



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Engineering an Usutu virus specific scFv antibody using
computational saturated mutagenesis“

verfasst von / submitted by

Marina Pletzer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2019

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 834

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Molekulare Biologie

Betreut von / Supervisor:

ao. Univ.-Prof. Dr. Timothy Skern

Acknowledgements

An dieser Stelle möchte ich mich bei all jenen bedanken, die durch ihre fachliche und persönliche Unterstützung zum Fertigstellen dieser Arbeit beigetragen haben.

Zuerst möchte ich mich bei Tim für die Möglichkeit bedanken an diesem Projekt zu arbeiten, für seine Unterstützung das ganze letzte Jahr über und zahlreichen Ratschläge, wodurch ich unglaublich viel über Virologie und Strukturbiologie lernen konnte. Ebenso möchte ich mich bei Amelie bedanken, die mich bei der praktischen Arbeit im Labor betreut hat, für die zahlreichen Diskussionen wenn die Experimente nicht so funktioniert haben wie sie sollten und dafür, dass sie mir ihr Wissen weitergegeben hat. Auch bei meinen Kollegen möchte ich mich vor allem für die schöne und tolle Zeit bedanken, die ich im Labor und am Institut hatte und auch dafür, dass sie mir bei allen möglichen Fragen immer wieder weitergeholfen haben.

Bei meinen Freunden möchte ich mich insbesondere für die emotionale Unterstützung bedanken, die immer ein offenes Ohr für meine Sorgen und Probleme haben und mich auch immer aufmuntern.

Und zu guter Letzt, der größte Dank gebührt meiner Familie, die mich nun jahrelang während meines gesamten Studiums auf allen Ebenen unterstützt hat und die es mir erst ermöglicht hat hier zu sein wo ich jetzt bin.

Danke!

Contents

1	Abstract	7
2	Introduction	8
2.1	Virology	8
2.2	The envelope glycoprotein	10
2.3	Neutralization of WNV	11
2.4	Comparison of Usutu and West Nile virus envelope glycoprotein	13
2.5	Relevance for humans	14
2.6	Computational approach.....	15
2.7	Single-chain variable fragment (scFv) antibodies	17
3	Aims	19
4	Material and Methods.....	20
4.1	Expression and purification of D-III of USUV A, B, H and WNV	20
4.2	Crystallization of D-III of USUV A and B	21
4.3	Crystal structure of USUV A and B.....	22
4.4	Computational saturated mutagenesis	24
4.5	Cloning of scFvE16-mut _{Arg}	25
4.6	Expression and purification of scFvE16 and scFvE16-mut _{Arg} in inclusion bodies.....	26
4.7	Refolding of scFvE16/scFvE16-mut _{Arg} by dialysis.....	27
4.8	On-column refolding of scFvE16.....	28
4.9	Cloning of scFvE16 for periplasmic expression.....	29
4.9.1	Cloning of pET25b-PelB-scFvE16.....	29

4.9.2	<i>Cloning of pET25b-PelB-scFvE16-spacer</i>	30
4.9.3	<i>Cloning of pETduet-OmpA-scFvE16, pETduet-OmpT-scFvE16 and pETduet-PhoA-scFvE16</i> ..	32
4.10	Cold osmotic shock	35
4.11	Expression of scFvE16 into the bacterial periplasm	36
5	Results	37
5.1	Expression and purification of D-III of USUV A, B, H and WNV	37
5.2	Crystal structure of the D-III of USUV A and B	39
5.3	Computational saturated mutagenesis	41
5.3.1	<i>D-III of USUV A in complex with E16-L52_HY-S30_LR-Y92_LR-T94_LR (E16-mutA)</i>	44
5.3.2	<i>D-III of USUV H in complex with E16-W33_HD-L52_HF-S30_LR-W50_LD (E16-mutH)</i>	46
5.4	Expression and purification of scFvE16 as insoluble inclusion bodies	49
5.5	Expression of scFvE16 into the bacterial periplasm	54
6	Discussion	57
6.1	Structure of D-III of USUV A and B	57
6.2	Computational saturated mutagenesis	58
6.3	Purification of E16 as an scFv antibody	61
7	References	64
8	Appendix: Zusammenfassung	69

1 Abstract

The African mosquito-borne Usutu virus (USUV), a member of the family of *Flaviviridae*, shows antigenic cross-reactivity with West Nile virus (WNV), suggesting a similar structure of their surface proteins. The E16 antibody binds to domain III (D-III) of the Envelope protein (E protein) of WNV and is capable of neutralizing WNV. Thus, we wanted to engineer an USUV specific single-chain variable fragment (scFv) antibody based on the E16 antibody using computational saturated mutagenesis. Such a USUV specific antibody could offer a potential treatment for USUV infections or could be applied in diagnostics. To this end, we purified and solved the crystal structure of the D-III of several USUV strains, which serve as the antigen and as a basis for our computational approach to engineer an USUV specific antibody. By computationally introducing single mutations on the E16 antibody and the subsequent calculation of the difference in energy, we were able to obtain models of E16 antibody mutants with an enhanced binding affinity towards USUV-D-III. Further, we tried to establish a purification protocol for the production of the scFv version of the E16 antibody (scFvE16) in *E. coli*. However, the scFvE16 was expressed only as insoluble inclusion bodies and the yield of soluble scFvE16 after refolding was not satisfactory. Nevertheless, the establishment of such a purification protocol depicts a crucial step for the production and experimental validation of our computationally designed E16 mutants.

2 Introduction

Usutu virus (USUV) was isolated from a mosquito (*Culex neavei*) for the first time in 1959 in South Africa¹. During the following decades, USUV was reported in several African countries and drew attention to itself in Europe for the first time in 2001, where USUV was responsible for the death of blackbirds (*Turdus merula*) and great grey owls (*Strix nebulosa*) within the area surrounding Vienna, Austria². At the present, USUV is thought to have established a transmission cycle involving local bird and mosquito species, enabling the virus to spread through Europe³, where it is currently endemic in several countries⁴. Consequently, USUV has been found in several bird and mosquito species in European countries, such as Hungary⁵, Switzerland⁶, Spain⁷, Belgium⁸ or France⁹, in bats in Germany¹⁰ and serologically in a horse in Serbia¹¹. However, the identification of USUV in frozen samples from dead birds from 1996 found in Italy suggests the introduction of USUV in Europe even prior to 2001¹². Further, 21 cases of USUV infections in humans in Africa and Europe have been reported until today, progressing either with a wide range of symptoms or asymptotically⁴.

2.1 Virology

USUV belongs to the genus of flaviviruses within the family of *Flaviviridae* and is an arthropod-borne virus (arbovirus), spread through an enzootic cycle including mosquitos and birds^{13,14}. As a member of the Japanese encephalitis virus (JEV) serocomplex, USUV is closely related to viruses such as West Nile virus (WNV), Murray Valley encephalitis virus (MVEV) or Saint Louis encephalitis virus (SLEV)¹³. Further human pathogens, such as Zika virus (ZIKV)

and Dengue virus (DENV) within the genus flavivirus also represent close relatives to USUV. Hence, USUV belongs to a number of viruses associated with human health issues.

In general, flaviviruses are spherical shaped viruses covered by a lipid envelope derived from the host cell membrane. The virus particle is 40 – 60 nm in diameter and contains the positive-sense RNA genome with ~11,000 nucleotides (nt) in length. The genomic organization is comparable to other flaviviruses and includes a 5' cap structure, a unique open reading frame (ORF) and two untranslated regions (UTRs); however, there is no 3' poly(A) tail¹⁵. The ORF is predicted to be translated into a 3,434 amino acids long polyprotein, which is post-translationally processed into three structural (capsid, envelope and pre-membrane) and seven non-structural (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) proteins. The structural proteins are encoded on the 5' end of the viral genome and form the viral particle¹⁶. The central core of the particle is composed of capsid protein (C), which is also associated to the viral RNA. The envelope glycoprotein (E) binds to the host cell receptor and mediates membrane fusion past entering the cell via the endosomal pathway¹⁷. The pre-membrane protein (prM) is essential for the assembly of the viral particle and its maturation¹⁸. The function of the non-structural proteins can be presumed based on the similarity to other flaviviruses; NS1 is crucial for the viral genome replication and virion maturation. The small, hydrophobic proteins NS2A, NS2B, NS4A and NS4B are essential for virus assembly and maturation. NS3 is associated with enzymatic activities and acts as a protease, helicase and nucleoside triphosphatase. The N-terminus of NS5 acts as methyltransferase and the C-terminus as RNA-dependent RNA polymerase^{19,20}.

2.2 The envelope glycoprotein

The E protein plays an essential role during viral entry and viral infection in the life cycle of flaviviruses. After binding to the host cell receptor, the virus enters the cell via the endosomal pathway. The acidic environment in the endosome causes structural rearrangement of the E protein leading to the fusion of the viral lipid membrane and the host cell membrane followed by the release of the viral capsid into the cytoplasm¹⁷.

The structure of the E protein of several flaviviruses, such as WNV²¹, ZIKV²² or Dengue virus type 2 (DENV2)²³ has been determined to date. The flavivirus E protein contains predominantly β -strands and is structured into three distinct domains: Domain I (D-I) represents the central part, adopts a globular barrel structure and exhibits a conformational change under acidic conditions in the endosome. Domain II (D-II) is defined by an elongated structure and contains the hydrophobic fusion loop, which is crucial for membrane fusion post endocytosis. The immunoglobulin(Ig)-like domain III (D-III) is assumed to be involved in receptor binding and depicts a major target of neutralizing antibodies. D-I is connected on one side with D-II via four peptide linkers and on the other side with D-III via one peptide linker²⁴.

A recent study on the E protein of USUV (USUV-E) reveals a similar overall structure of the protein. However, in comparison to the E protein of other flaviviruses (e.g. WNV, ZIKV or DENV), USUV-E adopts a more elongated conformation due to a larger angle between D-I and D-II. Therefore, USUV-E depicts the most elongated E protein structure solved. Structural investigation demonstrated D-I folding into a compact barrel structure containing nine β -strands and one α -helix. Interestingly, this helix has been reported to be present only in E proteins from members of the JEV serocomplex (e.g. WNV or JEV). D-II is composed out

of 8 β -strands and 4 helices, exhibiting an elongated conformation and D-III consists of 6 β -strands adopting an Ig-like fold²⁵.

2.3 Neutralization of WNV

Several monoclonal antibodies (mAbs) with strongly neutralizing effects specific for the E protein of WNV (WNV-E) were generated. Yeast surface display was applied to map the domains involved in antibody binding, which revealed that most mAbs specifically recognize D-III as an epitope. One of these generated antibodies, E16, was able to protect mice from lethal WNV challenge five days post infection. Additionally, plaque reduction neutralization assays demonstrated the ability of E16 to neutralize all WNV strains with a 50% plaque reduction neutralization titer (PRNT₅₀) of 4-18 ng and a 90% plaque reduction neutralization titer (PRNT₉₀) of 53-297 ng. Subsequently, E16 was humanized to exploit its full potential as a therapeutic against WNV²⁶.

A crystal structure of the E16 antigen-binding (Fab) fragment in complex with D-III at 2.5 Å resolution revealed the structural basis of antibody binding and WNV neutralization (PDB ID: 1ZTX) (Fig. 1.A). Generally, a tight network of hydrogen bonds dominates the protein-protein binding interface. D-III interacts with E16 via four discontinuous segments with a total of 16 residues involved in the formation of contacts: Tyr302, Ser306, Lys307, Ala308, Phe309, Thr330, Gly31, Thr332, Asp333, Ala365, Thr366, Ala367, Asn368, Gly389, Glu390 and Gln391 (Fig. 1.B)²⁷. Four of these 16 residues located at the core of the binding site are crucial for antibody recognition: Ser306, Lys307, Thr330 and Thr332²⁶. Sequence analysis displayed a nearly complete conservation of the residues structurally defined as E16 epitope among several WNV strains, constituting the ability of the E16 antibody to neutralize various WNV strains. However, sequence comparison with other flaviviruses (e.g. JEV, SLEV

or DENV) shows high diversity within the four segments. Consequently, E16 is generally not capable of cross-neutralization of other flaviviruses²⁷. Nevertheless, plaque reduction neutralization assays in a more recent study demonstrated cross-reactivity between WNV and USUV subtype CAR-1969 with a E16 antibody titer of <5000 ng. Three other USUV strains tested, were not neutralized by the E16 antibody even at high concentration²⁸.

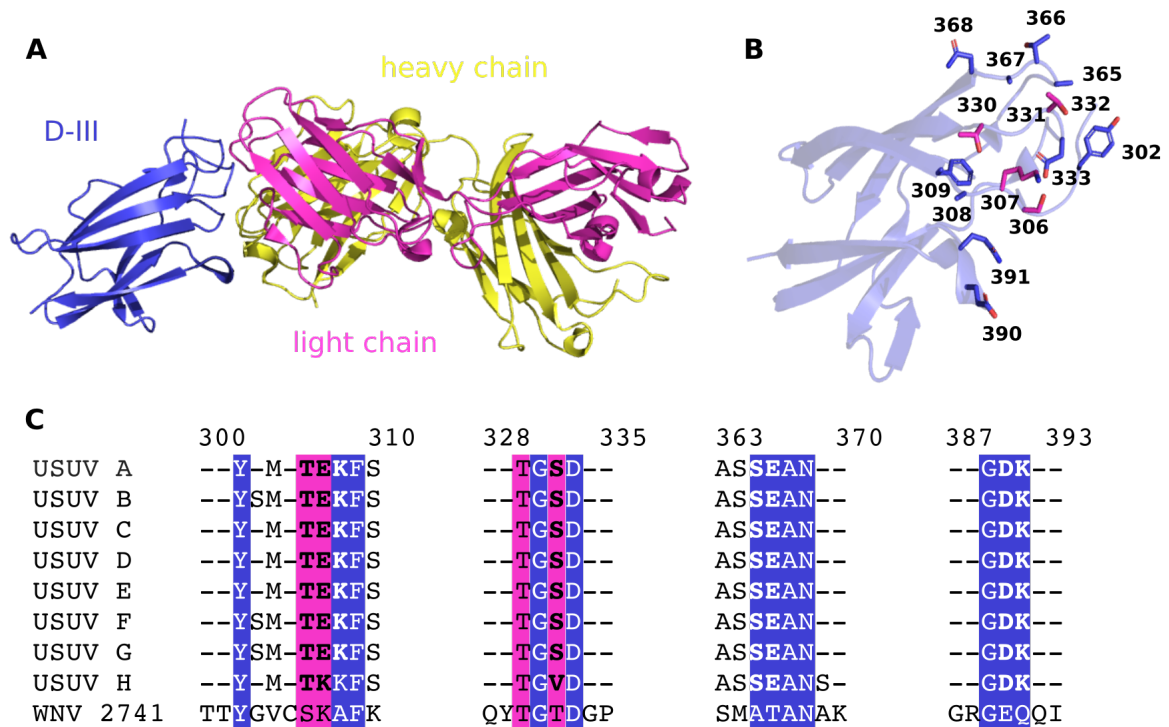


Fig. 1. Residues involved in the binding of E16 in West Nile virus compared to those of different Usutu virus groups. (A) Crystal structure of WNV-D-III in complex with the E16 Fab fragment (PDB ID: 1ZTX). (B) The structure of WNV-D-III showing the residues involved in the binding. The positions that are essential for the binding to E16 are colored in magenta. (C) The alignment of the segments comprising the residues involved in contacting the E16 of WNV-D-III with the D-III of all the USUV groups shows that several residues in crucial positions differentiate (bold). Residues involved in the binding are highlighted blue; residues crucial for binding are highlighted magenta.

2.4 Comparison of Usutu and West Nile virus envelope glycoprotein

WNV represents one of the closest relatives to USUV within the genus flavivirus²⁹. Cross-reaction between USUV and WNV in plaque reduction neutralization assays indicates similar antigenic features of E protein D-III. The amino acid sequence comparison of the E protein between seven USUV strains, USUV subtype CAR-1969 and WNV 2741 shows a sequence identity of 98-100 % among the seven different USUV strains, 95-96% when comparing these seven strains to USUV subtype CAR-1969 and 77% and 78% when comparing WNV 2741 to the seven USUV strains and USUV subtype CAR-1969, respectively²⁸.

In general, based on the D-III amino acid sequence, all USUV strains can be clustered into eight groups, USUV A – H (Fig. 1.C). The comparison of the D-III amino acid sequence of these eight USUV groups with each other and with WNV 2741 demonstrates the following: When comparing only the 16 residues relevant for E16 binding, USUV group A – G show identical residues at these positions with eight identical versus eight non-identical amino acids compared to WNV 2741. However, USUV H, which contains only USUV subtype CAR-1969 as a member, differs in two positions (307 and 332) when compared to the other USUV groups and shows nine identical versus seven non-identical residues when compared to WNV 2741 (unpublished data). Thus, USUV subtype CAR-1969 (USUV H) is more similar to WNV 2741 than USUV strains from the remaining groups A – G, which potentially explains the neutralization with E16 and, consequently, makes E16 an ideal basis for the design of a new antibody binding USUV.

2.5 Relevance for humans

So far, most USUV infections affected animals, especially birds as the example of Austria shows, where USUV caused the death of several bird species in 2001². However, in 1981 in the Central African Republic, the first human case of an USUV infection was reported with the patient suffering from fever and rash³⁰. During the following decades, several other human USUV infections were reported in Africa, as well as in Europe, with a total of 21 cases so far⁴. Infections may proceed asymptotically; however, in some cases, symptoms vary broadly from moderate symptoms (e.g. fever, rash) to severe neurological disorders (e.g. meningoencephalitis)^{14,31}.

The second case was described in 2004 in Burkina Faso with a patient exhibiting only moderate symptoms, similar to the first incident¹⁴. Nevertheless, two patients with severe neurological symptoms and suffering from meningoencephalitis were reported in 2009 in Italy. Both patients had been affected by an immune-deficiency previously; one patient was suffering from diffuse large B cell lymphoma and the second patient obtained an orthotopic liver transplant prior to the infection^{32,33}. In Italy, further cases were reported in which USUV was detected in cerebrospinal fluid from three patients with acute meningoencephalitis³⁴. Another three patients with an USUV-related neuroinvasive infection were reported in Croatia in 2013³⁵. Additionally, USUV was found in samples from healthy blood donors when screened for WNV in Germany³⁶, as well as in Austria³⁷.

Currently, USUV does not represent a threat to human health and only a few cases of human USUV infections have been reported compared to other flaviviruses. Nevertheless, USUV has spread in several European and African countries, similar to WNV and is presently considered as endemic⁴. Nonetheless, fast propagation of flaviviruses depicts a major risk for

human health, as the Zika virus outbreak in 2015/16 emphasizes. ZIKV was first isolated from a rhesus monkey in 1947 and although the virus was widely distributed in Africa and Asia already in the 20th century, ZIKV received only little medical attention^{38,39}. Only about a dozen cases had been reported with patients suffering from mild febrile illness⁴⁰. In 2015/16, however, not only 440,000 to 1,300,000 suspected ZIKV infections were reported in Brazil in connection with common symptoms, such as fever, rash, arthralgia or conjunctivitis, but also a significant increased number of newborns with microcephaly was documented^{41,42}. As a consequence, the surveillance of the development and distribution of seemingly harmless flaviviruses, such as USUV, is essential.

2.6 Computational approach

Protein-Protein interactions (PPIs) are essential for several biological processes, such as signal transduction, transcription/translation, the assembly of huge molecular machineries and many more. Consequently, the study of PPIs, their behavior, evolution and manipulation depicts a central topic not only for fundamental biology, but also for applied research, especially in the field of drug design. Understanding the molecular forces dominating binding affinity and binding specificity in PPIs still depicts a challenging task⁴³. Advances in molecular biology over the last years facilitated the determination of the contribution of single residues to binding energetics⁴⁴. Nevertheless, recently developed computational methods to predict binding affinity experienced difficulties to distinguish between native PPIs and computationally designed PPIs⁴⁵. Additionally, even the determination of features that differentiate between strong, medium and weak PPIs remains challenging⁴⁶.

Nonetheless, several useful approaches have been developed for the investigation of PPIs. One option depicts the estimation of the PPI landscape; based on the introduction of mutations within the binding interface, an enhancement or decrease in binding affinity can be estimated⁴⁴. However, the experimental investigation of PPI landscapes is time consuming and costly due to a large number of mutants that need to be assessed for binding affinity. Consequently, a computational approach represents an attractive alternative for the generation of binding landscapes and recent advances in the prediction of affinity-enhancing mutations using structure-based computational methods suggest a solid approximation to reality⁴⁷. Further, compared to experimental PPI design and engineering, such structure-based computational methods depict a more rational approach⁴⁸.

Computational saturated mutagenesis developed by Sharabi et al.⁴⁹ represents such a method enabling the exploration of PPI landscapes, the investigation of PPI evolution and the prediction of mutations with affinity-enhancing or affinity-decreasing effects. The crystal structure of PPIs represents the basis for computational calculations. Single mutations are introduced on one of the binding partners at positions located within the binding interface, followed by the estimation of the binding affinity. Accordingly, the intermolecular energy (ΔE_{inter}) is calculated by subtracting the energy of the single protein chains from the complex energy. Subsequently, ΔE_{inter} of the mutant is subtracted from the wild type to calculate the difference in free energy of binding ($\Delta \Delta G_{\text{bind}}$). The dissociation constant (K_d), and therefore, the binding affinity is directly related to ΔG_{bind} allowing the estimation of affinity-enhancing and affinity-decreasing mutations. As a consequence, computational saturated mutagenesis enables the identification of positions within the binding interface that are evolutionary optimized or mutations with a potential for further optimization, respectively. In the case of

protein design and evolution, such computational methods represent a tool to narrow down the number of possibilities that need to be tested experimentally compared to conventional *in vitro* approaches, such as phage or ribosome display, which are based on the selection of binders from a large library of protein mutants⁴⁹.

2.7 Single-chain variable fragment (scFv) antibodies

Compared to a tetrameric antibody constituted of two heavy chains and two light chains with a molecular weight around 150 kD, an scFv antibody is a fusion protein solely composed of the variable domains of the heavy chain (V_H) and the light chain (V_L), fused via a flexible linker (Fig. 2). Usually, such a linker contains 10-25 amino acids and is rich in glycine and serine to ensure flexibility and occasionally comprises glutamic acid or lysine to improve solubility. Consequently, an scFv antibody maintains its original specificity, but is tremendously reduced in size with a molecular weight around 27 kD^{50,51}.

The reduction in size and the presence as a single peptide depicts several advantages. The production is facilitated and the expression and purification from *E. coli* is possible due to reduced complexity⁵². Additionally, as a single peptide, scFvs simplify the application of phage display, since the antigen-binding domain can be expressed and displayed on the surface of the phage⁵³. Ultimately, scFvs exhibit lower immunogenicity compared to complete antibody constructs, promoting their use in therapeutic and diagnostic applications⁵¹.

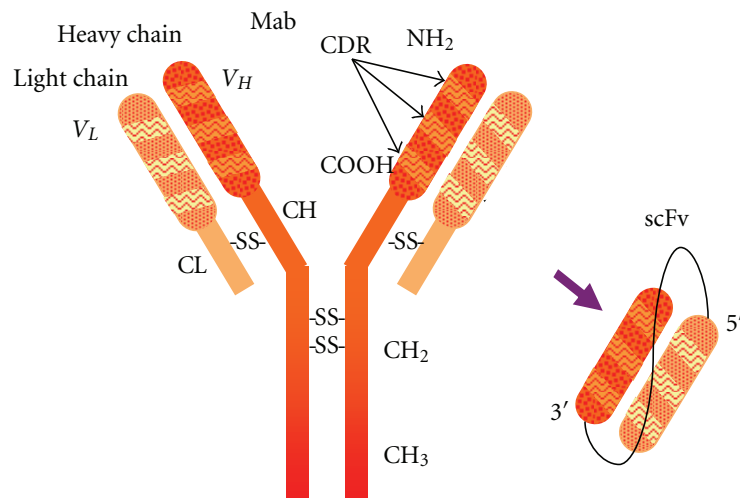


Fig. 2. Antibody model. The model shows the composition of a tetrameric antibody (Mab) compared to a recombinantly generated scFv. The antibody constitutes of two heavy chains comprising the variable (V_H) and constant (C_H) domain and two light chains comprising a variable (V_L) and constant (C_L) domain. The complementary-determining regions (CDRs) are located on the variable domains. The chains are linked to each other via disulfide bonds. On the contrary, the scFv displays a fusion protein composed of only the variable domains of the light and heavy chain. Image from Ahmad et al. *Clinical and Developmental Immunology*. (2012)

3 Aims

The aim of the project was to generate a Usutu virus specific antibody. We wanted to make use of the structural similarity and amino acid sequence similarity between the E protein of USUV and WNV. Therefore, we used the WNV neutralizing antibody, E16, as a basis to design and engineer a USUV specific antibody using computational saturated mutagenesis. Compared to conventional methods for antibody generation (e.g. immunization of animals or phage display), a computational approach allows a more rational design, followed by the experimental assessment of only a few potential binding antibodies. Accordingly, the project further depicts an opportunity for the investigation of alternative approaches for antibody design and engineering.

Computational saturated mutagenesis is a structure-based method and a crystal structure is required to perform calculations of mutations located within the binding interface of a protein complex with binding affinity-enhancing or -decreasing effects. The structure of WNV-D-III in complex with the E16 Fab fragment was solved previously by Nybakken et al²⁷. However, we wanted to replace the WNV-D-III within this structure with the USUV-D-III to calculate potential USUV specific antibodies. Therefore, the solution of the structure of USUV-D-III was crucial for the project.

Finally, for the assessment of these potential USUV specific antibodies, the establishment of a protocol for antibody production and purification was necessary. Consequently, we engineered an scFv antibody version of the humanized E16 antibody and used this construct to set up a reliable purification procedure.

4 Material and Methods

4.1 Expression and purification of D-III of USUV A, B, H and WNV

The pET11d vector containing the coding sequence (residues 299-400) of E protein D-III of USUV A, B, H and WNV (USUV-A-D-III, USUV-B-D-III, USUV-H-D-III and WNV-D-III) was used for expression and purification from *E. coli*. The *E. coli* strain soluBL21 was used for the expression and purification of USUV-A-D-III, USUV-B-D-III and USUV-H-D-III; the *E. coli* strain BL21 for the expression and purification of WNV-D-III.

The recombinant plasmid was transformed via heat shock (42°C, 45 sec) into *E. coli* soluBL21 and BL21 competent cells. A single colony was picked to inoculate 80 ml Luria-Bertani (LB) medium for overnight growth at 37°C. The subsequent 80 ml over night culture was used for the inoculation of 2 L of fresh LB medium and cells were allowed to grow at 37°C. When the OD₆₀₀ reached 0.5-0.8, the expression was induced with 0.5 mM isopropylthiogalactoside (IPTG) and the culture was incubated at 23°C, 180 rpm for 4 h. Cells were harvested at 6,720 g for 15 min at 4°C and re-suspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA, 5% glycerol). Cell disruption was performed with an EmulsiFlex-C3 homogenizer.

First, the lysate was washed twice (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5% glycerol, 1% Triton X-100, 5 mM DTT), followed by a washing step to remove the DNA and RNA (50 mM Tris-HCl, pH 8.0, 1 M NaCl 5% glycerol) from the pellet. The re-suspended pellet was centrifuged after each step at 48,000 g for 15 min at 4°C and the supernatant was discarded. After another washing step with lysis buffer to remove excess salt, the pellet was re-

suspended in solubilization buffer (50 mM Tris-HCl, pH 8.0, 6 M urea, 5% glycerol) and spun over night at 4°C to denature and solubilize D-III from the inclusion bodies.

The solubilized D-III was separated from the insoluble fraction by centrifugation at maximum speed (48,000 g) for 15 min at 4°C. The supernatant was loaded into 3,500 MWCO dialysis tubes (Thermo Fisher Scientific) after adjusting the concentration to 1 mg/ml for refolding. A two-step dialysis at 4°C was performed: (1) Dialysis against 1.5 L refolding buffer A (500 mM Tris-HCl, pH 8.0, 150 mM NaCl, 3 M urea, 5% glycerol) and (2) dialysis against 1.5 L refolding buffer B (500 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5% glycerol). For size-exclusion chromatography (SEC), the samples were dialyzed against 1.5 L SEC buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl) at 4°C. Each dialysis buffer was exchanged once after at least 1 h of dialysis. SEC was performed using a HiLoad 26/600 S75 column (GE Healthcare) equilibrated with SEC buffer. Samples corresponding to 1 L expression volume were concentrated to 15 ml after dialysis and loaded onto the SEC column. Fractions were analyzed with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE). The fractions containing the protein were pooled and desalted (10 mM Tris-HCl, pH 8.0 at 4°C). The sample was concentrated up to 13 mg/ml using a 3,000 MWCO concentrator (Amicon) for crystallization trials.

4.2 Crystallization of D-III of USUV A and B

The purified USUV-A-D-III was concentrated to 9 mg/ml, 10 mg/ml and 11 mg/ml using a 3,000 MWCO concentrator (Amicon); the purified USUV-B-D-III was concentrated to 8 mg/ml, 10 mg/ml and 13 mg/ml. Crystallization trials were performed with 200 nl and 400 nl protein, respectively, at the different concentrations mixed with 200 nl of reservoir solution and equilibrated against 70 µl reservoir solution by sitting drop vapor diffusion at

room temperature. Crystallization screening was performed using optimized buffer conditions based on the conditions that lead to the crystallization of USUV-H-D-III (unpublished data) (Tab. 1) and the commercially available crystal screen from Hampton Research.

Tab. 1. Buffer composition used for crystallization trials.

Condition	Salt	Buffer	Precipitant
1	0.2 M Ammonium acetate	0.1 M Sodium citrate tribasic dihydrate	30% Polyethylene glycol 4k
2		0.1 M Sodium citrate tribasic dihydrate	30% Polyethylene glycol 8k
3		0.1 M Sodium cacodylate trihydrate	30% Polyethylene glycol 6k
4		0.1 M MES monohydrate	30% Polyethylene glycol 8k
5	0.2 M Ammonium sulfate	0.1 M Sodium cacodylate trihydrate	30% Polyethylene glycol 8k
6			30% Polyethylene glycol 6k
7			30% Polyethylene glycol 4k
8			30% Polyethylene glycol 6k
9	0.2 M Ammonium sulfate	0.1 M MES monohydrate	30% Polyethylene glycol 8k
10			30% Polyethylene glycol 4k
11		0.1 M Sodium citrate tribasic dihydrate	30% Polyethylene glycol 6k
12		0.1 M Sodium citrate tribasic dihydrate	30% Polyethylene glycol 4k

4.3 Crystal structure of USUV A and B

Diffraction-quality crystals of USUV-A-D-III were obtained in a condition consisting of 0.2 M lithium sulfate monohydrate, 0.1 M Tris-HCl (pH 8.5) and 30% (w/v) polyethylene glycol 4k and of USUV-B-D-III in a condition consisting of 0.1 M sodium citrate tribasic dihydrate (pH 5.6), 0.2 M ammonium acetate and 30% (w/v) polyethylene glycol 4k.

Diffraction data were collected at European Synchrotron Radiation Facility (ESRF). Crystals were cryo-protected by soaking in reservoir solution containing 20% (v/v) glycerol prior to flash-cooling in liquid nitrogen.

The structure was determined by molecular replacement using Phaser⁵⁴ and the previously determined structure of USUV-H-D-III served as a homology model (unpublished data). Auto-building of the model and refinement were carried out with Phenix^{55,56} and Coot⁵⁷ software. Stereochemistry and structure quality was analyzed using MolProbity⁵⁸.

Data collection and refinement statistics are summarized in Table 2.

Tab. 2. Crystallographic data.

	D-III USUV A	D-III USUV B
DATA COLLECTION		
Source	ESRF	ESRF
Wavelength (Å)	0.968	0.972
Resolution (Å)	100.0 – 1.66 (1.7 – 1.66)	n.a.
Space group	C 2 2 2 ₁	P 2 2 2 ₁
Unit cell (Å, °)	a = 31.2, b = 60.4, c = 85.4 α=β=γ=90	a = 47.06, b = 55.54, c = 128.47 α=β=γ=90
Molecules/a.u.	1	3
Total reflections	64,421 (3,316)	n.a.
Unique reflections	8,674	n.a.
Completeness (%)	98.4 (85.2)	n.a.
R _{merge}	6.2 (129.6)	n.a.
R _{sigma}	n.a.	n.a.
Redundancy	6.46 (4.46)	n.a.
<I/σ _I >	16.82 (1.18)	n.a.
Wilson B factor (Å ²)	29	n.a.
REFINEMENT		
R _{work}	18.25	19.39
R _{free}	22.12	22.30
R.m.s.d. bonds (Å)	0.007	0.007
R.m.s.d. angles (°)	0.935	0.898
RAMACHANDRAN PLOT		
favoured (%)	98.99	96.70
allowed (%)	1.01	3.30
outliers (%)	0.00	0.0

4.4 Computational saturated mutagenesis

The WNV-D-III in complex with the E16 Fab fragment (PDB ID: 1ZTX) was replaced with USUV-A-D-III and USUV-H-D-III using PyMol⁵⁹. The docking was optimized with RosettaDock⁶⁰. The emerging model of USUV-A-D-III and USUV-H-D-III in complex with E16 Fab fragment was used for the remaining calculations. Water molecules and heteroatoms were removed and hydrogen atoms were added using MolProbity⁵⁸. The structure was energy minimized with 50 steps of atom displacement using conjugated gradient minimization. The binding interface positions were defined as all E16 residues within a 5 Å radius of the 16 residues essential for WNV-D-III to bind to E16. Positions classified as binding interface positions were optimized using all 20 amino acids as possible candidates for mutations. A shell was defined around the interfacial residues to allow repacking of the side chain conformations within this shell after the introduction of each mutation. A shell cutoff of 3.5 Å was used for the interfacial residues of E16 and 4.5 Å for the interfacial residues of USUV-A-D-III and USUV-H-D-III. The total energy was minimized using the dead-end elimination (DEE) algorithm⁶¹ and the total energy (E_{tot}) of the complex was determined. The energy of the two single chains was calculated using the same parameters. Subsequently, the intermolecular energy was obtained by subtracting the energy of the single chains from the total energy:

$$E_{\text{inter}} = E_{\text{tot}} - E_{\text{D-III}} - E_{\text{E16}}$$

The calculation was repeated for the wild type complex to obtain $\Delta\Delta G_{\text{bind}}$ by subtracting the total energy of the mutated complex from the total energy of the wild type complex:

$$\Delta\Delta G_{\text{bind}} = E_{\text{inter}}^{\text{WT}} - E_{\text{inter}}^{\text{mut}}$$

The $\Delta\Delta G_{\text{bind}}$ values obtained for each position were used to estimate the binding affinity of potential E16 mutants binding USUV-A-D-III and USUV-H-D-III.

All calculations were preformed in collaboration with the Shifman lab at the Hebrew University of Jerusalem.

4.5 Cloning of scFvE16-mut_{Arg}

Site-directed mutagenesis was performed with the plasmid pETduet-scFvE16 (Fig. 3.A) to introduce an arginine at position 92 and 94 of the light chain (Y92_LR and T94_LR) with the following primers: 5'-CAG CAG CAC CGC ACC CGC CCG CTG ACC-3' and 5'-GCA ATA ATA CAG GGC CAG ATC CTC G-3' (mutation is underlined) using the Q5[®] Site-Directed Mutagenesis kit (New England BioLabs[®]) (Fig. 3.B).

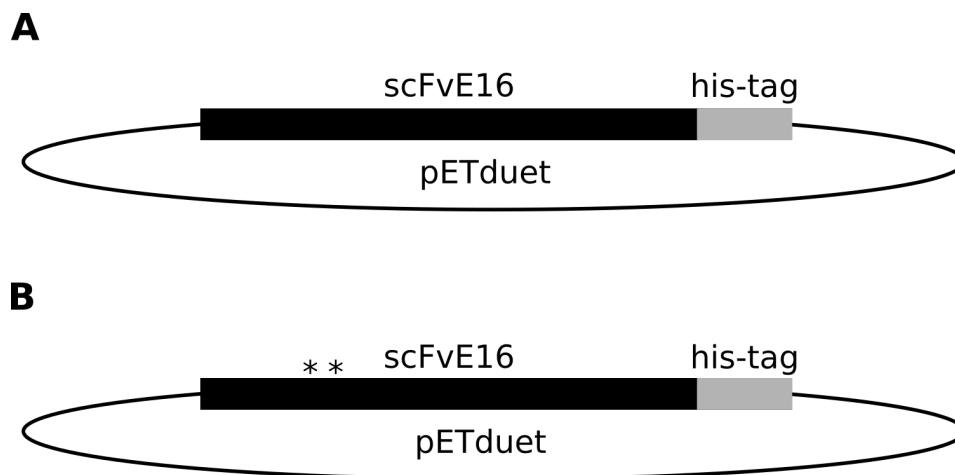


Fig. 3. scFvE16 constructs. (A) pETduet vector containing the coding sequence for scFvE16 including a C-terminal his-tag. (B) Asterisks highlighting the introduction of an arginine at position 92 and 94 of the light chain.

The plasmid was amplified by PCR (Tab. 3 and 4) and the emerging PCR product was ligated using the KLD enzyme mix (Tab. 5). The KLD reaction was incubated for 5 min at room temperature prior to transformation into *E. coli* Top10 competent cells via heat shock (42°C, 45 sec). Single colonies were picked to inoculate 8 ml of LB medium and incubated at 37°C

over night. The plasmid was extracted using the Wizard® Plus SV Minipreps DNA purification system (Promega) and the sequence of pETduet-scFvE16-mut_{Arg} was confirmed by Sanger sequencing (Microsynth).

Tab. 3. PCR reaction.

Reagent	Volume (μl)
Q5 buffer	10
GC enhancer	10
dNTPs	1
Forward primer (100 μM)	0.5
Reverse primer (100 μM)	0.5
Plasmid (150 ng/μl)	0.1
Q5 Polymerase	0.5
Nuclease-free water	27.4

Tab. 4. Cycling conditions.

Step	Temperature (°C)	Time (sec)
Initial denaturation	98	30
25x	Denaturation	10
	Annealing	20
	Elongation	210
Final extension	72	120
Hold	10	-

Tab. 5. KLD reaction.

Reagent	Volume (μl)
KLD reaction mix	5
KLD enzyme mix	1
PCR product	1
Nuclease-free water	3

4.6 Expression and purification of scFvE16 and scFvE16-mut_{Arg} in inclusion bodies

The pETduet vector containing the scFvE16 and scFvE16-mut_{Arg} construct with a C-terminal hexa histidine-tag (his-tag) was transformed into *E. coli* BL21 competent cells via heat shock (42°C, 45 sec). A single colony was picked to inoculate 80 ml LB medium and incubated at 37°C over night. The over night culture was used to inoculate 2 L of fresh LB

medium and the cells were incubated at 37°C. The expression was induced with 0.5 mM IPTG when the OD₆₀₀ reached 0.6-0.8. The cells were harvested at 6,720 g for 15 min at 4°C and re-suspended in lysis buffer (50mM Tris-HCl, pH 8.0, 1 mM EDTA). The cells were lysed by sonication and subsequent centrifugation at 30,000 g for 20min at 4°C (repeated three times in total). The lysed cells were washed with washing buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 M urea) with and without 2% Triton X-100. The re-suspended pellet was centrifuged after each washing step at 30,000 g for 20 min at 4°C and the supernatant was discarded. In a final step, the sample was re-suspended in solubilization buffer (50 mM Tris-HCl, pH 8.0, 7 M urea, 10 mM DTT) and spinned at 4°C over night to solubilize scFvE16/scFvE16-mut_{Arg} from the inclusion bodies. The sample was centrifuged at 30,000 g for 20 min at 4°C to remove insoluble precipitants.

4.7 Refolding of scFvE16/scFvE16-mut_{Arg} by dialysis

The solubilized scFvE16/scFvE16-mut_{Arg} was refolded by stepwise dialysis at 4°C using 7,000 MWCO dialysis tubes (Thermo Fisher Scientific) after adjusting the concentration to 1 mg/ml. The sample was dialyzed against 1.5 L of dialysis buffer A (500 mM Tris-HCl, pH 8.0, 150 mM NaCl, 3.5 M urea, 5% glycerol) and 1.5 L of dialysis buffer B (500 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5% glycerol). Each buffer was exchanged once after at least 1 h of dialysis. After dialysis, the sample was centrifuged at 30,000 g for 20 min at 4°C and was analyzed on SDS PAGE.

Refolding was performed with different refolding buffers on a smaller scale with only scFvE16. Therefore, 8 ml of solubilized scFvE16 at 0.5 mg/ml were loaded into 7,000 MWCO dialysis tubes (Thermo Fisher Scientific). In a first dialysis step, the sample was dialyzed against 0.5 L dialysis buffer A. The sample was split into two 4 ml aliquots and dialyzed

against refolding buffer 200 (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 500 mM urea, 200 mM arginine, 2 mM reduced glutathione, 1 mM oxidized glutathione, 20% glycerol) and refolding buffer 400 (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 500 mM urea, 400 mM arginine, 20% glycerol), respectively. In a final dialysis step, both samples were dialyzed against PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4). Each buffer was exchanged once after at least 1 h of dialysis at 4°C. After each dialysis step, the sample was centrifuged at 20,000 g and a small aliquot was kept for SDS PAGE analysis. Both samples were loaded on a S200 Increase 10/300 column (GE Healthcare) equilibrated with PBS.

4.8 On-column refolding of scFvE16

HisTrap[™] High Performance 5 ml column (GE Healthcare) was equilibrated with 5 CV (column volumes) binding buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 7 M urea). Solubilized scFvE16 was bound to the column and washed with 10 CV binding buffer. The urea concentration was gradually reduced by the substitution with refolding buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 400 mM arginine) over 5 CV at 0.1 ml/min flow rate from 7 to 0 M urea. The column was washed with 10 CV refolding buffer prior to elution with 2 CV of elution buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 500 mM imidazole, 400 mM arginine). Fractions containing protein were pooled and analyzed on SDS PAGE. Different refolding buffers were tested containing 200 mM arginine instead of 400 mM or detergent instead of arginine (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 1 mM β -mercaptoethanol, 1% Triton X-100). Subsequently, also the elution buffer was supplemented with the according reagents.

4.9 Cloning of scFvE16 for periplasmic expression

4.9.1 Cloning of pET25b-PelB-scFvE16

The plasmid pETduet-scFvE16 and the pET25b(+) vector containing the PelB signal sequence were digested with the restriction enzymes *Hind*III and *Nco*I (New England BioLabs®) (Tab. 6.1 and 6.2). The restriction enzyme digest was incubated at 37°C for 1 h followed by heat inactivation at 65°C for 10 min. The scFvE16 insert and the digested pET25b(+) vector were gel extracted from a 1% agarose gel using the Wizard® SV gel and PCR clean-up system (Promega). Ligation was performed with 100 ng of the purified vector and 56 ng of the purified insert using the T4 DNA ligase (New England BioLabs®) (Tab. 7). The ligation reaction was incubated at room temperature for 20 min, followed by the heat inactivation of the ligase at 65°C for 10 min and transformation into *E. coli* Top10 competent cells via heat shock (42°C, 45 sec). Single colonies were picked to inoculate 6 ml of LB medium and allowed to grow at 37°C over night. The pET25b-PelB-scFvE16 plasmid (Fig. 4) was purified using the Wizard® Plus SV Minipreps DNA purification system (Promega) and the insert was confirmed by Sanger sequencing (Microsynth).

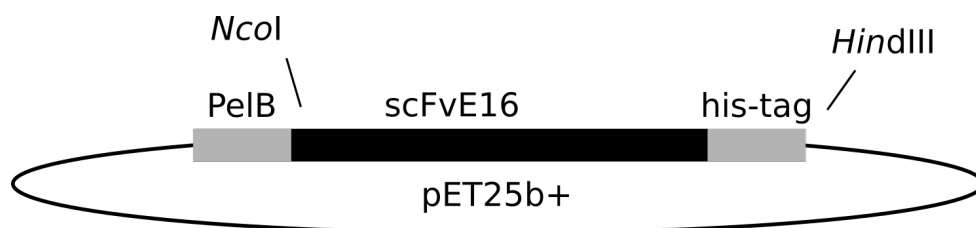


Fig. 4. pET25b-PelB-scFvE16. pET25b(+) vector containing scFvE16 fused to the PelB at the N-terminus and to a his-tag at the C-terminus.

Tab. 6.1. Restriction enzyme digest pET25b(+).

Reagent	Volume (μl)
CutSmart (10x)	10
pET25b(+) (174 ng/μl)	23
<i>Hind</i> III	1
<i>Nco</i> I	1
Nuclease-free water	65

Tab. 6.2. Restriction enzyme digest pETduet-scFvE16.

Reagent	Volume (μl)
CutSmart (10x)	20
pETduet-scFvE16 (160 ng/μl)	125
<i>Hind</i> III	1
<i>Nco</i> I	1
Nuclease-free water	53

Tab. 7. Ligation reaction.

Reagent	Volume (μl)
T4 DNA ligase buffer (10x)	1
T4 DNA ligase	1
Vector (138 ng/μl)	0.7
Insert (258 ng/μl)	0.2
Nuclease-free water	7.1

4.9.2 Cloning of pET25b-PelB-scFvE16-spacer

To increase the distance between the ribosomal binding site (RBS) and the PelB start codon a 35 bp long oligonucleotide was cloned in between using restriction enzyme digest and ligation (Fig. 5).

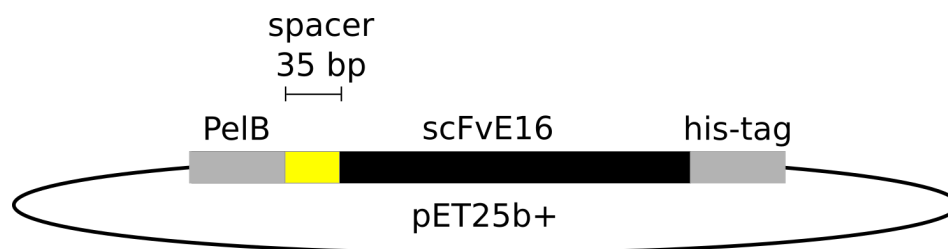


Fig. 5. pET25b-PelB-scFvE16-spacer. pET25b(+) vector containing scFvE16 fused to the PelB at the N-terminus with a 35 nucleotide long spacer inbetween and a C-terminal his-tag.

The two complementary oligonucleotides 5'-TAT GAA ATA CCT GCT GCC GAC CGC TGC TGC TGG CA-3' and 5'-TAT GCC AGC AGC GGT CGG CAG CAG GTA TTT CA-3' including the sticky-end overhangs for ligation (underlined) were phosphorylated using T4

polynucleotide kinase (New England BioLabs®) (Tab. 8). The reaction was incubated at 37°C for 30 min followed by 98°C for 30 sec and 37°C for 5 min to also anneal the two single stranded oligonucleotides. For digestion, the plasmid pET25b-PelB-scFvE16 was incubated with *NdeI* at 37°C for 1 h (Tab. 9). 1 µl of calf-intestinal alkaline phosphatase (CIP) was added to the reaction and incubated at 37°C for 1 h to avoid re-ligation. The enzymes were heat inactivated at 65°C for 10 min. The digested vector was purified from a 1% agarose gel using Wizard® Plus SV Minipreps DNA purification system (Promega). For ligation, 100 ng of purified and digested vector were mixed with 2 µl of the phosphorylation/annealing reaction (Tab. 10) and incubated at room temperature for 1 h, followed by heat inactivation at 65°C for 10 min and transformation into *E. coli* Top10 competent cells via heat shock (42°C, 45 sec). Single colonies were picked to inoculate 8 ml of LB medium and cells were allowed to grow at 37°C over night. The plasmid was purified using the Wizard® Plus SV Minipreps DNA purification system (Promega) and the insertion of the spacer was confirmed by Sanger sequencing (Microsynth).

Tab. 8. Phosphorylation/annealing reaction.

Reagent	Volume (µl)
T4 DNA ligase buffer (10x)	2
T4 polynucleotide kinase	1
Oligonucleotides (1:10)	4
Nuclease-free water	9.5

Tab. 9. Restriction enzyme digest.

Reagent	Volume (µl)
CutSmart (10x)	10
pET25b-PelB-scFvE16 (138 ng/µl)	29
<i>NdeI</i>	1
Nuclease-free water	60

Tab. 10. Ligation reaction.

Reagent	Volume (µl)
T4 DNA ligase buffer (10x)	2
T4 DNA ligase	1
Vector (33 ng/µl)	3
Oligonucleotides	2
Nuclease-free water	12

4.9.3 Cloning of *pETduet-OmpA-scFvE16*, *pETduet-OmpT-scFvE16* and *pETduet-PhoA-scFvE16*

The coding sequence for the signal peptides OmpA, OmpT and PhoA (Tab. 11) was inserted at the N-terminus of the scFvE16. Therefore, the scFvE16 insert was amplified from the *pETduet-scFvE16* vector by two PCR steps with primers generating overhangs corresponding to the signal peptide sequence (Tab. 12) (Fig. 6). The extended insert contained restriction enzyme sites for *Bam*HI and *Hind*III and the primers were designed to delete the start codon of the scFvE16 during the first amplification step.

Tab. 11. Signal sequences.

	Coding Sequence	Signal peptide sequence
OmpA	ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTG GTTTCGCTACCGTAGCGCAGGCC	MKKTAIATAVALAGFATVAQA
OmpT	ATGCGGGCGAACTTCTGGGAATAGTCCTGACAACCCCTA TTGCGATCAGCTCTTTTGCT	MRAKLLGIVLTTPIAISSFA
PhoA	ATGAAACAAAGCACTATTGCAGTGGCACTCTTACCGTTAC TGTTTACCCCTGTGACAAAAGCC	MKQSTIALALLPLLFTPVTKA

Tab. 12. Primer list. Restriction enzyme sites are underlined.

Primer	Sequence
OmpA_1	5' -TGG CAC TGG CTG GTT TCG CTA CCG TAG CGC AGG CCG ACA TCG TGA TGA CC-3'
OmpA_2	5' - AGC CAG GAT CCA TGA AAA AGA CAG CTA TCG CGA TTG CAG TGG CAC TGG CTG GTT-3'
OmpT_1	5' -TCC TGA CAA CCC CTA TTG CGA TCA GCT CTT TTG CTG ACA TCG TGA TGA CC-3'
OmpT_2	5' -AGC CAG GAT CCA TGC GGG CGA AAC TTC TGG GAA TAG TCC TGA CAA CCC CTA-3'
PhoA_1	5' -ACT CTT ACC GTT ACT GTT TAC CCC TGT GAC AAA AGC CGA CAT CGT GAT GAC C-3'
PhoA_2	5' -AGC CAG GAT CCA TGA AAC AAA GCA CTA TTG CAC TGG CAC TCT TAC CGT TAC T-3'
Rev_primer	5' -GGC CGC AAG CTT TTA ATG GTG ATG GTG GTG ATG GCC CT-3'

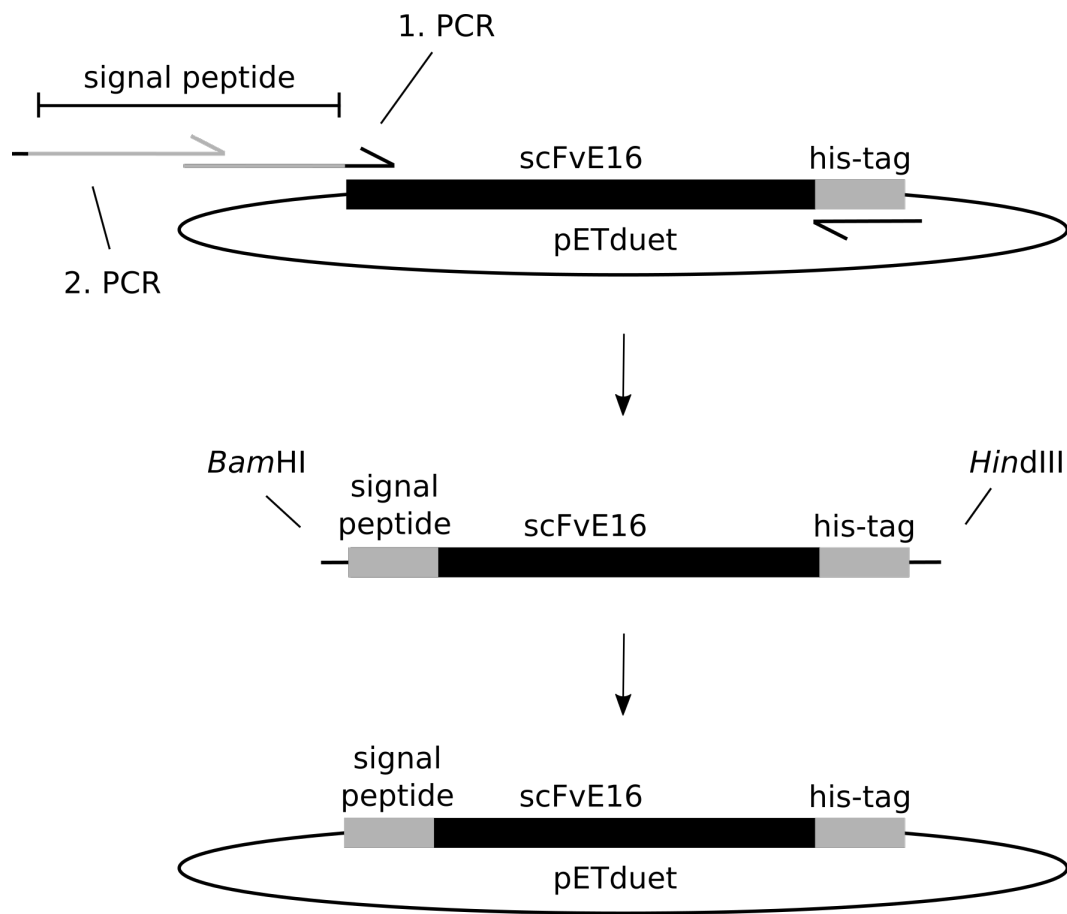


Fig. 6. Cloning strategy for the insertion of a N-terminal signal peptide sequence. The coding sequence for the scFvE16 including a C-terminal his-tag was amplified in two PCR steps to introduce a N-terminal signal peptide sequence and recognition sites for the restriction enzyme digest.

For the first PCR step the following forward primers were used: OmpA_1, OmpT_1 and PhoA_1 (Tab. 13 and 14); to complete the signal peptide sequence, a second PCR step was performed with OmpA_2, OmpT_2 and PhoA_2 as forward primers (Tab. 14 and 15). The reversed primer was the same for all reactions.

Tab. 13. First PCR reaction.

Reagent	Volume (μl)
Q5 buffer	10
dNTPs	1
Forward primer (100 μM)	0.5
Reverse primer (100 μM)	0.5
pETduet-scFvE16 (120 ng/μl)	0.1
Q5 Polymerase	0.5
Nuclease-free water	37.4

Tab. 14. Cycling conditions for both PCR steps.

Step	Temperature (°C)	Time (sec)
Initial denaturation	98	30
25x	Denaturation	10
	Annealing	50
	Elongation	30
Final extension	72	120
Hold	10	-

Tab. 15. Second PCR reaction.

Reagent	Volume (μl)
Q5 buffer	10
dNTPs	1
Forward primer (100 μM)	0.5
Reverse primer (100 μM)	0.5
PCR product	0.1
Q5 Polymerase	0.5
Nuclease-free water	37.4

After each PCR step, the insert was purified from a 1% agarose gel using Wizard® Plus SV Minipreps DNA purification system (Promega). The insert and the empty vector were digested with *Bam*HI and *Hind*III (New England BioLabs®) at 37°C for 1 h (Tab. 16.1 and 16.2). Finally, the plasmid was ligated using the T4 DNA ligase (New England BioLabs®) (Tab. 17). The reaction was incubated at room temperature for 20 min, followed by heat inactivation at 65°C for 10 min and transformation into *E. coli* Top10 competent cells using heat shock

(42°C, 45 sec). Single colonies were picked to inoculate 8 ml of LB medium and incubated at 37°C over night. The plasmid was purified using the Wizard® Plus SV Minipreps DNA purification system (Promega) and the insert of the signal peptide sequence was confirmed by Sanger sequencing (Microsynth).

Tab. 16.1. Restriction enzyme digest pETduet.

Reagent	Volume (μl)
CutSmart (10x)	5
pETduet (222 ng/μl)	9
<i>Hind</i> III	1
<i>Bam</i> HI	1
Nuclease-free water	34

Tab. 16.2. Restriction enzyme digest inserts.

Reagent	Volume (μl)
CutSmart (10x)	5
Insert	50 ng
<i>Hind</i> III	1
<i>Bam</i> HI	1
Add nuclease-free water	50

Tab. 17. Ligation.

Reagent	Volume (μl)
T4 DNA ligase buffer (10x)	2
T4 DNA ligase	1
Vector (20 ng/μl)	2.5
Insert	23 ng
Add nuclease-free water	10

4.10 Cold osmotic shock

Cells from a 30 ml culture were harvested at 3,000 g for 15 min at 4°C, re-suspended in ice-cold TES buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 20% sucrose) and incubated on ice for 30 min. The re-suspended cells were centrifuged at 6,200 g for 15 min at 4°C. The supernatant represents the periplasmic fraction and was retained for further analysis. The pellet was re-suspended in 5 ml of ice-cold 5 mM MgSO₄ and incubated on ice for 30 min. The centrifugation step was repeated. The supernatant represents the cytoplasmic fraction and was used for further analysis. The remaining pellet was re-suspended in water and analyzed as well.

4.11 Expression of scFvE16 into the bacterial periplasm

Plasmids containing the coding sequence for PelB-scFvE16, PelB-scFvE16-spacer, OmpA-scFvE16, OmpT-scFvE16 and PhoA-scFvE16, and the empty vectors pETduet and pET25b(+) were transformed into *E. coli* BL21 competent cells via heat shock (42°C, 45 sec). Single colonies were picked to inoculate 10 ml of LB medium and cells were allowed to grow at 37°C over night. 30 ml of fresh LB medium were inoculated with 1 ml of over night culture and expression was performed under different conditions (Tab. 18). Cells were harvested at 3,000 g for 15 min at 4°C and either re-suspended in lysis buffer or the periplasmic fraction was extracted via cold osmotic shock. The cells re-suspended in lysis buffer were centrifuged at 3,000 g for 15 min at 4°C and aliquots of the lysate and supernatant were used for SDS PAGE and western blot analysis. Fractions obtained from cold osmotic shock were also analyzed via SDS PAGE and western blot.

Tab. 18. Expression conditions.

Condition	IPTG (mM)	Temperature (°C)	Time (h)	Construct
1	0.5	24	4	pET25b(+)-scFvE16
2		24	16	
3		30	4	
4		37	4	
5		16	24	
6	0.1	16	24	
7		24	16	
8		30	4	
9		37	4	
10	0.5	16	18	pET25b(+)-scFvE16-spacer
11		30	4	
12		37	4	
13	0.2	24	5	pETduet-OmpA-scFvE16, pETduet-OmpT-scFvE16,
14		37	5	
15	1	24	5	pETduet-PhoA-scFvE16
16		24	5	

5 Results

5.1 Expression and purification of D-III of USUV A, B, H and WNV

For crystallization trials and future experiments to validate a potential USUV specific antibody, we purified the D-III of USUV A, B, H and WNV (USUV-A-D-III, USUV-B-D-III, USUV-H-D-III and WNV-D-III) from *E. coli*. The overexpression of USUV-A-D-III, USUV-B-D-III, USUV-H-D-III and WNV-D-III in *E. coli* leads to the formation of insoluble inclusion bodies, since no soluble D-III protein was detected in the supernatant of the lysate. However, we were able to recover the D-III from inclusion bodies using a purification protocol including the solubilization with high concentration of urea and refolding by dialysis (Fig. 7.B).

Subsequently, after the lysis of the cells, the pellet was washed several times to remove soluble components, lipids and nucleic acids. During these washing steps, the soluble and insoluble fractions were analyzed on SDS PAGE (Fig. 7.A). The D-III has an expected molecular weight of 11 kD. The gel shows that the D-III was found only in the insoluble fraction during the purification process (Fig. 7.A, lanes 1, 3, 5, 7, 9 and 11). However, after the solubilization of the protein aggregates the D-III was detected in the supernatant (Fig. 7.A, lane 12). Subsequent to the removal of the remaining insoluble components, the sample contained only minor impurities even though no chromatography had been applied.

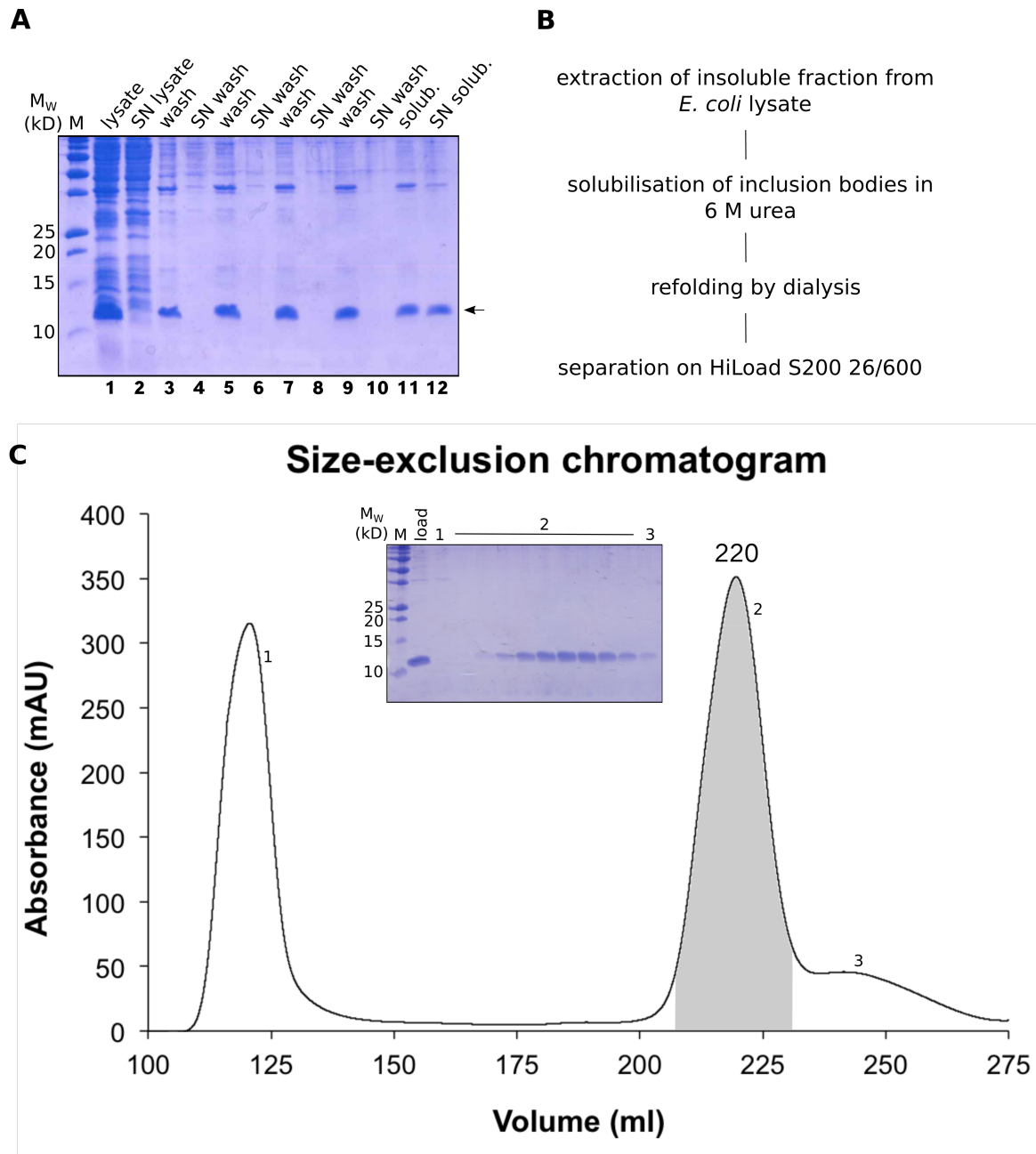


Fig. 7. Purification of D-III of USUV A, B, H and WNV from inclusion bodies. Data from the purification of D-III of WNV is shown representative for all D-III purifications. (A) SDS PAGE showing several washing steps during the purification. In the last step, the D-III was solubilized in 6 M urea. (B) Scheme of the purification process. (C) After refolding, the D-III is applied on size-exclusion chromatography. The peak containing the D-III is indicated in grey. The fractions were analyzed on SDS PAGE.

Refolding of D-III domain was achieved by stepwise dialysis. During dialysis the urea concentration was gradually removed, so the protein could adapt its native fold. Minor

protein precipitates emerging during the refolding procedure were removed and the sample then was applied on a size-exclusion column to separate the D-III from residual impurities (Fig. 7.C). The D-III elutes around 220 ml as a single peak; the fractions were analyzed on SDS PAGE. Fractions containing the D-III were pooled and concentrated. The final yields for the USUV-A-D-III, USUV-B-D-III, USUV-H-D-III and WNV-D-III were between 10 and 24 mg/L culture. USUV-A-D-III and USUV-B-D-III was further used for crystallization trials. USUV-H-D-III and WNV-D-III was flash frozen and stored at -80°C for future experiments.

5.2 Crystal structure of the D-III of USUV A and B

We performed crystallization trials with USUV-A-D-III and USUV-B-D-III based on the conditions that lead to the crystallization of USUV-H-D-III (see material and methods). The structure of USUV-H-D-III was solved in our lab previously with a resolution of 1.19 Å (unpublished data). From our crystallization trials, we obtained several diffracting crystals and were able to collect a full data set for both proteins. We solved the phase problem using molecular replacement with USUV-H-D-III as a homology model and were able to obtain a 1.90 Å structure for USUV-A-D-III and a 1.67 Å structure for USUV-B-D-III (Fig. 8.A).

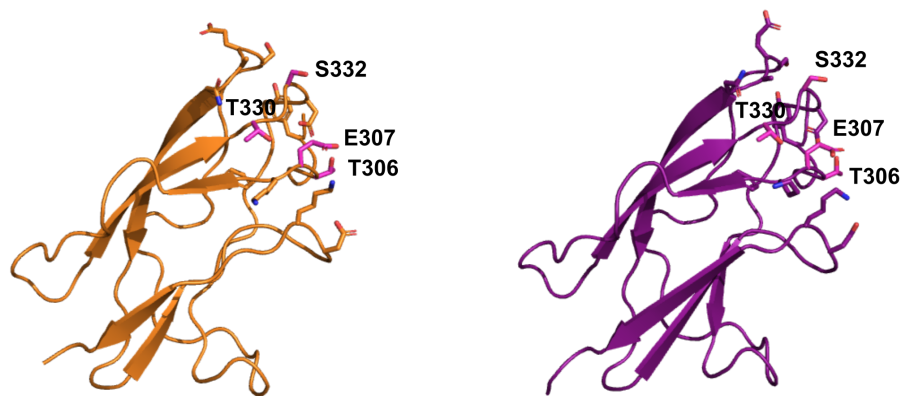
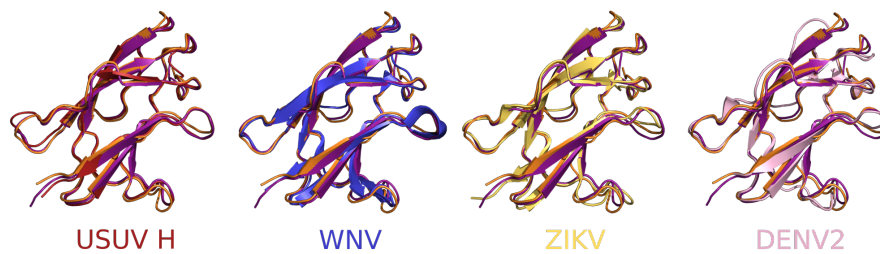
A**B**

Fig. 8. Crystal structure of D-III. (A) Cartoon representation of USUV-A-D-III (orange) and USUV-B-D-III (purple). Side chains analogous to the residues involved in the binding between WNV-D-III and E16 are shown as sticks and the residues crucial for binding are highlighted in magenta. (B) Structural alignment of USUV-A-D-III and USUV-B-D-III with their closest structural relatives. From left to right: D-III of USUV H, D-III of WNV (PDB: 1ZTX), D-III of ZIKV (PDB: 6DFI) and D-III of DENV2 (PDB: 6FLC).

Our analyses show that USUV-A-D-III and USUV-B-D-III adapt an immunoglobulin-like fold, which is generally the case for the D-III of the E protein of viruses belonging to the genus of flavivirus. The structural alignment of USUV-A-D-III, USUV-B-D-III and USUV-H-D-III demonstrates their close structural similarity with RMSD values below 1 Å (Tab. 19). This is not surprisingly, since they share sequence similarities of 95 to 99%. However, both structures superimpose well with WNV-D-III, also with RMSD values below 1 Å and sequence similarities of 74 to 75%. Hence, WNV-D-III represents the closest structural relative of another flavivirus followed by the D-III of Zika virus (ZIKV, PDB: 6DFI) and Dengue virus type

2 (DENV2, PDB: 6FLC) (Fig. 8.B). The comparison with the D-III of ZIKV shows a RMSD value of 1.15 Å (101 residues were aligned out of 101) with a sequence identity of 52% and 0.96 Å (100/102) with a sequence identity of 54% for USUV-A-D-III and USUV-B-D-III, respectively. In the case of the D-III of DENV2 a RMSD value of 1.34 Å (97/98) with a sequence identity of 44% for USUV-A-D-III and 1.41 Å (98/98) with a sequence identity of 45% for USUV-B-D-III were determined. RMSD values were calculated using PDBeFold⁶².

Tab. 19. RMSD values (Å) of D-III of USUV and its closest structural relatives. The values in brackets display the number of aligned residues (N_{align}) compared to the total number of residues of the smaller protein aligned (N_{res}).

	WNV	USUV A	USUV B	USUV H	ZIKV	DENV2
WNV	-	0.94 (100/101)	0.66 (99/101)	0.95 (101/101)	1.03 (99/102)	1.35 (96/98)
USUV A	-	-	0.73 (101/101)	0.84 (101/101)	1.15 (101/101)	1.34 (97/98)
USUV B	-	-	-	0.69 (102/102)	0.96 (100/102)	1.41 (98/98)
USUV H	-	-	-	-	1.17 (101/102)	1.46 (98/98)
ZIKV	-	-	-	-	-	1.43 (97/98)
DENV2	-	-	-	-	-	-

The high structural similarity between the D-III of USUV and WNV allowed us to use the structure of WNV-D-III in complex with the E16 Fab fragment as a basis for our computational approach. However, since the structure of USUV-B-D-III was not available when we started the calculations, USUV-A-D-III and USUV-H-D-III were used as an input.

5.3 Computational saturated mutagenesis

We used the crystal structure of WNV-D-III in complex with the E16 Fab fragment (PDB: 1ZTX) as a basis for the computational saturated mutagenesis calculations. To this end, we replaced WNV-D-III with the structure of USUV-A-D-III and USUV-H-D-III to calculate

potential E16 mutants binding USUV-A-D-III and USUV-H-D-III as an antigen instead of WNV-D-III. Before applying computational saturated mutagenesis, the complex was uploaded to the RosettaDock server⁶⁰ to optimize rigid-body orientation and side chain conformations of our model. Subsequently, we set a radius of 5 Å around the analogous residues within USUV-A-D-III and USUV-H-D-III that are involved in the complex formation of WNV-D-III and E16 (shown by Nybakken et al.²⁷) and determined the E16 positions within this radius as binding interface positions. Mutations to all 20 amino acids were introduced at these positions; for each introduced mutation, $\Delta\Delta G_{\text{bind}}$ was determined. The results are summarized and visualized in a heat map (Fig. 9). Mutations causing the destabilization of the complex at a certain position are colored in red with $\Delta\Delta G_{\text{bind}}$ values of 2.0 kcal/mol or higher; mutations causing minor destabilization are colored in yellow with $\Delta\Delta G_{\text{bind}}$ values between 0.5 and 2.0 kcal/mol. On the contrary, mutations that are considered to be neutral or have the WT amino acid are colored in green with $\Delta\Delta G_{\text{bind}}$ values between -0.5 and 0.5 kcal/mol. Finally, mutations at positions that are increasing the stability of the complex are colored in blue with $\Delta\Delta G_{\text{bind}}$ values of -0.5 kcal/mol or lower. Additionally, the difference in energy of E16 (ΔE_{E16}) was calculated to estimate the potential of each mutation to increase or decrease the stability of the E16.

In general, we observed that mutations at several positions would not increase the stability of the complex and therefore, would not contribute to an enhancement in the binding affinity between the D-III and the E16. Indeed, the WT amino acid reassuringly represents the optimum amino acid at each of these positions. On the contrary, we could determine positions at which the mutation of the WT amino acid to a different amino acid increased the binding affinity. In these cases, it was mostly observed that a stabilization of

the complex was not achieved by one amino acid in particular, but by one of several different amino acids that all resulted in a negative $\Delta\Delta G_{\text{bind}}$ value. Positions especially interesting for the design of a USUV specific antibody are highlighted in the heat map.

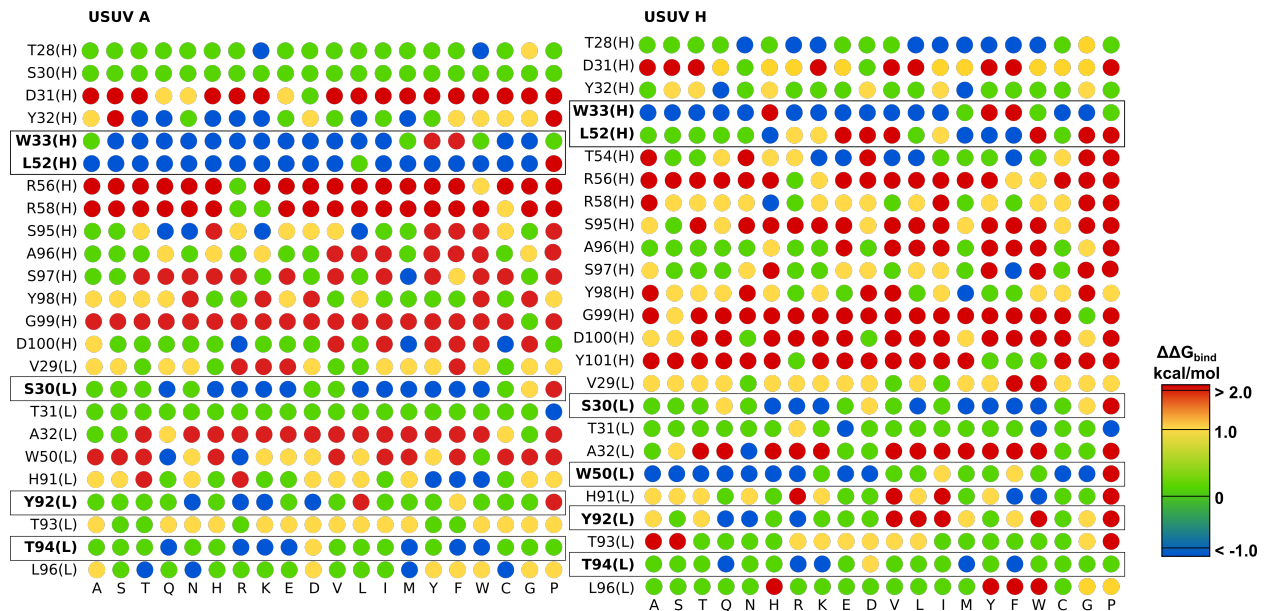


Fig. 9. Heat map after computational saturated mutagenesis. The interface positions in the E16 antibody that were mutated during the calculations are displayed on the left (in brackets H for heavy chain; L for light chain), the amino acid mutations on the bottom. The resulting $\Delta\Delta G_{\text{bind}}$ values are color-coded. Positions especially interesting for increasing the binding affinity are in boxes.

The resulting energy values refer to the complex containing a single mutation. Therefore, it was challenging to estimate how the combination of several mutations would influence the complex stability, as well as the stability of the E16. Thus, we combined mutations with stabilizing effects, but tried to exclude those mutations showing strong destabilization of the E16 itself. Subsequently, only mutations at positions with mild destabilizing effect towards the E16 or stabilizing effects were considered. To obtain estimates on how the combination of mutations would influence the binding interface, we performed another run on the RosettaDock server after introducing selected mutations. This

procedure allowed us to obtain a model with ideal docking between the D-III and the E16 mutant and optimized side chain rotamers, which is crucial for those situated in the binding interface. After the docking, the binding of the D-III towards the E16 was analyzed considering the general orientation of the D-III towards the E16 and the potential polar contacts formed between them.

5.3.1 D-III of USUV A in complex with E16-L52_HY-S30_LR-Y92_LR-T94_LR (E16-mutA)

We obtained one potential E16 mutant binding USUV-A-D-III with the following positions mutated: L52_HY, S30_LR, Y92_LR and T94_LR. The mutations were selected randomly, but with negative $\Delta\Delta G_{\text{bind}}$ values and we tried to choose mutations with a negative ΔE_{E16} or only minor destabilizing effects to make a compromise. Each of these mutations causes, when considered as a single mutation a decrease in $\Delta\Delta G_{\text{bind}}$ of at least -1.7 kcal/mol (Tab. 20). The values for ΔE_{E16} are either negative or below 1.0 kcal/mol.

Tab. 20. Energy values for a potential E16 mutant binding USUV-A-D-III.

	$\Delta\Delta G_{\text{bind}}$	$\Delta E_{\text{complex}}$	ΔE_{E16}
L52_HY	-5.319	-5.262	-0.178
S30_LR	-3.257	-3.486	0.082
Y92_LR	-1.704	-0.562	0.925
T94_LR	-2.520	-1.925	0.916

After introducing the mutations, we uploaded USUV-A-D-III in complex with E16-mutA to the RosettaDock server and analyzed the binding interface. First, the model shows that USUV-A-D-III interacts with the E16-mutA via the four binding loops as the D-III analogue of WNV (Fig. 10.A).

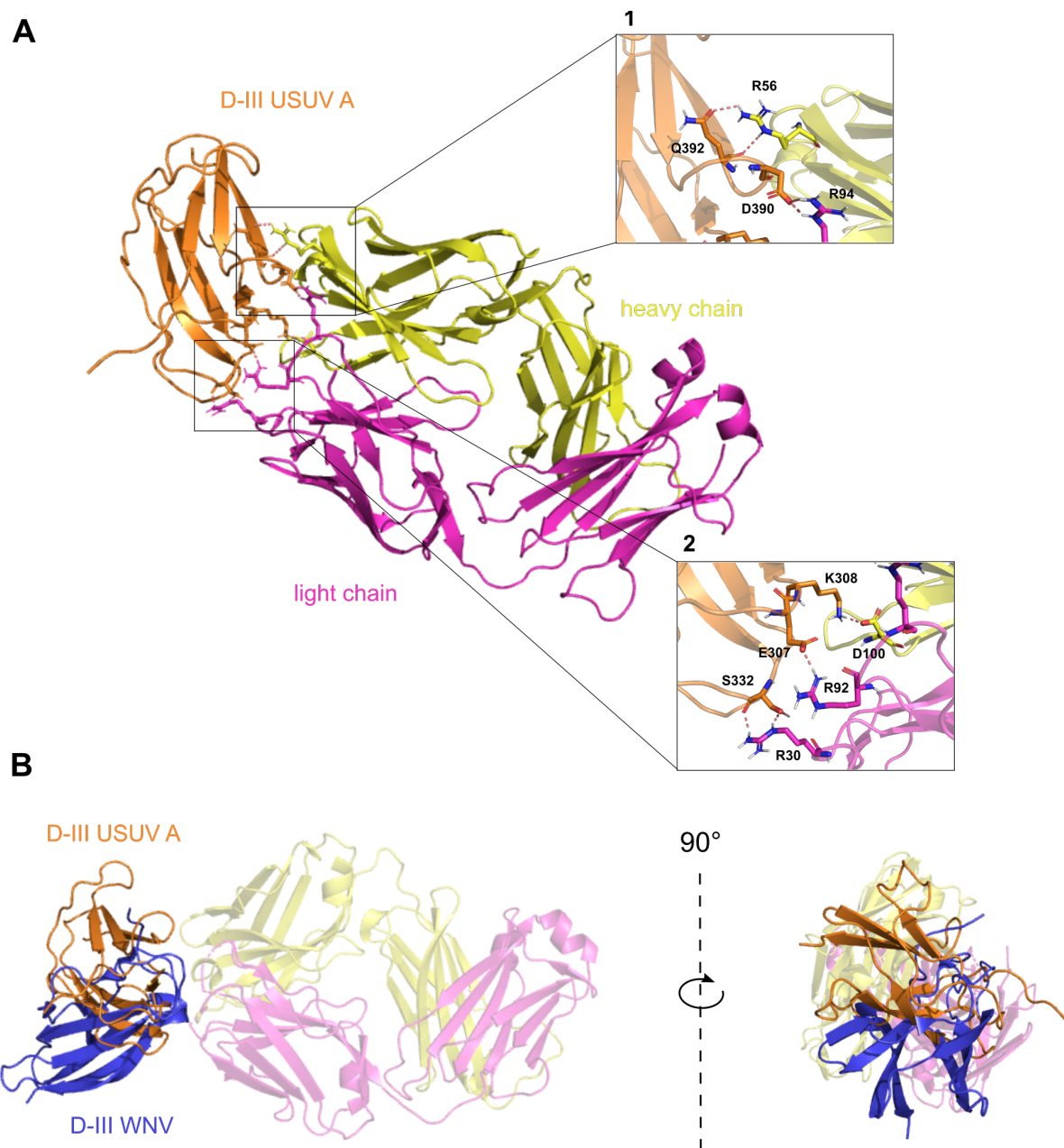


Fig. 10. Analysis of a potential E16 mutant in complex with the D-III of USUV A. (A) Model of USUV-A-D-III in complex with E16-mutA. Residues involved in the formation of potential polar contacts are shown in insets. (B) The structural alignment of the crystal structure of E16 in complex with WNV-D-III and the model complex of the E16-mutA and USUV-A-D-III shows a different positioning of the antigen.

Further, eight potential polar contacts are formed compared to two polar contacts between the USUV-A-D-III and WT E16. When looking at the binding interface in more detail,

three out of four of the positions mutated are directly involved in the formation of polar contacts. This is the case for R30_L, R92_L and R94_L, all situated in the light chain of the antibody. R92_L is involved in the formation of a polar contact to the D-III-E307 (Fig. 10.A, inset 2). This contact is of special interest when considering the specificity of a potential antibody, since the glutamic acid at this position is characteristic for USUV-A-D-III. In USUV-H-D-III and WNV-D-III a lysine is found at the very same position. The remaining contacts are formed between D100_H and D-III-K308, R30_L and D-III-S332 (Fig. 10. A, inset 2), R94_L and D-III-D390, and R56_H and D-III-Q392 (Fig. 10.A, inset 1), whereby S332, K308 and D390 are specific for USUV A in respect to WNV. Therefore, not only was the number of polar contacts increased, but also it was possible to generate contacts, which are specific for D-III of USUV A. Additionally, most contacts formed are of electrostatically nature.

When comparing the orientation of USUV-A-D-III towards the E16-mutA with the orientation of WNV-D-III towards the E16, we observed a clear change (Fig. 10.B). Compared to WNV-D-III, USUV-A-D-III is rotated by approximately 90° and tilted towards the heavy chain.

5.3.2 *D-III of USUV H in complex with E16-W33_HD-L52_HF-S30_LR-W50_LD (E16-mutH)*

We were also able to obtain a potential E16 mutant binding USUV-H-D-III by combining several single mutations from the computational calculations. Consequently, we generated an E16 mutant (E16-mutH) by introducing mutations at the following positions: W33_HE, L52_HF, S30_LR and W50_LD. The mutations were chosen under the same criteria as for the design of a USUV A specific antibody. In the case of USUV H, the $\Delta\Delta G_{\text{bind}}$ values are generally higher compared to USUV A, in the range of -2.8 and -0.6 kcal/mol (Tab. 21) and, consequently, the binding affinity lower. Additionally, the values for ΔE_{E16} are higher as well,

especially for the mutation at position 33 (heavy chain) from tryptophan to glutamic acid with 3.717 kcal/mol. However, since the removal provoked the loss of polar contacts within the binding interface, we kept this mutation.

Tab. 21. Energy values for a potential E16 mutant binding D-III of USUV H.

	$\Delta\Delta G_{\text{bind}}$	$\Delta E_{\text{complex}}$	ΔE_{E16}
W33_HE	-2.764	-0,198	3.717
L52_HF	-0.610	-1.775	-0.845
S30_LR	-2.792	-2.368	0.491
W50_LD	-2.527	-1.655	1.134

For the analysis of the complex formation between USUV-H-D-III and E16-mutH, we applied the same procedure as for USUV-A-D-III and E16-mutA and uploaded the complex to the RosettaDock server after introducing the mutations. Repeatedly, the model shows an interaction of USUV-H-D-III with the E16-mutH via the four binding loops. The complex exhibits nine potential polar contacts within the binding interface compared to five polar contacts of the initial complex including the WT E16 (Fig. 11.A). The mutations at position 30 from serine to arginine (S30_LR) and at position 50 from tryptophan to aspartic acid (W50_LD) are directly involved in the formation of polar contacts. R30_L contacts D-III-H395 and D50_L forms a contact with D-III-K308 (Fig. 11.A, inset 1). The Y92_L forms a contact with the backbone of D-III-F311 (Fig. 11.A, inset 1). On the heavy chain, the following contacts are formed: T54_H and R56_H contact D-III-E366, S97_H and Y98_H contact D-III-K307, and D100_H contacts D-III-N368 (Fig. 11.A, inset 2). The interaction with E366 and K308 are specific for USUV-H-D-III when compared to WNV and the interaction with K307 is specific when compared to USUV A.

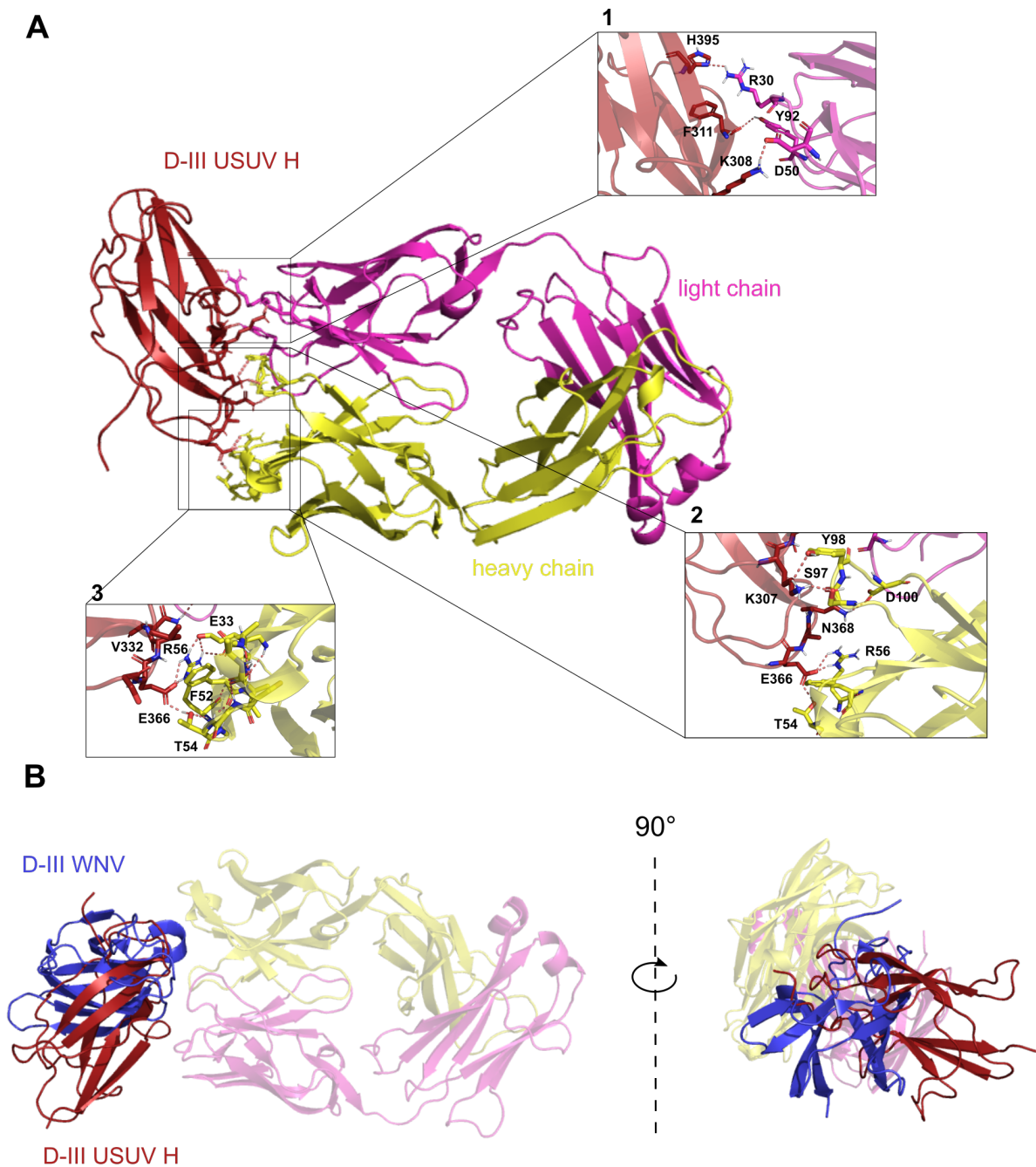


Fig. 11. Analysis of a potential E16 mutant in complex with the D-III of USUV H. (A) Cartoon representation of the model of USUV-H-D-III in complex with E16-mutH after RosettaDock. Insets show the binding interface in more detail. (B) The comparison of WNV-D-III in complex with E16 and USUV-H-D-III in complex with E16-mutH illustrates the different positioning of the antigen relative to the E16.

The mutation at position 52 from the aliphatic leucine to the aromatic phenylalanine (F52_H) appears to make hydrophobic contributions to the stability of the complex (Fig. 11. A,

inset 3). The introduction of phenylalanine leads to the gain of three additional polar contacts compared to leucine at the same position and appears to stabilize the surrounding residues in a certain orientation so that contacts can form. The glutamic acid at position 33 (E33_H) does not contribute to the complex stability in respect of the formation of polar contacts directly with the D-III (Fig. 11.A, inset 3). However, its removal correlates with the destabilization of the complex and the loss of polar contacts elsewhere in the interface. We assumed that the tryptophan present in the WT might lead to steric clashes due to its size and, consequently, to a different orientation inhibiting the formation of several polar contacts. Additionally, E33_H belongs to the residues surrounding F52_H and is involved in the formation of a tight network of polar contacts with the residual surrounding residues. The resulting orientation might be responsible for the contact formation and its stabilization between D-III-E366 and T54_H and R56_H.

As for USUV-A-D-III, we also observed a different orientation of USUV-H-D-III towards the E16 mutant compared to WNV-D-III (Fig. 11.B.). The D-III is rotated by 90°C, but in the opposite direction compared to USUV-A-D-III and tilted towards the light chain of the antibody.

5.4 Expression and purification of scFvE16 as insoluble inclusion bodies

Similar to the expression of D-III, scFvE16 protein is present as insoluble inclusion bodies when expressed in *E. coli*. The expression in various *E. coli* strains and under different expression conditions failed to improve the solubility of the protein. Therefore, we decided to establish a purification procedure based on the solubilization of the scFvE16 with high concentrations of urea followed by refolding.

Following cell lysis, several washing steps were applied to remove soluble impurities, lipids and nucleic acids. During the purification procedure, aliquots of the insoluble and soluble fraction were taken and analyzed on SDS PAGE (Fig. 12.A).

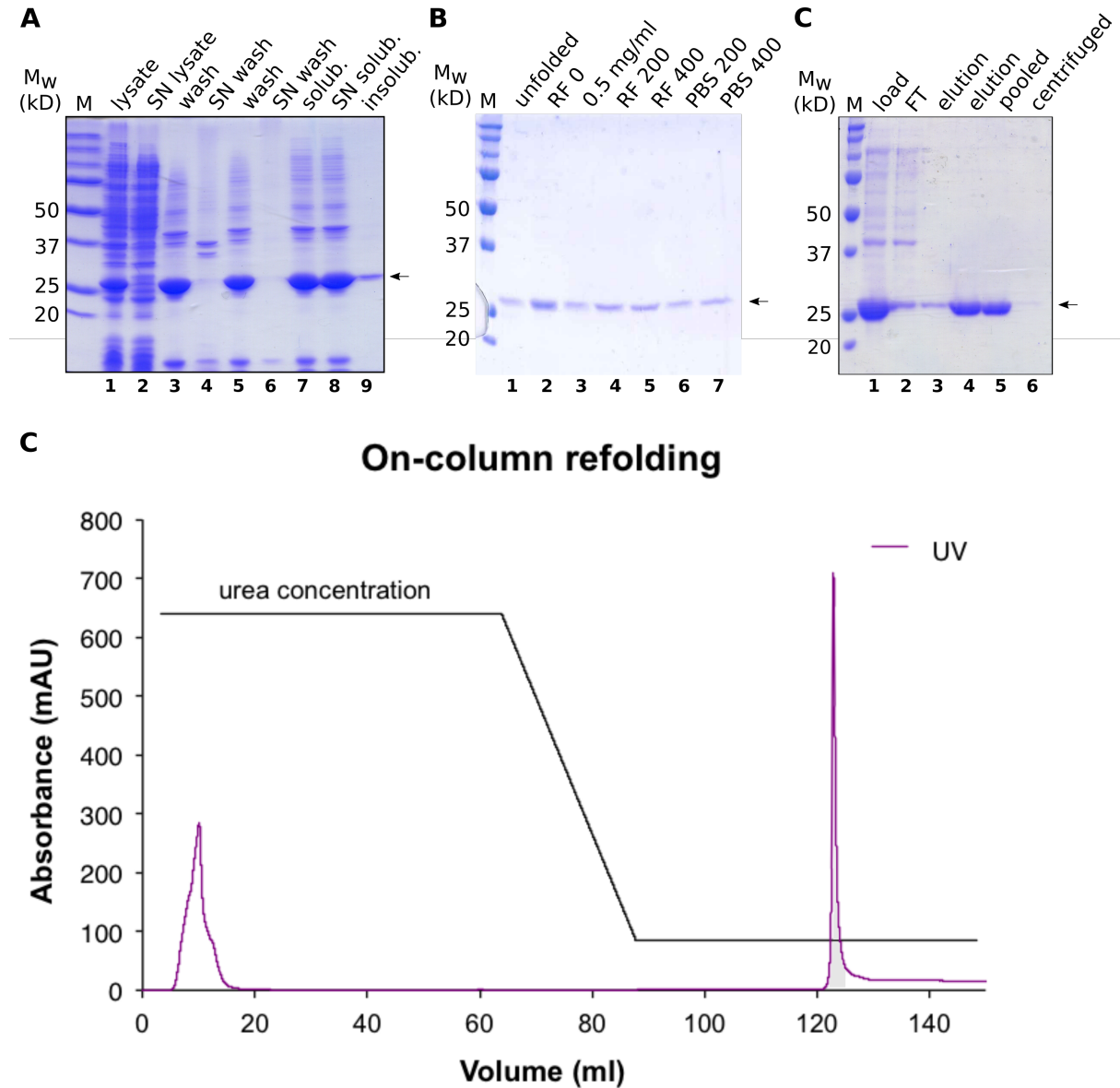


Fig. 12. Purification of scFvE16 from inclusion bodies and refolding. (A) SDS PAGE showing the washing steps during the purification of the scFvE16. In the last step, the scFvE16 is solubilized in 7 M urea. (B) During the refolding, the scFvE16 is analyzed on SDS-PAGE after each dialysis step. (C) SDS PAGE showing the fractions from scFvE16 on-column refolding. (D) Chromatogram of scFvE16 on-column refolding. After gradually reducing the urea concentration from 7 to 0 M, the scFvE16 is eluted with 500 mM imidazole from the HisTrap column.

The scFvE16 has an expected molecular weight of 27 kD and was observed only in the insoluble fraction (Fig. 12.A, lanes 1, 3, 5 and 7). After the solubilization procedure, the scFvE16 was detected in the supernatant (Fig. 12.A, lane 8).

For the refolding of the scFvE16, the urea concentration needs to be gently reduced and protein precipitation needs to be avoided. Several refolding methods exist and are well established. However, the success and refolding efficiency differ between individual proteins. We applied two different strategies to reduce the urea and obtain correctly folded protein: (1) dialysis and (2) on-column refolding, in which the protein is bound to a solid phase to avoid aggregation while reducing the urea concentration.

During dialysis, the urea concentration was reduced in a two-step procedure from 7 M to 3.5 M and from 3.5 M to 0 M. To increase the chances of soluble protein, the dialysis buffer contained a high salt concentration (0.5 M Tris-HCl and 0.15 M NaCl) to compensate for the reduction of the urea. However, during the second dialysis step when removing the urea completely, the protein strongly started to precipitate. The aggregates were removed by centrifugation, but SDS PAGE analyses showed that it was not possible to recover the scFvE16.

We applied the same strategy on a version of the E16-mutA, where we mutated residues at position 92 and 94 of the light chain to an arginine using site directed mutagenesis (scFvE16-mut_{Arg}). The idea was to investigate whether scFvE16 will be more soluble due to the introduction of charged amino acids. Analyses of the purification procedure on SDS PAGE demonstrate solid expression levels of the construct and the presence of the scFvE16-mut_{Arg} in the supernatant after the solubilization with 7 M urea (Fig. 13.A, lane 8). After the solubilization, we applied the same refolding procedure as

mentioned previously and analyzed the sample on SDS PAGE and western blot after removing the precipitates (Fig. 13.B). On SDS PAGE a light band was detectable around 27 kD (Fig. 13.B, lane 2). Nevertheless, western blot analysis could not confirm the presence of refolded scFvE16-mut_{Arg}. However, we detected a signal above 50 kD, which might be a potential dimer. Since the detection of a dimer is unlikely on a denaturing gel, we sent the sample for mass spectrometry analysis to validate this theory. However, due to problems during the measurement we were not able to obtain any results.

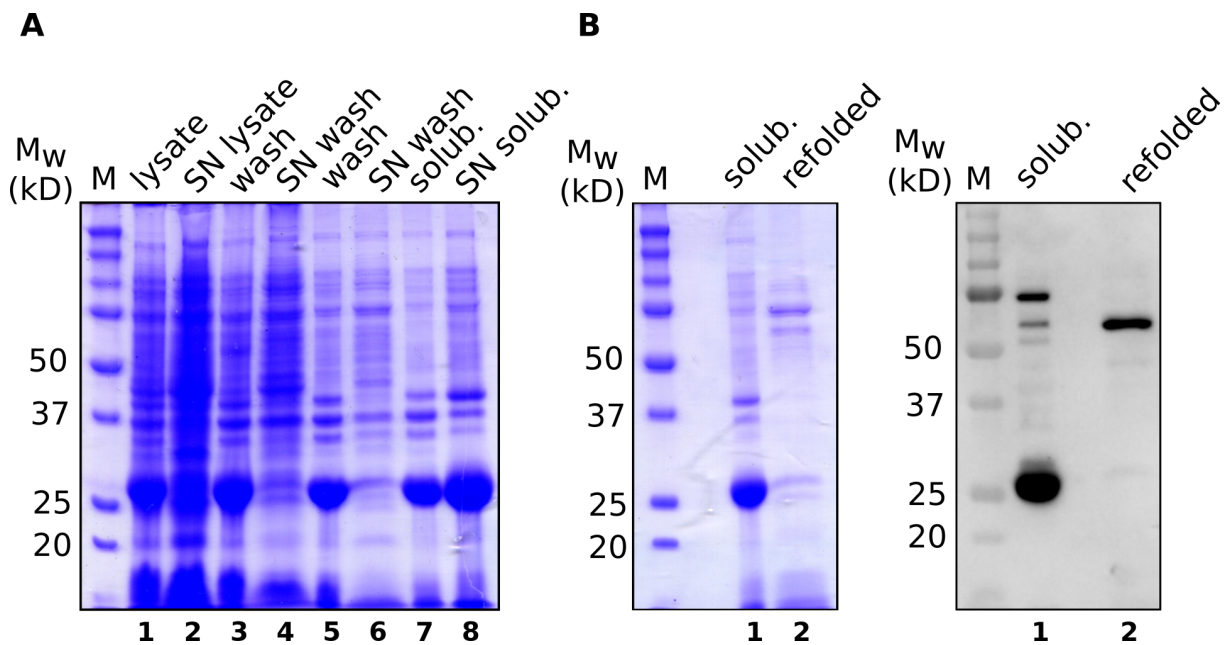


Fig. 13. Purification of scFvE16-mut_{Arg}. (A) Purification procedure analyzed on SDS PAGE. The scFvE16-mut_{Arg} is only detected in the insoluble fraction until solubilized with 7 M urea. (B) SDS PAGE and corresponding western blot of the scFvE16-mut_{Arg} before and after refolding.

Consequently, we supplemented the dialysis buffer with arginine, which is reported to increase the solubility of certain proteins⁶³. The first dialysis step was repeated as previously, since the scFvE16 was stable with only minor precipitation up to 3.5 M urea. The sample was split and refolding was continued using two different dialysis buffers: (1) containing 400 mM arginine and (2) containing 200 mM arginine, 2 mM reduced glutathione (GSH) and 1 mM

oxidized glutathione (GSSG). Since two disulfide bonds are present in the scFvE16, the redox system could improve the chances of correctly folded protein and reduce aggregation. In a final step, the protein was dialyzed into PBS. After each dialysis step, aggregates were removed and the sample was analyzed on SDS PAGE (Fig. 12.B). We were able to detect the scFvE16 in soluble form after each dialysis step. Therefore, we continued with size-exclusion chromatography to remove potential impurities not detectable on the gel and potential aggregates. However, when applying the sample on size-exclusion chromatography, we were not able to recover the scFvE16 from the column. Supposedly, the refolded scFvE16 was not stable enough under the conditions used and eventually precipitated on the column.

Our observations on the behavior of the scFvE16 during dialysis indicate that the protein inclines to precipitate. Therefore, we applied on-column refolding. The scFvE16 construct contains a his-tag on the C-terminus, which enables us to bind it to a Ni^{2+} affinity column. Bound to a solid phase, the individual proteins are spatially separated from each other, decreasing the chances of aggregation during the refolding procedure. Additionally, by controlling the flow rate, the urea concentration can be reduced gradually and slowly, while the concentration of solubility increasing components, such as arginine, are simultaneously increased. After the removal of the denaturant, we eluted the scFvE16 from the column with 0.5 M imidazole. We were able to detect a narrow peak and SDS PAGE confirmed that the peak contained the scFvE16 (Fig. 12.C and D). However, after the removal of the imidazole, the protein started precipitating and was, consequently, not stable in solution (Fig. 12.C, lane 6). Further, the variation of different arginine concentrations and

different additives, such as Triton X-100 or glycerol, did not improve the stability of the protein.

5.5 Expression of scFvE16 into the bacterial periplasm

Besides the expression as insoluble inclusion bodies followed by refolding, periplasmic expression represents the most widely used approach to produce scFvs in *E.coli*. To this end, a signal peptide is fused to the N-terminus of the protein, which directs the target to the bacterial secretion system. The oxidizing conditions present in the bacterial periplasm facilitate the formation of disulfide bonds and hence, increase the chances to obtain correctly folded protein compared to the cytoplasm. We made use of this approach by introducing the signal peptides PelB, OmpA, OmpT and PhoA on the N-terminus of the scFvE16. Further, we decided to test four different signal peptides, since their efficiency varies between individual target proteins. Subsequently, we performed small-scale test expressions under different conditions with varying temperature, expression time and IPTG concentration (see material and methods).

In the case of PelB-scFvE16, we were able to detect a band around 27 kD, which corresponds to the expected molecular weight (Fig. 14.A). The cells were lysed using cold osmotic shock and the lysate, periplasmic fraction (PPF), spheroblasts, cytoplasmic fractions (CF) and insoluble fraction were analyzed on SDS PAGE. However, the PelB-scFvE16 was detected only in the insoluble fractions (Fig. 14.A, lane 1,3 and 5), which lead to the conclusion that it was not successfully secreted into the periplasm. To gain a more detailed insight, we sent the sample for mass spectrometry. Analyses revealed that the scFvE16 was the most abundant protein in the sample; however, only minor amounts were expressed containing the PelB signal peptide. Hence, the scFvE16 was not secreted into the periplasm

and was mainly present as insoluble inclusion bodies. Since our construct contained two start codons, we assumed that the start codon of the PelB signal peptide was recognized with lower efficiency and translation was mainly initiated at the second start of the coding sequence. Thus, we increased the distance between the ribosomal binding site (RBS) and the first start codon by 35 nucleotides to increase the probability of the ribosome recognizing the PelB start codon. Small-scale test expressions were repeated using the construct with increased distance between the RBS and the PelB start codon (Fig. 14.B). We were able to detect a band at approximately the correct molecular weight in the PPF and CF (Fig. 14.B, lane 4 and 5). However, these results could not be confirmed by western blot analysis. Mass spectrometry analyses revealed that the protein detected on SDS PAGE was β -lactamase, a protein found in the bacterial periplasm providing antibiotics resistance⁶⁴.

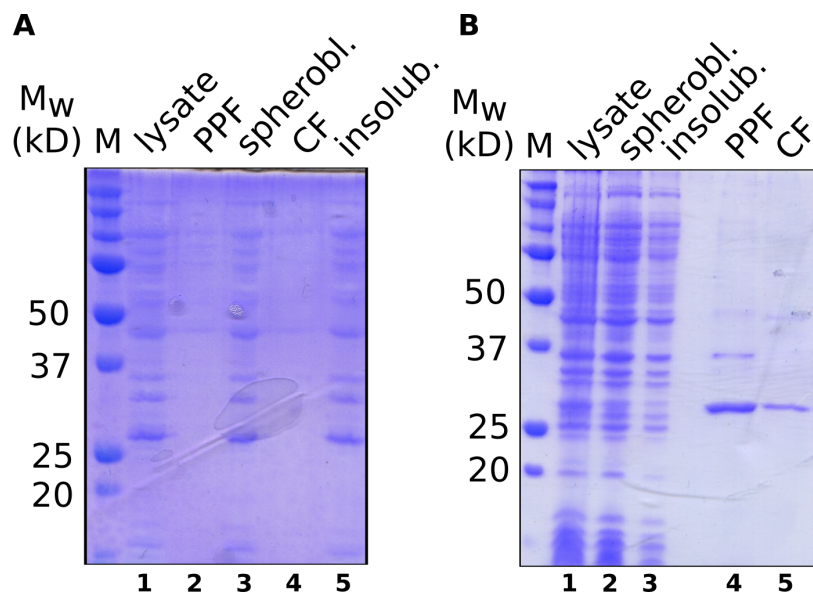


Fig. 14. PelB-scFvE16 expression. (A) SDS PAGE analyses of PelB-scFvE16 after cold osmotic shock of the cells. The expression was induced with 0.5 mM IPTG and the culture was incubated for 4 h at 37°C. The gel is representative for all other expression conditions tested (see material and methods). (B) Analyses of the osmotic shock fractions of PelB-scFvE16 after increasing the distance between the RBS and the PelB start codon. Expression was induced with 0.2 mM IPTG and the culture was incubated for 18 h at 16°C. The gel is representative for all other expression conditions tested (see material and methods).

To avoid the presence of two start codons within the construct, we removed the scFvE16 start codon during the cloning procedure of the remaining signal peptides, OmpA, OmpT and PhoA. Subsequently, we repeated small-scale test expressions with OmpA-scFvE16, OmpT-scFvE16 and PhoA-scFvE16 under varying conditions. Repeatedly, we analyzed the whole lysate and the supernatant on SDS PAGE and via western blot in comparison to cultures containing the empty plasmid (Fig. 15). However, the comparison with the empty vector did not show a difference in expression and we were not able to obtain a signal from Western Blot analysis. As a result, the constructs were not expressed or their expression level is below the detection limit.

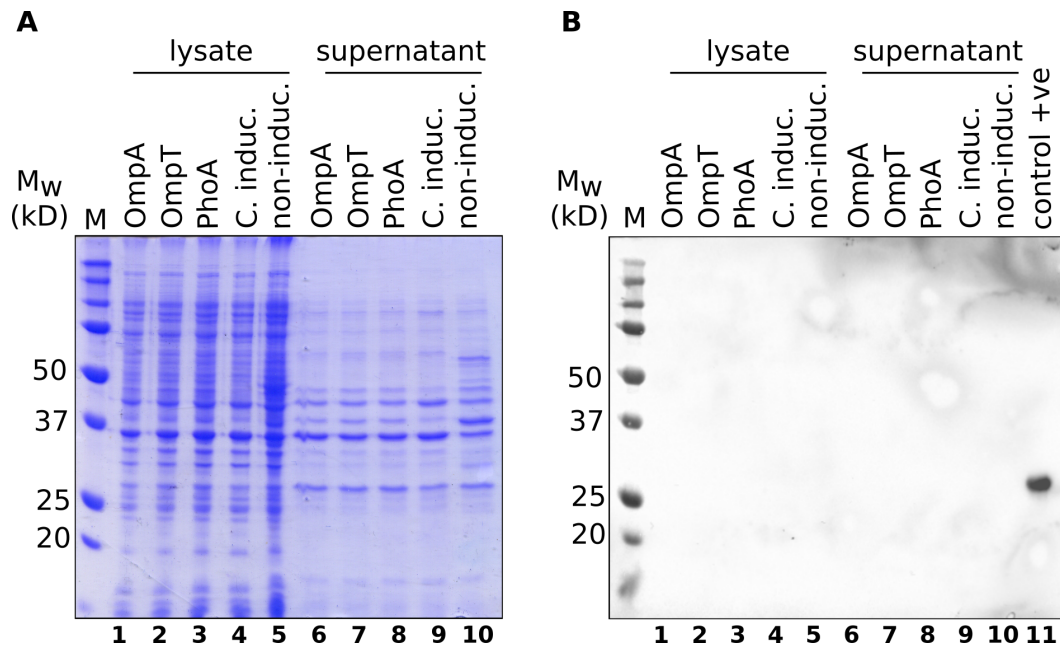


Fig. 15. Expression of OmpA-scFvE16, OmpT-scFvE16 and PhoA-scFvE16. (A) SDS PAGE analyses of the lysate and supernatant of scFvE16 fused to different signal peptides after expression for 5 h at 30°C. The gel is representative for all expression conditions tested (see material and methods). (B) Corresponding western blot. C. induc. = control with the empty vector induced with IPTG; non-induc. = control with the empty vector not induced.

6 Discussion

6.1 Structure of D-III of USUV A and B

The solution of the structure of USUV-A-D-III and USUV-B-D-III was essential for proceeding with our computational saturated mutagenesis approach and was necessary as a basis for the calculations, although we were not able to apply the approach on USUV-B-D-III.

Generally, the D-III of the E protein of flaviviruses presents the ideal target for antibody binding, as it is the outermost protein domain on the surface of the virus particle. Hence, several antibodies for other flaviviruses, such as WNV or ZIKV, exist targeting the D-III^{22,26}. Due to the sequence similarity and structural similarity between WNV-D-III and USUV-D-III, we could make use of the E16 antibody targeting WNV as a basis. The four binding loops determined to be involved in the binding between WNV-D-III and the E16 exhibit close structural similarity when compared to USUV-A-D-III (Fig. 16.A). Therefore, the binding of the antibody mainly depends on the identity of the side chains situated within these loops. All together, 8 residues out of the 16 residues involved in the complex formation differentiate between WNV-D-III and USUV-A-D-III. Residues differentiating not only in their amino acid identity, but also in their amino acid property are of special interest for the design of an Usutu virus specific antibody. The glutamic acid at position 307 of USUV-A-D-III, which corresponds to a lysine in WNV-D-III, emphasizes the importance of targeting such residues for antibody specificity. By mutating the tyrosine at position 92 of the light chain of the E16 to an arginine the formation of an electrostatic interaction between the glutamic acid and the arginine becomes feasible (Fig. 16.B). On the contrary, the lysine of

WNV-D-III would not be able to form this electrostatic interaction due to the opposite charge and thus, a contact is generated specifically for USUV-A-D-III.

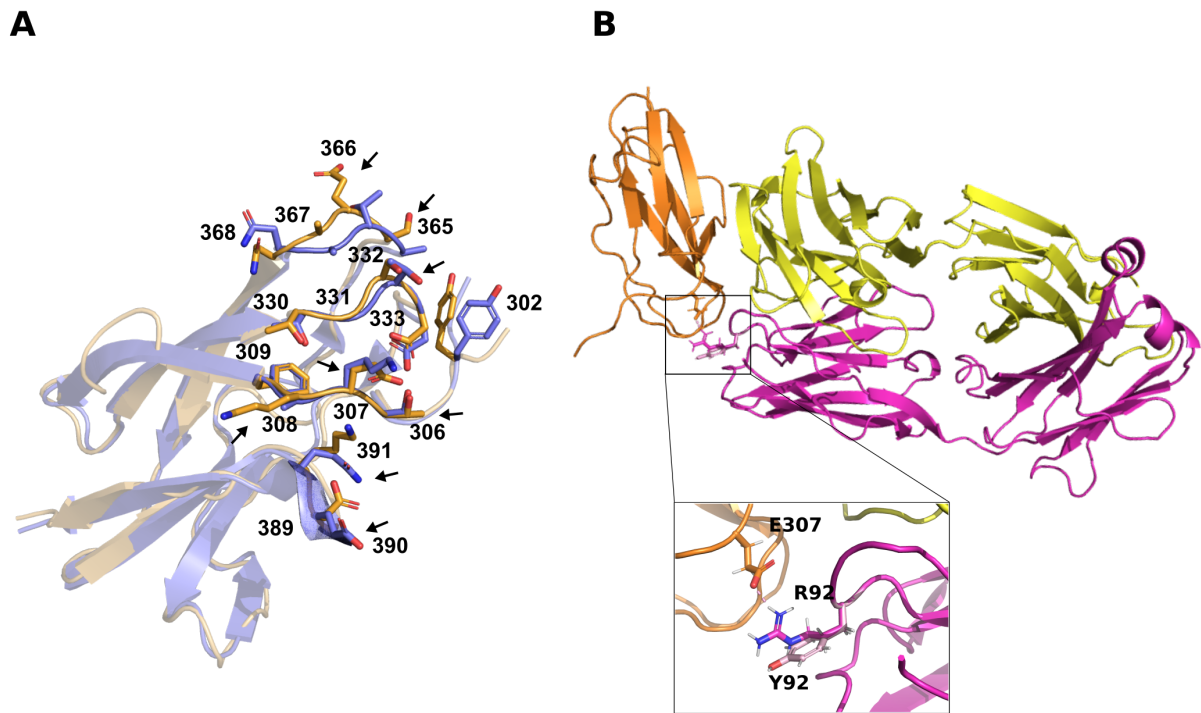


Fig. 16. Structural alignment of D-III of USUV A and WNV. (A) The residues involved in the binding between WNV-D-III (blue) and E16 are shown as sticks, as well as the analogous residues of USUV-A-D-III (orange). Residues differentiating in their amino acid identity are marked with an arrow. (B) USUV-A-D-III in complex with the E16-mutA. The inset shows the difference between the E16-mutA and WT E16 at position 92 of the light chain and the electrostatic interaction with E307 generated.

6.2 Computational saturated mutagenesis

Generally, computational saturated mutagenesis was implemented as a method to optimize existing protein-protein interactions by introducing single mutations^{47,49}. Thus, it depicts a computational approach to investigate the evolution of protein binding interfaces. The method allows the determination of evolutionary optimized positions within a binding interface, as well as positions having the potential to be further optimized by introducing single mutations and the subsequent calculation of the difference in energy. However, in our case, we attempted to computationally generate binding between two proteins.

Therefore, we based our calculations on the structure of the complex of WNV-D-III and E16 and replaced the initial binding partner, WNV-D-III, with the crystal structure of USUV-A-D-III and USUV-H-D-III, respectively. Due to the sequence similarity, as well as the structural similarity between WNV-D-III and USUV-A-D-III/USUV-H-D-III, computational saturated mutagenesis represented a reasonable approach to engineer an antibody.

Based on our calculations, we accomplished to generate E16 mutants and models of these mutants in complex with USUV-A-D-III and USUV-H-D-III. Essentially, our models showed an increased number of polar contacts between the D-IIIs and the E16 mutants compared to models including the D-IIIs and WT E16. However, we were not able to generate mutants with a similar tight network of polar contacts, as it is present between WNV-D-III and E16. Nevertheless, we could demonstrate that positions calculated to improve the binding affinity between the D-III of USUV A/H and E16 do actually lead to changes within the binding interface.

Further, we generated models of WNV-D-III in complex with E16-mutA and E16-mutH and applied RosettaDock. The mutations we introduced to generate binding towards the D-IIIs of USUV reduced the interaction network dramatically. In case of WNV-D-III and E16-mutA, eight polar contacts are identified within the binding interface, thus comparable to the interaction between USUV-A-D-III and E16-mutA. However, the number of contacts was reduced by almost half compared to 15 between WNV-D-III and the WT E16. Consequently, WNV-D-III eventually still binds E16-mutA, but probably with reduced binding affinity. The model of WNV-D-III and E16-mutH showed a reduction of polar contacts to five. Hence, it is questionable whether the E16-mutH binds WNV-D-III after all. Thus, our analyses indicate

that our design is not only enhancing the binding affinity towards our target, but also simultaneously reducing the affinity to the initial binding partner.

Nevertheless, our models indicate a change of the relative orientation of the D-III and the E16 mutants in the case of USUV A and H compared to WNV-D-III binding E16. The estimation of the influence of the altered orientation on the binding affinity remains challenging, mainly depending on the accuracy of the docking algorithm. In case the docking is accurate, the altered orientation would contribute to the formation of new polar contacts after introducing the mutations. On the contrary, in case the altered orientation does not represent reality, the polar contacts generated would probably not form and the E16 mutants would not bind USUV-D-III. Therefore, the experimental validation of the E16 mutants is crucial not only to test the binding itself, but also to examine how the mutations influence the relative orientation in reality and to review the accuracy of our predictions using the RosettaDock server.

To optimize the calculation procedure, an interesting approach would depict to perform an initial run introducing single mutations, as we did. However, after analyzing the results and the combination of several mutations, the implementation of a second run of computational saturated mutagenesis with these E16 mutants as initial binding partners, seems reasonable. Consequently, such a procedure would facilitate the estimation of the effect of the combination of mutations regarding the binding affinity, as well as, the stability of the E16 mutants. Besides, additional positions can be determined to increase the binding affinity towards the target further, while simultaneously decreasing the affinity towards the initial binding partner and therefore, enhancing the specificity. This stepwise optimization

could improve the chances to obtain feasible mutants capable of binding the target experimentally.

Finally, of course, an essential step in this approach is missing and that is to test the computationally achieved models experimentally. Nevertheless, we were able to generate models of USUV-A-D-III and USUV-H-D-III in complex with computationally generated E16 mutants showing enhanced binding affinity and specificity towards the target. Hence, the optimization and verification of this approach could present a possible alternative to conventional antibody design and engineering. The interface between the antibody and the target can be specifically manipulated reducing the number of possible antibodies to be tested experimentally. The only requirement is an already existing antibody against a similar target as a basis for computational saturated mutagenesis.

6.3 Purification of E16 as an scFv antibody

In general, the refolding of a protein is a challenging procedure and the exact process needs to be established for each protein individually by trial and error. Refolding was applied successfully on several proteins so far, such as in our case, the refolding of the D-IIIs, as we were able to purify them from inclusion bodies to sufficient amounts. Likewise, we tried to establish a refolding strategy for the scFvE16 using dialysis and on-column refolding. However, during the dialysis, the scFvE16 precipitates heavily after the reduction of the urea to a critical threshold and, consequently, was not recoverable in its native form. The addition of arginine, a solubility enhancing component, lead to the detection of scFvE16 on SDS PAGE after the refolding procedure. Nevertheless, the protein was still not stable enough in solution to proceed with the purification process. The screening of the chemical space of the scFvE16 using dynamic light scattering (DLS) or differential scanning fluorimetry (DSF) would

perhaps enable to identify buffer conditions with stability enhancing effects. However, due to the small amounts of recoverable scFvE16, we were limited in performing buffer screens. Therefore, we proceeded with the application of on-column refolding, having the advantage that the target protein is bound to a solid phase in its denatured state and thus, reduces the probability of aggregation during the refolding. Similar to the refolding by dialysis, the scFvE16 was not stable in solution after the elution from the column and precipitated after a few hours at 4°C and upon the removal of imidazole from the elution buffer. Consequently, we were not able to establish a refolding approach leading to stable and soluble scFvE16.

A second approach to express and purify scFvs from *E. coli* depicts the introduction of a N-terminal signal peptide directing the target protein to the bacterial secretion system followed by the purification from the bacterial periplasm. We made use of this approach by fusing the scFvE16 to four different signal peptides: PelB, OmpA, OmpT and PhoA. All these signal peptides represent well established systems for periplasmic expression and have been successfully used for the purification of proteins before⁶⁵⁻⁶⁷. However, the secretion efficiency and the expression level vary between the signal peptides and the individual target protein. Therefore, it is difficult to estimate which signal peptide is the most suitable for a specific target. In our case, none of the tested signal peptides led to satisfactory expression levels of the scFvE16. The PelB-scFvE16 was expressed without the signal peptide, which led to the formation of insoluble inclusion bodies in the cytoplasm. By increasing the number of nucleotides between the RBS and the PelB start codon, we wanted to ensure the initiation of the translation at the first start codon. However, expression levels of this modified construct were too low to be detectable. A similar situation was observed

for the scFvE16 fused to OmpA, OmpT and PhoA and we were not able to detect any expression.

In general, recovery from inclusion bodies and periplasmic expression are two well established and widely successfully applied methods for the expression and purification of scFvs in *E. coli*. As scFvs usually tend to the formation of insoluble inclusion bodies when expressed in bacteria, both approaches present a technique for the production of antibodies in a prokaryotic system to large amounts. However, as our case shows, the success and applicability strongly depends on the individual scFv. Thus, for the production of soluble and correctly folded scFvE16, it will be necessary to apply a eukaryotic expression system, such as insect cells or mammalian cells. Although it is difficult to obtain yields with eukaryotic expression systems comparable to bacterial expression systems, the chances are higher to obtain soluble and correctly folded protein. In the case of scFvE16, the solubility and folding of the protein represent the main limitations in the production. Thus a switch to a eukaryotic expression system seems reasonable.

7 References

1. Woodall, J. P. The viruses isolated from arthropods at the East African Virus Research Institute in the 26 years ending December 1963. *Proc E Afr Acad* **2**, 141–146 (1964).
2. Weissenböck, H. *et al.* Emergence of Usutu virus, an African mosquito-borne flavivirus of the Japanese encephalitis virus group, Central Europe. *Emerg Infect Dis* **8**, 652–656 (2002).
3. Weissenböck, H. *et al.* Usutu virus activity in Austria, 2001 – 2002. *Microbes Infect.* **5**, 1132–1136 (2003).
4. Gaibani, P. & Rossini, G. An overview of Usutu virus. *Microbes Infect.* **19**, 382–387 (2017).
5. Ursu, K. *et al.* Emergence of Usutu virus in Hungary. *J. Clin. Microbiol.* **45**, 3870–3874 (2007).
6. Steinmetz, H. W. *et al.* Emergence and establishment of Usutu virus infection in wild and captive avian species in and around Zurich, Switzerland — Genomic and pathologic comparison to other central European outbreaks. *Vet. Microbiol.* **148**, 207–212 (2011).
7. Pinna, E. De, Day, M., Lawrence, J. & Jones, J. Usutu virus sequences in *Culex pipiens* (Diptera: Culicidae), Spain. *Emerg Infect Dis* **14**, 861–863 (2008).
8. Garigliany, M. *et al.* Detection of Usutu virus in a bullfinch (*Pyrrhula pyrrhula*) and a great spotted woodpecker (*Dendrocopos major*) in north-west Europe. *Vet. J.* **199**, 191–193 (2014).
9. Lecollinet, S. *et al.* Dual emergence of Usutu virus in common blackbirds, eastern France, 2015. *Emerg Infect Dis* **22**, 2225–2227 (2016).
10. Cadar, D. *et al.* Usutu virus in bats, Germany, 2013. *Emerg Infect Dis* **20**, 2013–2015 (2014).
11. Lupulovic, D. *et al.* First serological evidence of West Nile virus activity in horses in Serbia. *Vector-Borne Zoonotic Dis* **11**, 1303–1305 (2011).
12. Weissenböck, H., Bakonyi, T., Rossi, G., Mani, P. & Nowotny, N. Usutu virus, Italy,

1996. *Emerg Infect Dis* **19**, 1996–1999 (2013).
13. Poidinger, M., Hall, R. A. & Mackenzie, J. S. Molecular characterization of the Japanese encephalitis serocomplex of the flavivirus genus. *Virology* **421**, 417–421 (1996).
 14. Nikolay, B., Diallo, M., Boye, C. S. & Sall, A. A. Usutu virus in Africa. *Vector-Borne Zoonotic Dis* **11**, 1417–1423 (2011).
 15. Rice, C. M. *et al.* Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science*. **229**, 726–733 (1985).
 16. Gould, E. A., Kolodziejek, J. & Nowotny, N. Complete genome analysis and molecular characterization of Usutu virus that emerged in Austria in 2001 Comparison with the South African Strain SAAR-1776 and other flaviviruses. **328**, 301–310 (2004).
 17. Smit, J. M., Moesker, B., Rodenhuis-Zybert, I. & Wilschut, J. Flavivirus cell entry and membrane fusion. *Viruses* **3**, 160–171 (2011).
 18. Li, L. *et al.* The flavivirus precursor membrane-envelope protein complex: structure and maturation. *Science*. **319**, 1830–1835 (2008).
 19. Murray, C. L., Jones, C. T. & Rice, C. M. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nat. Rev. Microbiology* **6**, 699–708 (2008).
 20. Watterson, D., Modhiran, N. & Young, P. R. The many faces of the flavivirus NS1 protein offer a multitude of options for inhibitor design. *Antiviral Res.* **130**, 7–18 (2016).
 21. Nybakken, G. E., Nelson, C. A., Chen, B. R., Diamond, M. S. & Fremont, D. H. Crystal structure of the West Nile virus envelope glycoprotein. *J. Virol.* **80**, 11467–11474 (2006).
 22. Qin, C., Qi, J. & Gao, G. F. Structures of the Zika virus envelope protein and its complex with a flavivirus broadly protective antibody. *Cell Host Microbe* **19**, 696–704 (2016).
 23. Modis, Y., Ogata, S., Clements, D. & Harrison, S. C. A ligand-binding pocket in the dengue virus envelope glycoprotein. *PNAS* **100**, 6986–6991 (2003).
 24. Zhang, X., Jia, R., Shen, H., Wang, M. & Yin, Z. Structures and functions of the envelope glycoprotein in flavivirus infections. *Viruses* **9**, 1–14 (2017).
 25. Chen, Z. *et al.* Crystal structure of Usutu virus envelope protein in the pre-fusion state. *Virol. J.* **15**, 1–8 (2018).

26. Oliphant, T. *et al.* Development of a humanized monoclonal antibody with therapeutic potential against West Nile virus. *Nat. Med.* **11**, 522–530 (2005).
27. Nybakken, G. E. *et al.* Structural basis of West Nile virus neutralization by a therapeutic antibody. *Nature* **437**, 764–768 (2005).
28. Nikolay, B. *et al.* Validation of a structural comparison of the antigenic characteristics of Usutu virus and West Nile virus envelope proteins. *Virus Res.* **189**, 87–91 (2014).
29. Ashraf, U. *et al.* Usutu virus: an emerging flavivirus in Europe. *Viruses* **7**, 219–238 (2015).
30. Institut Pasteur de Dakar. Rapport sur le fonctionnement technique de l'Institut Pasteur de Dakar. 1982–1984 (1984).
31. Vázquez, A. *et al.* Usutu virus – potential risk of human disease in Europe. *Euro Surveill* **16**, 1–5 (2011).
32. Cavrini, F. *et al.* Usutu virus infection in a patient who underwent orthotopic liver transplantation, Italy, August–September 2009. *Euro Surveill* **14**, (2009).
33. Pecorari, M. *et al.* First human case of Usutu virus neuroinvasive infection, Italy, August–September 2009. *Euro Surveill* **14**, 9–10 (2009).
34. Cavrini, F. *et al.* A rapid and specific real-time RT-PCR assay to identify Usutu virus in human plasma, serum, and cerebrospinal fluid. *J. Clin. Virol.* **50**, 221–223 (2011).
35. Vilibic-Cavlek, T. *et al.* First evidence of simultaneous occurrence of West Nile virus and Usutu virus neuroinvasive disease in humans in Croatia during the 2013 outbreak. *Infection* **42**, 689–695 (2014).
36. Cadar, D. *et al.* Blood donor screening for West Nile virus (WNV) revealed acute Usutu virus (USUV) infection, Germany, September 2016. *Euro Surveill* **22**, 1–5 (2017).
37. Bakonyi, T. *et al.* Usutu virus infections among blood donors, Austria, July and August 2017 – Raising awareness for diagnostic challenges. *Euro Surveill* **22**, 1–6 (2017).
38. Dick, G. W. A., Kitchen, S. F. & Haddow, A. J. Zika Virus I. isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* 509–520 (1952).
39. Faye, O. *et al.* Molecular evolution of Zika virus during its emergence in the 20th century. *PLoS Negl. Trop. Dis.* **8**, 1–10 (2014).
40. Hayes, E. B. Zika virus outside Africa. *Emerg Infect Dis* **15**, 1347–1350 (2009).
41. Hennessey, M., Fischer, M. & Staples, J. E. Zika virus spreads to new areas — region of

- the Americas, May 2015 – January 2016. *Morb. Mortal. Wkly. Rep.* **65**, 55–58 (2016).
42. Schuler-Faccini, L. *et al.* Possible Association Between Zika Virus Infection and Microcephaly — Brazil, 2015. *Morb. Mortal. Wkly. Rep.* **65**, 59–62 (2016).
 43. Fleishman, S. J. *et al.* Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science*. **332**, 816–821 (2012).
 44. Pal, G., Kouadio, J. K., Artis, D. R., Kossiakoff, A. A. & Sidhu, S. S. Comprehensive and quantitative mapping of energy landscapes for protein-protein interactions by rapid combinatorial scanning. *J. Biol. Chem.* **281**, 22378–22385 (2006).
 45. Fleishman, S. J. *et al.* Community-wide assessment of protein-interface modeling suggests improvements to design methodology. *J. Mol. Biol.* **414**, 289–302 (2011).
 46. Bahadur, R. P., Chakrabarti, P., Rodier, F. & Janin, J. A dissection of specific and non-specific protein – protein interfaces. *J. Mol. Biol.* **336**, 943–955 (2004).
 47. Aizner, Y. *et al.* Mapping of the binding landscape for a picomolar protein-protein complex through computation and experiment. *Structure* **22**, 636–645 (2014).
 48. Schreiber, G. & Fleishman, S. J. Computational design of protein – protein interactions. *Curr. Opin. Struct. Biol.* **23**, 903–910 (2013).
 49. Sharabi, O., Erijman, A. & Shifman, J. M. *Computational methods for controlling binding specificity. Methods in Enzymology* **523**, (Elsevier Inc., 2013).
 50. Bird, R. E. *et al.* Single-chain antigen-binding proteins. *Science*. **242**, 423–426 (1988).
 51. Ahmad, Z. A. *et al.* scFv antibody: principles and clinical application. *Clin. Dev. Immunol.* **2012**, (2012).
 52. Wang, R. *et al.* Engineering production of functional scFv antibody in E. coli by co-expressing the molecule chaperone Skp. *Cell. Infect. Microbiol.* **3**, 1–9 (2013).
 53. Hammers, C. M. & Stanley, J. R. Antibody phage display: technique and applications. *J Invest Dermatol* **134**, 1–13 (2014).
 54. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
 55. Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica* **64**, 61–69 (2008).
 56. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for

- macromolecular structure solution. *Acta Crystallographica* **66**, 213–221 (2010).
57. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica* **60**, 2126–2132 (2004).
 58. Chen, V. B. *et al.* MolProbity : all-atom structure validation for macromolecular crystallography. *Acta Crystallographica* **66**, 12–21 (2010).
 59. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
 60. Lyskov, S. & Gray, J. J. The RosettaDock server for local protein-protein docking. *Nucleic Acids Res.* **36**, 233–238 (2008).
 61. Desmet, J., De Maeyer, M., Hazes, B. & Lasters, I. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature* **356**, 539–542 (1992).
 62. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica* **60**, 2256–2268 (2004).
 63. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A. & Lian, L. A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.* **126**, 8933–8939 (2004).
 64. Wilke, M. S., Lovering, A. L. & Strynadka, N. C. J. b-lactam antibiotic resistance: a current structural perspective. *Curr. Opin. Microbiol.* **8**, 525–533 (2005).
 65. Chi, W., Kim, H., Yoo, H., Kim, Y. P. & Hong, S. Periplasmic expression, purification, and characterization of an anti-epidermal growth factor receptor antibody fragment in *Escherichia coli*. *Biotechnol. Bioprocess Eng.* **21**, 321–330 (2016).
 66. Kiran, S., Suresh, M. R. & Wuest, F. R. Improved soluble expression of a single-chain antibody fragment in *E. coli* for targeting CA125 in epithelial ovarian cancer. *Protein Expr. Purif.* **102**, 27–37 (2014).
 67. Zhang, H. *et al.* Cultivation of recombinant *Escherichia coli* for secretory production of human epidermal growth factor under control of PL promoter. *Enzyme Microb. Technol.* **40**, 708–715 (2007).

8 Appendix: Zusammenfassung

Usutu Virus (USUV) gehört zur Familie der *Flaviviridae* und wird von Mücken übertragen. USUV weist antigenetische Kreuzreaktivität mit West Nile Virus (WNV) auf, was eine ähnliche Struktur der Oberflächenproteine vermuten lässt. Basierend auf dem E16 Antikörper, welcher die domain III (D-III) des Envelope protein (E protein) von WNV bindet und diesen neutralisiert, wollen wir einen Antikörper in Form eines single-chain variable fragments (scFV), spezifisch gegen USUV entwickeln. Solch ein Antikörper hätte Potential um gegen mögliche USUV Infektionen eingesetzt werden zu können oder könnte in der Diagnostik Verwendung finden. Für das Design eines USUV-spezifischen Antikörpers verwenden wir einen computergestützten Ansatz, „computational saturated mutagenesis“, wobei die Kristallstruktur des USUV-D-III Antigens als Basis für die Computerberechnungen dient. Dabei werden rechnerisch einzelnen Mutationen im E16 Antikörper eingefügt und anschließend die Energiedifferenz berechnet um abzuschätzen ob die Mutation in einer erhöhten oder verminderten Bindungsaffinität zum USUV-D-III resultiert. Folglich haben wir das D-III von verschiedenen USUV Stämmen in *E. coli* aufgereinigt, kristallisiert und dessen Struktur gelöst. Zusätzlich haben wir versucht ein Aufreinigungsprotokoll zur Herstellung der scFv Version des E16 Antikörpers (scFvE16) in *E. coli* zu etablieren. Bei der Expression des scFvE16 in *E. coli* kommt es allerdings zur Bildung von unlöslichen Proteinaggregaten und auch anschließendes *in vitro* refolding führt zu keinem zufriedenstellenden Ergebnis. Ein solches Protokoll wäre dennoch notwendig für die Herstellung zukünftiger potentieller

USUV-spezifischer Antikörper resultierend aus den Computerberechnungen und deren experimenteller Evaluierung.