and T. Hartmann. "Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase." In: *Proceedings of the National Academy of Sciences of the United States of America* 96.26 (1999), pp. 14777–14782.
- [191] D. Ober et al. "Homospermidine synthase of *Rhodospseudomonas viridis*: Substrate specificity and effects of the heterologously expressed enzyme on polyamine metabolism of *Escherichia coli*." In: *The Journal of General and Applied Microbiology* 42.5 (1996), pp. 411–419.
- [192] D. Ober et al. "Molecular evolution by change of function. Alkaloid-specific homospermidine synthase retained all properties of deoxyhypusine synthase except binding the eIF5A precursor protein." In: *The Journal of biological chemistry* 278.15 (2003), pp. 12805–12812.
- [193] K. Okonechnikov, O. Golosova, and M. Fursov. "Unipro UGENE: A unified bioinformatics toolkit." In: *Bioinformatics (Oxford, England)* 28.8 (2012), pp. 1166–1167.
- [194] M. E. Olsen et al. "Polyamines and Hypusination Are Required for Ebolavirus Gene Expression and Replication." In: *mBio* 7.4 (2016).
- [195] U. Omasits et al. "Protter: Interactive protein feature visualization and integration with experimental proteomic data." In: *Bioinformatics (Oxford, England)* 30.6 (2014), pp. 884–886.
- [196] D. J. Oshannessy et al. "Determination of Rate and Equilibrium Binding Constants for Macromolecular Interactions Using Surface Plasmon Resonance: Use of Nonlinear Least Squares Analysis Methods." In: *Analytical biochemistry* 212.2 (1993), pp. 457–468.
- [197] C. Papworth et al. "Site-directed mutagenesis in one day with >80% efficiency." In: *Strategies* 9.3 (1996), pp. 3–4.
- [198] J. H. Park et al. "Molecular cloning, expression, and structural prediction of deoxyhypusine hydroxylase: A HEAT-repeat-containing metalloenzyme." In: *Proceedings of the National Academy of Sciences of the United States of America* 103.1 (2006), pp. 51–56.
- [199] J. H. Park et al. "Post-translational modification by β -lysylation is required for activity of *Escherichia coli* elongation factor P (EF-P)." In: *The Journal of biological chemistry* 287.4 (2012), pp. 2579–2590.
- [200] J. H. Park et al. "Production of active recombinant eIF5A: Reconstitution in *E. coli* of eukaryotic hypusine modification of eIF5A by its coexpression with modifying enzymes." In: *Protein engineering, design & selection : PEDS* 24.3 (2011), pp. 301–309.
- [201] J. H. Park et al. "Reversal of the deoxyhypusine synthesis reaction. Generation of spermidine or homospermidine from deoxyhypusine by deoxyhypusine synthase." In: *The Journal of biological chemistry* 278.35 (2003), pp. 32683–32691.
- [202] M. H. Park. "Regulation of biosynthesis of hypusine in Chinese hamster ovary cells. Evidence for eIF-4D precursor polypeptides." In: *The Journal of biological chemistry* 262.26 (1987), pp. 12730–12734.
- [203] M. H. Park. "The essential role of hypusine in eukaryotic translation initiation factor 4D (eIF-4D). Purification of eIF-4D and its precursors and comparison of their activities." In: *The Journal of biological chemistry* 264.31 (1989), pp. 18531–18535.
- [204] M. H. Park. "The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A)." In: *Journal of biochemistry* 139.2 (2006), pp. 161–169.
- [205] M. H. Park, H. L. Cooper, and J. E. Folk. "Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor." In: *Proceedings of the National Academy of Sciences of the United States of America* 78.5 (1981), pp. 2869–2873.

- [206] M. H. Park, H. L. Cooper, and J. E. Folk. "The biosynthesis of protein-bound hypusine (N epsilon -(4-amino-2-hydroxybutyl)lysine). Lysine as the amino acid precursor and the intermediate role of deoxyhypusine (N epsilon -(4-aminobutyl)lysine)." In: *The Journal of biological chemistry* 257.12 (1982), pp. 7217–7222.
- [207] M. H. Park and J. E. Folk. "Biosynthetic labeling of hypusine in mammalian cells. Carbon-hydrogen bond fissions revealed by dual labeling." In: *The Journal of biological chemistry* 261.30 (1986), pp. 14108–14111.
- [208] M. H. Park and E. C. Wolff. "Cell-free synthesis of deoxyhypusine. Separation of protein substrate and enzyme and identification of 1,3-diaminopropane as a product of spermidine cleavage." In: *The Journal of biological chemistry* 263.30 (1988), pp. 15264–15269.
- [209] M. H. Park, E. C. Wolff, and J. E. Folk. "Hypusine: Its post-translational formation in eukaryotic initiation factor 5A and its potential role in cellular regulation." In: *BioFactors (Oxford, England)* 4.2 (1993), pp. 95–104.
- [210] M. H. Park et al. "A new non-radioactive deoxyhypusine synthase assay adaptable to high throughput screening." In: *Amino acids* 49.11 (2017), pp. 1793–1804.
- [211] M. H. Park et al. "Antiproliferative effects of inhibitors of deoxyhypusine synthase. Inhibition of growth of Chinese hamster ovary cells by guanidyl diamines." In: *The Journal of biological chemistry* 269.45 (1994), pp. 27827–27832.
- [212] M. H. Park et al. "Comparison of the activities of variant forms of eIF-4D. The requirement for hypusine or deoxyhypusine." In: *The Journal of biological chemistry* 266.13 (1991), pp. 7988–7994.
- [213] M. H. Park et al. "Eukaryotic initiation factor 4D. Purification from human red blood cells and the sequence of amino acids around its single hypusine residue." In: *The Journal of biological chemistry* 261.31 (1986), pp. 14515–14519.
- [214] M. H. Park et al. "The mammalian hypusine-containing protein, eukaryotic initiation factor 4D. Structural homology of this protein from several species." In: *The Journal of biological chemistry* 259.7 (1984), pp. 4563–4565.
- [215] M. Parmar et al. "Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure." In: *Biochimica et biophysica acta* 1860.2 (2018), pp. 378–383.
- [216] P. H. Patel et al. "The Drosophila deoxyhypusine hydroxylase homologue nero and its target eIF5A are required for cell growth and the regulation of autophagy." In: *The Journal of cell biology* 185.7 (2009), pp. 1181–1194.
- [217] M. Y. Pavlov et al. "Slow peptide bond formation by proline and other N-alkylamino acids in translation." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.1 (2009), pp. 50–54.
- [218] M. A. Paz, B. M. Torrelío, and P. M. Gallop. "Hydralazine inhibition of the post-translational hydroxylation of deoxyhypusine, a polyamine-derived amino acid." In: *Biochemical pharmacology* 33.5 (1984), pp. 779–785.
- [219] A. E. Pegg. "Mammalian polyamine metabolism and function." In: *IUBMB life* 61.9 (2009), pp. 880–894.
- [220] L. Peil et al. "Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.38 (2013), pp. 15265–15270.
- [221] V. Pelechano and P. Alepuz. "eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences." In: *Nucleic acids research* 45.12 (2017), pp. 7326–7338.
- [222] C. Pellerin et al. "Identification of a New Group of Bovine Viral Diarrhea Virus Strains Associated with Severe Outbreaks and High Mortalities." In: *Virology* 203.2 (1994), pp. 260–268.
- [223] S. Peternel et al. "Engineering inclusion bodies for non denaturing extraction of functional proteins." In: *Microbial cell factories* 7 (2008), p. 34.
- [224] V. W. Pollard and M. H. Malim. "The HIV-1 Rev protein." In: *Annual review of microbiology* 52 (1998), pp. 491–532.
- [225] T. L. Poole et al. "Pestivirus translation initiation occurs by internal ribosome entry." In: *Virology* 206.1 (1995), pp. 750–754.
- [226] A. Rajkovic et al. "Cyclic Rhamnosylated Elongation Factor P Establishes Antibiotic Resistance in Pseudomonas aeruginosa." In: *mBio* 6.3 (2015), e00823.
- [227] F. K. Ramsey and W. H. Chivers. "Mucosal disease of cattle." In: *Veterinary Clinics of North America: Food Animal Practice* 34 (1953), pp. 629–634.
- [228] S. Raymond and L. Weintraub. "Acrylamide gel as a supporting medium for zone electrophoresis." In: *Science (New York, N.Y.)* 130.3377 (1959), p. 711.

- [229] L. Redecke et al. "Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser." In: *Science (New York, N.Y.)* 339.6116 (2013), pp. 227–230.
- [230] A. Reimann et al. "Repeated evolution of the pyrrolizidine alkaloid-mediated defense system in separate angiosperm lineages." In: *The Plant cell* 16.10 (2004), pp. 2772–2784.
- [231] J. Renart, J. Reiser, and G. R. Stark. "Transfer of proteins from gels to diazobenzoyloxymethyl-paper and detection with antisera: A method for studying antibody specificity and antigen structure." In: *Proceedings of the National Academy of Sciences of the United States of America* 76.7 (1979), pp. 3116–3120.
- [232] V. Richter et al. "A systematic worldwide review of the direct monetary losses in cattle due to bovine viral diarrhoea virus infection." In: *Veterinary journal (London, England : 1997)* 220 (2017), pp. 80–87.
- [233] J. F. Ridpath, S. R. Bolin, and E. J. Dubovi. "Segregation of bovine viral diarrhoea virus into genotypes." In: *Virology* 205.1 (1994), pp. 66–74.
- [234] J. F. Ridpath and J. D. Neill. "Detection and characterization of genetic recombination in cytopathic type 2 bovine viral diarrhoea viruses." In: *Journal of Virology* 74.18 (2000), pp. 8771–8774.
- [235] G. Rinck et al. "A cellular J-domain protein modulates polyprotein processing and cytopathogenicity of a pestivirus." In: *Journal of Virology* 75.19 (2001), pp. 9470–9482.
- [236] R. D. Robbins et al. "Inhibition of deoxyhypusine synthase enhances islet beta cell function and survival in the setting of endoplasmic reticulum stress and type 2 diabetes." In: *The Journal of biological chemistry* 285.51 (2010), pp. 39943–39952.
- [237] O. Rosorius et al. "Nuclear pore localization and nucleocytoplasmic transport of eIF-5A: Evidence for direct interaction with the export receptor CRM1." In: *Journal of cell science* 112 (Pt 14) (1999), pp. 2369–2380.
- [238] M. Ruhl et al. "Eukaryotic initiation factor 5A is a cellular target of the human immunodeficiency virus type 1 Rev activation domain mediating trans-activation." In: *The Journal of cell biology* 123.6 Pt 1 (1993), pp. 1309–1320.
- [239] T. Rümehapf et al. "Processing of the envelope glycoproteins of pestiviruses." In: *Journal of Virology* 67.6 (1993), pp. 3288–3294.
- [240] U. Sandholzer et al. "cDNA and derived amino acid sequence of the hypusine containing protein from *Dicystostelium discoideum*." In: *FEBS letters* 246.1-2 (1989), pp. 94–100.
- [241] F. Sanger, S. Nicklen, and A. R. Coulson. "DNA sequencing with chain-terminating inhibitors." In: *Proceedings of the National Academy of Sciences of the United States of America* 74.12 (1977), pp. 5463–5467.
- [242] H. Schirmer et al. "Genetic and antigenic characterization of an atypical pestivirus isolate, a putative member of a novel pestivirus species." In: *The Journal of general virology* 85.Pt 12 (2004), pp. 3647–3652.
- [243] S. Schmeiser et al. "Morphogenesis of pestiviruses: New insights from ultrastructural studies of strain Giraffe-1." In: *Journal of Virology* 88.5 (2014), pp. 2717–2724.
- [244] C. Schmidt et al. "Structure of the hypusylated eukaryotic translation factor eIF-5A bound to the ribosome." In: *Nucleic acids research* 44.4 (2016), pp. 1944–1951.
- [245] J. Schnier et al. "Translation initiation factor 5A and its hypusine modification are essential for cell viability in the yeast *Saccharomyces cerevisiae*." In: *Molecular and cellular biology* 11.6 (1991), pp. 3105–3114.
- [246] R. Schönherr et al. "Real-time investigation of dynamic protein crystallization in living cells." In: *Structural dynamics (Melville, N.Y.)* 2.4 (2015), p. 041712.
- [247] M. H. Schreier, B. Erni, and T. Staehelin. "Initiation of mammalian protein synthesis. I. Purification and characterization of seven initiation factors." In: *Journal of molecular biology* 116.4 (1977), pp. 727–753.
- [248] M. Schroeder et al. "In silico design, synthesis, and screening of novel deoxyhypusine synthase inhibitors targeting HIV-1 replication." In: *ChemMedChem* 9.5 (2014), pp. 940–952.
- [249] A. P. Schuller et al. "eIF5A Functions Globally in Translation Elongation and Termination." In: *Molecular cell* 66.2 (2017), pp. 194–205.e5.
- [250] C. R. Schultz et al. "Synergistic drug combination GC7/DFMO suppresses hypusine/spermidine-dependent eIF5A activation and induces apoptotic cell death in neuroblastoma." In: *The Biochemical journal* 475.2 (2018), pp. 531–545.
- [251] P. H. Seeburg et al. "Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone." In: *Nature* 270.5637 (1977), pp. 486–494.
- [252] G. V. Semisotnov et al. "Study of the molten globule intermediate state in protein folding by a hydrophobic fluorescent probe." In: *Biopolymers* 31.1 (1991), pp. 119–128.

- [253] N. B. Shah and T. M. Duncan. "Bio-layer interferometry for measuring kinetics of protein-protein interactions and allosteric ligand effects." In: *Journal of visualized experiments : JoVE* 84 (2014), e51383.
- [254] P. Shah and E. Swiatlo. "A multifaceted role for polyamines in bacterial pathogens." In: *Molecular microbiology* 68.1 (2008), pp. 4–16.
- [255] N. C. Shaner et al. "Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein." In: *Nature biotechnology* 22.12 (2004), pp. 1567–1572.
- [256] T. Shiba et al. "Hypusine, a new amino acid occurring in bovine brain." In: *Biochimica et Biophysica Acta (BBA) - General Subjects* 244.3 (1971), pp. 523–531.
- [257] T. Shiba et al. "Synthesis and Stereochemistry of Hypusine, a New Amino Acid in Bovine Brain." In: *Bulletin of the Chemical Society of Japan* 55.3 (1982), pp. 899–903.
- [258] B.-S. Shin et al. "Amino acid substrates impose polyamine, eIF5A, or hypusine requirement for peptide synthesis." In: *Nucleic acids research* 45.14 (2017), pp. 8392–8402.
- [259] H. Sievert et al. "A novel mouse model for inhibition of DOHH-mediated hypusine modification reveals a crucial function in embryonic development, proliferation and oncogenic transformation." In: *Disease models & mechanisms* 7.8 (2014), pp. 963–976.
- [260] P. Simmonds et al. "ICTV Virus Taxonomy Profile: Flaviviridae." In: *The Journal of general virology* 98.1 (2017), pp. 2–3.
- [261] Z. Smit-McBride et al. "Protein synthesis initiation factor eIF-4D. Functional comparison of native and unhyposinated forms of the protein." In: *The Journal of biological chemistry* 264.31 (1989), pp. 18527–18530.
- [262] Z. Smit-McBride et al. "Sequence determination and cDNA cloning of eukaryotic initiation factor 4D, the hypusine-containing protein." In: *The Journal of biological chemistry* 264.3 (1989), pp. 1578–1583.
- [263] F. Sobott et al. "Subunit exchange of multimeric protein complexes. Real-time monitoring of subunit exchange between small heat shock proteins by using electrospray mass spectrometry." In: *The Journal of biological chemistry* 277.41 (2002), pp. 38921–38929.
- [264] C. Steentoft et al. "Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology." In: *The EMBO journal* 32.10 (2013), pp. 1478–1488.
- [265] F. W. Studier. "Protein production by auto-induction in high density shaking cultures." In: *Protein expression and purification* 41.1 (2005), pp. 207–234.
- [266] A. Szabo et al. "A zinc finger-like domain of the molecular chaperone DnaJ is involved in binding to denatured protein substrates." In: *The EMBO journal* 15.2 (1996), pp. 408–417.
- [267] K. D. Tartof and C. A. Hobbs. "Improved media for growing plasmid and cosmid clones." In: *Bethesda Research Laboratories Focus* 9.12 (1987).
- [268] N. Tautz et al. "Cytopathogenicity of a pestivirus correlates with a 27-nucleotide insertion." In: *Journal of Virology* 70.11 (1996), pp. 7851–7858.
- [269] N. Tautz et al. "Pathogenesis of mucosal disease: A cytopathogenic pestivirus generated by an internal deletion." In: *Journal of Virology* 68.5 (1994), pp. 3289–3297.
- [270] Y.-B. Teng et al. "Crystal structure of Arabidopsis translation initiation factor eIF-5A2." In: *Proteins* 77.3 (2009), pp. 736–740.
- [271] N. Thonghin et al. "Cryo-electron microscopy of membrane proteins." In: *Methods (San Diego, Calif.)* (2018).
- [272] P. B. Tiwari et al. "Analyzing surface plasmon resonance data: Choosing a correct biphasic model for interpretation." In: *The Review of scientific instruments* 86.3 (2015), p. 035001.
- [273] K. Tokatlidis et al. "High activity of inclusion bodies formed in *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D." In: *FEBS letters* 282.1 (1991), pp. 205–208.
- [274] Y. Tong et al. "Crystal structure of human eIF5A1: Insight into functional similarity of human eIF5A1 and eIF5A2." In: *Proteins* 75.4 (2009), pp. 1040–1045.
- [275] B. M. Torrelío, M. A. Paz, and P. M. Gallop. "The formation and stability of the hypusine containing protein in Chinese hamster ovary cells." In: *Biochemical and biophysical research communications* 145.3 (1987), pp. 1335–1341.
- [276] M. Trávén et al. "Primary bovine viral diarrhoea virus infection in calves following direct contact with a persistently viraemic calf." In: *Zentralblatt für Veterinärmedizin. Reihe B. Journal of veterinary medicine. Series B* 38.6 (1991), pp. 453–462.

- [277] A. T. Tuukkanen and D. I. Svergun. "Weak protein-ligand interactions studied by small-angle X-ray scattering." In: *The FEBS journal* 281.8 (2014), pp. 1974–1987.
- [278] S. Ude et al. "Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches." In: *Science (New York, N.Y.)* 339.6115 (2013), pp. 82–85.
- [279] A. Ullrich et al. "Rat insulin genes: Construction of plasmids containing the coding sequences." In: *Science (New York, N.Y.)* 196.4296 (1977), pp. 1313–1319.
- [280] T. C. Umland et al. "A new crystal structure of deoxyhypusine synthase reveals the configuration of the active enzyme and of an enzyme.NAD.inhibitor ternary complex." In: *The Journal of biological chemistry* 279.27 (2004), pp. 28697–28705.
- [281] L. Vayssié et al. "Molecular genetics of regulated secretion in paramecium." In: *Biochimie* 82.4 (2000), pp. 269–288.
- [282] S. Vilcek et al. "Characterization of a novel pestivirus originating from a pronghorn antelope." In: *Virus research* 108.1-2 (2005), pp. 187–193.
- [283] B. Vogelstein and D. Gillespie. "Preparative and analytical purification of DNA from agarose." In: *Proceedings of the National Academy of Sciences of the United States of America* 76.2 (1979), pp. 615–619.
- [284] J. Vonck and E. F. van Bruggen. "Architecture of peroxisomal alcohol oxidase crystals from the methylotrophic yeast *Hansenula polymorpha* as deduced by electron microscopy." In: *Journal of bacteriology* 174.16 (1992), pp. 5391–5399.
- [285] G. S. Waldo et al. "Rapid protein-folding assay using green fluorescent protein." In: *Nature biotechnology* 17.7 (1999), pp. 691–695.
- [286] I. M. Wallace et al. "M-Coffee: Combining multiple sequence alignment methods with T-Coffee." In: *Nucleic acids research* 34.6 (2006), pp. 1692–1699.
- [287] F. Weiland et al. "Localization of pestiviral envelope proteins E(rns) and E2 at the cell surface and on isolated particles." In: *The Journal of general virology* 80 (Pt 5) (1999), pp. 1157–1165.
- [288] E. Weiskircher et al. "Bovine viral diarrhea virus NS4B protein is an integral membrane protein associated with Golgi markers and rearranged host membranes." In: *Virology journal* 6 (2009), p. 185.
- [289] A.-M. Wesseling et al. "Variability of Pyrrolizidine Alkaloid Occurrence in Species of the Grass Subfamily Pooideae (Poaceae)." In: *Frontiers in plant science* 8 (2017), p. 2046.
- [290] M. Wiskerchen, S. K. Belzer, and M. S. Collett. "Pestivirus gene expression: The first protein product of the bovine viral diarrhea virus large open reading frame, p20, possesses proteolytic activity." In: *Journal of Virology* 65.8 (1991), pp. 4508–4514.
- [291] I. Wohlgenuth et al. "Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates." In: *The Journal of biological chemistry* 283.47 (2008), pp. 32229–32235.
- [292] E. C. Wolff, J. E. Folk, and M. H. Park. "Enzyme-substrate intermediate formation at lysine 329 of human deoxyhypusine synthase." In: *The Journal of biological chemistry* 272.25 (1997), pp. 15865–15871.
- [293] E. C. Wolff, M. H. Park, and J. E. Folk. "Cleavage of spermidine as the first step in deoxyhypusine synthesis. The role of NAD." In: *The Journal of biological chemistry* 265.9 (1990), pp. 4793–4799.
- [294] E. C. Wolff, J. Wolff, and M. H. Park. "Deoxyhypusine synthase generates and uses bound NADH in a transient hydride transfer mechanism." In: *The Journal of biological chemistry* 275.13 (2000), pp. 9170–9177.
- [295] E. C. Wolff et al. "Deoxyhypusine synthase from rat testis: Purification and characterization." In: *The Journal of biological chemistry* 270.15 (1995), pp. 8660–8666.
- [296] D. M. Worrall and N. H. Goss. "The formation of biologically active beta-galactosidase inclusion bodies in *Escherichia coli*." In: *Australian journal of biotechnology* 3.1 (1989), pp. 28–32.
- [297] Z. Wu et al. "Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces." In: *Journal of Virology* 86.20 (2012), pp. 10999–11012.
- [298] J. Xu et al. "Bovine viral diarrhea virus NS3 serine proteinase: Polyprotein cleavage sites, cofactor requirements, and molecular model of an enzyme essential for pestivirus replication." In: *Journal of Virology* 71.7 (1997), pp. 5312–5322.
- [299] S. Yamamoto, S. Nagata, and K. Kusaba. "Purification and characterization of homospermidine synthase in *Acinetobacter tartarogenes* ATCC 31105." In: *Journal of biochemistry* 114.1 (1993), pp. 45–49.
- [300] Y. C. Yang et al. "Deoxyhypusine/hypusine formation on a 21,000-dalton cellular protein in a *Neurospora crassa* mutant in vivo and in vitro." In: *Biochimica et biophysica acta* 1033.2 (1990), pp. 133–138.

- [301] M. Yao et al. "Crystal structure of hyperthermophilic archaeal initiation factor 5A: A homologue of eukaryotic initiation factor 5A (eIF-5A)." In: *Journal of biochemistry* 133.1 (2003), pp. 75–81.
- [302] K. Yeşilbağ, G. Alpay, and P. Becher. "Variability and Global Distribution of Subgenotypes of Bovine Viral Diarrhea Virus." In: *Viruses* 9.6 (2017).
- [303] Z. Yi et al. "Flavivirus replication complex assembly revealed by DNAJC14 functional mapping." In: *Journal of Virology* 86.21 (2012), pp. 11815–11832.
- [304] C. F. Zanelli and S. R. Valentini. "Is there a role for eIF5A in translation?" In: *Amino acids* 33.2 (2007), pp. 351–358.
- [305] G. Zhang, H. Flick-Smith, and J. W. McCauley. "Differences in membrane association and sub-cellular distribution between NS2-3 and NS3 of bovine viral diarrhoea virus." In: *Virus research* 97.2 (2003), pp. 89–102.
- [306] W. Zhong, L. L. Gutshall, and A. M. Del Vecchio. "Identification and characterization of an RNA-dependent RNA polymerase activity within the nonstructural protein 5B region of bovine viral diarrhea virus." In: *Journal of Virology* 72.11 (1998), pp. 9365–9369.