Project Code: BRG/05/2541

Project Title:

THE DENGUE VIRUS PROTEASE COMPLEX NS2B/NS3 – A NOVEL TARGET FOR THE STRUCTURE – BASED DESIGN OF IMPROVED ANTIVIRAL AGENTS

Principal Investigator: Gerd Katzenmeier, Ph.D.
Institute of Molecular Biology and Genetics
Mahidol University
email: frkgz@mahidol.ac.th

Co-investigator: Chanan Angsuthanasombat, Ph.D.
Institute of Molecular Biology and Genetics
Mahidol University

Period of Funding: October 01, 1997 - May 30, 2001

Abstract

The objective of this project is the detailed characterization of the dengue virus protease complex NS2B/NS3 by molecular biological and biochemical methods in order to analyse structure-activity relationships of the enzyme with a view to the rationale-based design of antiviral compounds.

During the course of this project, we have constructed recombinant clones of the NS2B/NS3 protease components in *Escherichia coli* which over-express the target proteins and we have developed biochemical methods for their purification and refolding. The dengue virus proteins NS2B, NS3(pro) and NS2B/NS3 from dengue virus type 2 were expressed as fusion proteins containing hexahistidine affinity tags and were purified to near-homogeneity by Ni - affinity metal chelate chromatography. Several methods for the refolding of the NS3 protease were explored and stepwise dialysis was used to renature the enzyme to its active

conformation. A major part of the project was the development of an assay system to monitor NS3 proteolytic activity. A number of methods were evaluated and currently a fluorometric assay with short peptides and a HPLC-based assay using dansylated 12mer peptide substrates which mimick authentic dengue virus polyprotein cleavage sites is used for the analysis of NS3 protease biochemical reaction parameters. We have shown that all peptides encompassing native polyprotein cleavage sites are specifically cleaved by the NS3 enzyme. The assay is suitable for the analysis of substrate requirements, structural determinants of cleavage activity and -specificity and for the evaluation of potential inhibitors.

An alternative expression system for the NS3 protein in the yeast *Pichia pastoris* was explored and we could show that NS3(pro) fused to the yeast alpha-factor was expressed at high levels albeit not secreted into the culture medium. A recombinant fusion protein containing the hydrophilic subdomain of NS2B fused to NS3(pro) was expressed in *E. coli* and we have demonstrated autoproteolytic processing at the NS2B-NS3 site after purification and refolding. This experimental system will be useful to further characterize the mechanism of cofactor activation of the NS3 protease. The project is now equipped with a number of useful tools and techniques to address the biochemical research problems associated with a structural and functional characterization of the DEN NS3 protease.

Keywords:

Dengue virus, polyprotein, NS2B, NS3, serine protease, cofactor, expression, purification, assay, substrate, HPLC, fluorescence

Introduction

Dengue viruses, members of the family *Flaviviridae*, are transmitted by mosquitoes of the species *Aedes aegypti* and *Aedes albopictus* and cause severe epidemics of diseases such as dengue fever (DF) and dengue hemorrhagic fever (DHF) / dengue shock syndrome (DSS) [**1**]. In terms of numbers of infected individuals, dengue fever is probably the most important human viral disease transmitted by arthropod vectors [**2**]. About 40% of the world population living in tropical and subtropical regions of the world, approximately 2.5 billion people, is at risk of infection and of the 1 million cases of dengue hemorrhagic fever per year, about 5% are fatal [**3**].

There is an urgent medical need for the development of a broadly effective and causative anti-viral therapy. A promising approach is the

development of small molecule drugs that inhibit viral enzymes involved in the replication of the virus. Potential target structures for the design of specific inhibitors are present on the dengue virus multifunctional nonstructural protein NS3. The 69 kDa NS3 protein (618 amino acid residues) contains several enzymatic activities activities including a serine protease within the N-terminal portion of 184 amino acid residues (*flavivirin*, EC 3.4.21.91) [**4**] and a polynucleotide - stimulated NTPase / helicase in the C-terminal part of the molecule [**5**]. The functional domains of protease and NTPase / helicase have been shown to overlap within a region of 20 amino acids between residues 160 and 180 of the DEN NS3 protein. The minimal NS3 protease domain was defined to 167 N-terminal residues of NS3 and a cluster of basic amino acids R_{184}KRK is essential for NTPase activity [**6**].

Dengue virus type 2, the most prevalent of 4 serotypes, contains a 10.723 nucleotides single-stranded RNA of positive polarity with a type I cap structure at the 5' - end, which codes for a single polyprotein precursor of 3391 amino acid residues with proteins arranged in the order C-prM-E-NS1-NS2-NS2A-NS2B-NS3-NS4A-NS4B-NS5 [**7**].

Individual viral proteins are generated by a series of co- and posttranslational processing events catalyzed both by host and viral proteases to give at least 10 mature proteins including the structural proteins, C, prM and E and 7 nonstructural proteins, which are expressed in infected cells [**8**, **9**]. The processing of the amino-terminal region of the polyprotein precursor including the structural proteins is catalyzed by a host cell signalase located within the endoplasmic reticulum which cleaves at the C ▽ prM, prM ▽ E, E▽ NS1 and NS4A ▽ NS4B junctions. Cleavage at the NS1 ▽ NS2A site is catalyzed by a host protease within the ER. The final cleavage of prM to M is mediated by another host protease located in an acidic compartment associated with the exocytic pathway and occurs at a late step of virion morphogenesis [**10**].

Co-expression studies have demonstrated that the proteolytic activity of dengue virus NS3 appeared to be dependent on the presence of NS2B [**11**]. An essential domain of 40 amino acid residues within a hydrophilic domain of NS2B has been shown to be responsible for activation *in vivo* and *in vitro* [**12**]. The hydrophobic regions of NS2B are dispensable for protease activity *in vitro*, however, they are required for membrane insertion, which apparently compartmentalizes the NS2B-NS3 functional activity to membranous structures of the ER and for efficient activation of NS3. The heterodimeric complex of NS2B / NS3 catalyzes cleavages at specific sites in the viral polyprotein which have Lys-Arg, Arg-Arg, Arg-Lys and in case of 2B ▽ 3 junction, Gln-Arg, at the P2 and

P1 positions followed by a short side chain amino acid (Gly, Ala, Ser) at the P1' position.

Analysis of polyprotein processing established that the NS3 serine protease, as a heterodimeric complex with the small amphipathic activator protein NS2B (14 kDa), catalyzes cleavages within the viral polyprotein at the dibasic sites 2A▼2B, 2B▼3, 3▼4A and 4B▼5 in the polyprotein [**13**]. The viral protease has been shown to catalyze additional cleavages within NS2A, the capsid protein C and within a C-terminal region of NS3 itself [**14**, **15**, **16**].

Experimental procedures for the cloning and overexpression in recombinant *E. coli*, purification and refolding of the NS2B-NS3 protease components have been established. All published purification methods make use of NS3 fusion proteins tagged with a poly-histidine sequence suitable for metal chelate affinity chromatography. The authors have described a purification method for the NS2B-NS3 full-length protein, NS2B and NS3(pro) using Ni - chelate chromatography which allows to obtain the target proteins from inclusion bodies at > 90% purity by a single step procedure [**17**].

The crystal structure of the dengue virus type 2 NS3 protease domain has been resolved at 2.1 A [**18**]. The x-ray structure revealed the two six-stranded β-barrel structure characteristic of serine proteases of the "SA" clan. Comparison of the DEN 2 protease structure with the structure of the hepatitis C virus (HCV) NS3 protease revealed significant differences between the two enzymes. The most noteable is that the structure of the DEN NS3 protease *without* bound NS2B cofactor resembles the structure of the HCV NS3 protease in its co-complex *with* the NS4A cofactor, indicating major differences in the mode of cofactor activation and substrate specificity [**19**]. HCV protease prefers a Cys residue at the P1 position of the substrates, whereas basic residues are preferred by the DEN NS3 protease. Although both the amino- and carboxyl-terminal β-barrels are 6-stranded in both DEN and HCV proteases, the strands in the N-terminal domain of DEN NS3 are shorter and the barrel is more deformed.

A recombinant fusion protein containing the the hydrophilic core region of NS2B (H-40 or "core fragment" CF-40) linked to the N-terminal 184 residues of the serine protease domain of NS3 (NS3[pro]) exhibits autoproteolytic processing at the NS2B/NS3 cleavage site upon purification and refolding from *E. coli* [**20**]. The activation effect has been studied with an assay system for *trans* - cleavage using tripeptide substrates labeled with an AMC (4-methylcoumaryl-7-amide) fluorescence reporter group. The presence of the NS2B core sequence resulted in a 3300 to 7600-fold activation of the NS3 protease activity. The observed k_{cat} / K_m values for the substrates tBoc-G-R-R-AMC and tBoc-Q-R-R-AMC were 172 and 107 s^{-1} M^{-1} for the NS2B-NS3(pro) fusion protein and 0.052

and 0.014 s^{-1} M^{-1} for the protease domain without NS2B (NS3[pro]), respectively. Significant activity was observed only with the NS2B-NS3(pro) fusion protein in trans - cleavage reactions with a cloned native NS4B/NS5 precursor substrate. However, a $(His)_6$NS3(pro) protein could cleave the chromogenic 'model' substrate BAPA (N-α-benzoyl-L-arginine-p-nitroanilide) after purification and refolding with greater specific activity than the NS2B-NS3(pro) fusion complex, indicating significant differences in the conformation of the substrate binding pocket between the NS2B-NS3 complex and NS3 alone. So far, no x-ray structure is available for the DEN NS2B-NS3 heterodimeric complex.

The project originally proposed had identified the following objectives to initiate research work on the dengue virus NS3 protease: **i**) cloning, overexpression and development of purification methods for the NS2B/NS3 two-component protease, **ii**) development of an enzymatic assay for the NS3 protease and evaluation of substrates and **iii**) collection of initial kinetic and mechanistic data for the enzyme. The report presented here shows that we could successfully establish methods for the purification and enzymatic analysis of the NS3 protease and that progress has been made towards an assay system useful for the characterization of substrate requirements for this important viral drug target enzyme.

Experimental Procedures

Cloning and expression of the NS2B, NS3(pro) and NS2B/NS3 proteins in *Escherichia coli*

Plasmid expression constructs. All dengue virus polyprotein sequences were generated by PCR. Oligonucleotide primers used for PCR were obtained from Bio-Synthesis, USA, and PCR reactions were run on a thermal cycler Gene Amp System 2400 (Perkin Elmer). Two half-genome cDNA clones of DEN 2 strain 16681 in plasmid pBluescript II KS constructed by RT PCR (Nopporn Sitthisombut, Dept. of Microbiology, Chiang Mai University) were used for the amplification of NS2B and NS3(pro) sequences; plasmid pD2/IC-30P was obtained from Dr. Sirithorn Butrapet, Center for Vaccine Development, Mahidol University, and was used as template for the amplification of NS2B-NS3. PCR products were analysed by restriction enzyme digestion and all PCR amplified regions were verified by DNA sequencing on a ABI Model 377 DNA Sequencer (Perkin Elmer). DEN 2 sequences were subcloned into expression vector pTrcHisA

(NS2B-NS3) to give pTH/NS2B-3 and pTrcHisB to give pTH2B, and pTH3p, respectively. Expression constructs were transformed into *E. coli* host strains TOP10 (Invitrogen), JM109, C41(DE3) [22] and BL21(DE3)pLysS. The enzymatically inactive mutant NS2B-NS3(pro)M with a S135A exchange in the NS3 sequence was generated by site-directed mutagenesis following the procedure of the QuickChange™ mutagenesis kit (Stratagene, LaJolla, USA). Complementary mutagenic oligonucleotide primers 5'-CCTGGAACTGCAGGATCTCC-AATTATCG-3' and 5'-CGATAATTGGAGATCCTGCAG-TTCCAGG-3' introducing simultaneously a *Pst*I site suitable for restriction screening were purchased from Genset Inc., Singapore and PCR reactions were run on a thermal cycler Gene Amp System 2400 (Perkin Elmer, Norwalk, USA). The sequence of the mutated NS3(pro) gene was verified by DNA sequencing using a Perkin Elmer ABI prism 377 sequencer (Perkin Elmer, Norwalk, USA).

Expression, purification and on-column refolding of dengue virus 2 proteins. DEN 2 fusion proteins were purified in analogy to published procedures [23, 24]. *E. coli* cells transformed with expression plasmids were grown in LB medium containing ampicillin (100 μg/ml) at 37°C until the optical density at 600 nm reached 0.6. Expression was induced with isopropyl-β-D-thiogalacto-pyranoside (IPTG, 0.1 mM) for 6-8 hrs. Cells were harvested by centrifugation, resuspended in phosphate buffered saline buffer containing lysozyme (100 μg/ml) and lysed in a French pressure cell at 14,000 p.s.i. The lysate was treated with RNaseA and DNase at 5 μg/ml and was kept on ice for 15 min. Inclusion bodies were harvested by centrifugation and the pellet fraction was washed 3 times with 20 mM sodium phosphate, pH 7.4, 1% (v/v) triton X-100. Inclusion bodies were solubilized either in guanidinium lysis buffer (6 M guanidinium hydrochloride, 20 mM sodium phosphate, pH 7.8, 500 mM NaCl) or in buffer A (8 M urea, 20 mM sodium phosphate, pH 7.8, 500 mM NaCl). The solution was applied to a Probond Ni^{2+} - IDA affinity column (Invitrogen) containing 2 ml of resin preequilibrated with 20 ml of buffer A. The column was washed twice with 5 ml of buffer A followed by washing with 20 ml of buffer B (buffer A with 20 mM imidazole). The column was washed with 5 ml of buffer C (buffer A without urea) and was kept overnight at 4°C. Contaminates were removed by 2 washes with buffer C containing 50 mM and 100 mM imidazole, respectively. DEN proteins fused to the polyHis tag were eluted from the column in buffer C with 500 mM imidazole. Fractions of 1 ml were collected and the elution profile was monitored at A_{280}. Peak fractions were pooled and desalted by using a PD10 column (Pharmacia) equilibrated with 50 mM Tris-HCl, pH 7.4. Protein concentrations were determined using a Bio-Rad protein quantitation kit and samples were analyzed on SDS-PAGE. Proteins separated on gels were

probed for the presence of the polyHis fusion tag on Western blots with Ni^{2+} - NTA conjugated to alkaline phosphatase. Preparations were stored in 50 mM Tris-HCl, pH 7.4, 40% glycerol at -20°C.

Cloning and expression of the NS3-substrate 'CSP'

The native dengue polyprotein precursor encompassing the NS2A/2B and NS2B/3 junctions between bases 3881 and 4574 on the dengue type 2 strain 16681 sequence were amplified by using plasmid pD2/IC-30P as a template for PCR. All primers were obtained from Biosynthesis Inc., USA, and sequences were 5'-GGAAGAGGATCCATTGGCAGTG-3' for the forward and 5'-ATAGGCTCAAGCTT-CCAGTTAAGCC-3' for the reverse primer, respectively. The PCR product encoding CSP (cleavage site probe) was cloned into pUC18 and verified by DNA sequencing. The fragment was subcloned into expression vector pTrcHisB. The CSP protein was expressed in *E. coli* JM 109 and cross-reacted with Ni-conjugated NTA in Western blots. The protein was purified to approximately 90% purity by Ni - chelate affinity chromatography on a ProBond™ resin (Invitrogen) and was labeled with fluorescein isothiocyanate. CSP (10 mg ml^{-1} in 50 mM Na_2CO_3, pH 8.5) was incubated with freshly prepared FITC at a final concentration of 50 μg ml^{-1} at 25°C for 3 hrs. in the dark. Unreacted FITC was removed by gel filtration on a PD10 column (Pharmacia) and the labeled CSP was stored at 4°C in 50 mM Tris-Cl, pH 7.4.

Cloning and expression of NS3(pro) with a C-terminal polyhistidine tag

The sequence encoding NS3(pro) was obtained by PCR using the plasmid pUC-NS3(pro) as a template and the primers 5'- GACCATGATTACGAATTCGAGCT-CGG-3' (forward) and 5'-GACTCTAGAGGATCCTCTTTCGGAAAATG-3' (reverse). The reverse primer changed the stop codon TAG present in the template sequence back to the original AAG codon. The pCR product was digested with BamHI and subcloned into pTH2A expression vector (Invitrogen). *E. coli* C41(DE3) was transformed and expression was induced with 1mM IPTG. The protein was purified under denaturing conditions as described above.

Cloning and expression of NS2B(H)-NS3(pro)

The fusion protein consisting of the hydrophilic core domain of NS2B (NS2B(H)) and the NS3 protease domain NS3(pro) was constructed by SOE-PCR (splicing overlap extension). Plasmid pTH/NS2B3f containing the full-length NS2B/NS3 protein was used as template for PCR. Primer NS2B(H)soe-f 5'-TGCT-

CACTGGAGGATCCGCCGATTTGGAACTGGAG-3' (bases 4276-4293 in DEN 2) incorporated a BamHI restriction site and primer NS2B(H)soe-r 5'-CTTCACTT-CCCACAGGTACCACAGTGTTTGTTCTTCCTC-3' (nucleotides 4498-4512) a KpnI restriction site. PCR reactions contained in 100 µl 20 ng template, 20 pMol of each primer, 200 mM dNTPs and 1.5 U Pfu DNA polymerase (Stratagene, USA). PCR was run for 30 cycles at 30 sec/96^{0}C, 30 sec/63^{0}C and 30 sec/68^{0}C. The 170 bp PCR product generates an overlapping sequence with the primer NS3(pro)soe-f 5'-GAACAAACACTGTGGTACCTGTGGGAAGTGAAGAAAC-3' (4404-4416) and the primer NS3(pro)soe-r 5'-CTTCTCTTTCAGGATCCCTAATCTTCGATCTCTCTGGGTTG-3' (5043-5061). For the PCR amplification of 550 bp NS3(pro), 30 cycles were run with annealing for 30 sec at 58^{0}C and extension for 1.10 min at 68^{0}C. The SOE-PCR was run with the two templates using the flanking primers NS2B(H)soe-f and NS3(pro)soe-r. 30 cycles of the SOE-PCR were performed with an annealing temperature of 58^{0}C/30 sec and an extension temperature of 68^{0}C/1.40 min. After digestion with BamHI, the resulting SOE-PCR product (0.73 kbp) was cloned into vector pTrcHisA digested with BamHI giving the plasmid pTH/NS2B(H)3(pro). The entire sequence of the fusion gene was verified by using an ABI 377 DNA sequencer (Perkin Elmer). Recombinant clones in in *E. coli* BL21 were analyzed by SDS-PAGE for protein expression under IPTG (1 mM) induction. Autoproteolytic cleavage of the fusion protein at the 2B/3 site was detected after refolding from 8M urea.

Cloning and expression of NS3(pro) in the methylotrophic yeast *Pichia pastoris*

Construction of expression plasmids. *Pichia pastoris* strains KM71H and GS115 were purchased from Invitrogen. All yeast protocols were performed according to the instructions given in the Invitrogen *Pichia* cloning manual. The plasmid pPICZαA was used for the construction of recombinant proteins containing a yeast α-factor secretion signal sequence. The plasmid pPICZB was used for the construction of recombinant clones which express the target proteins intracellular. All recombinant plasmids were generated by PCR using oligonucleotide primers (Genset, Singapore Biotech. Pte Ltd.). The secretion plasmid containing NS3(pro) fused the α-factor/polyhistidine tag by a $(glu\text{-}ala)_2$-repeat was constructed by using plasmid pTH-NS2B/3f as template for PCR and primers 5'-AGGGGAATTC-GTCGACAAAAGAGCCGGAGTATTGTGGG-3' and 5'-GGCCTTCTAGACTACTATCGGAA-AATGTCATCTTCG-3'. Typical PCR reactions contained in 100 µl 10 µM dNTPs, 40 pMol primers, 50 ng template plasmid and 3 U Pfu DNA polymerase (Stratagene, USA). 30 cycles were performed at

95°C/1.0 min, 40°C/30 sec and 72°C/1.5 min. An extra extension of 7 min at 72°C was added at the end of to the cycles. The PCR product was ligated into the yeast shuttle vectors and cells were transformed either by electro-transformation according to the manufacturer's instructions with a Gene-Pulser (Biorad) using 1.500 V, 25 μF and 200 Ω with pulse lengths of 5 to 10 msec, or by the 'Easy-Comp' method with polyethylene glycol under brief heat shock conditions (42°C for 10 min). Zeocin-resistant colonies were probed by PCR for the presence of stable genomic integration products by using the primers 5' AOX1 5'-GACTGGTTCCAATTGACAAGC-3' and 3' AOX1 5'-GCAAATGGCATTCT-GACATCC-3'.

Expression of NS3(pro) in Pichia. *Pichia pastoris* KM71H harbouring pPICZαA-3p was incubated in BMGY medium (1.34% yeast nitrogen base, 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1% glycerol) at 30°C under shaking at 300 rpm for 2 days. The culture was transferred to fresh BMGY medium and incubated until the OD_{600} was 5. The cells were harvested by centrifugation at 10.000 g for 10 min at room temperature and resuspended in 10 to 20 ml of BMMY medium (1.34% yeast nitrogen base, 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1% v/v MeOH). Incubation was continued for another 5 days with absolute MeOH added every 24 hrs to give a 3% final concentration. Samples of cells and the of the medium supernatant were collected every 24 hrs and kept at -80°C until they were subjected to protein analysis. The medium supernatant was concentrated by addition of 100% TCA to a final concentration of 1% and samples were kept on ice for 30 min. After centrifugation, the protein pellet was resuspended in 50 ml of 0.1 N NaOH. SDS-PAGE sample buffer (50 μl) was added and the mixture was heated to 95°C for 10 min. To analyze intracellular accumulation of the NS3(pro) fusion protein, cells were harvested at OD_{600}=1.0, suspended in sample buffer, heated to 95°C for 10 min and a sample volume corresponding to 0.2 OD_{600} were loaded onto a 15% SDS-PAGE gel. Protein bands from SDS-PAGE were transferred onto nitrocellulose filters by using a Trans-Blot SD semi-dry transfer cell (Biorad, USA). Blots were probed either with an anti-polyhistidine (C-terminal) antibody or with anti - c-myc antibody (Invitrogen) at 1:10.000 dilution for 2 hrs at room temperature. Bands were visualized by using alkaline-phosphatase conjugated anti - rabbit IgG (Sigma Chemistry) at 1:1000 dilution. NS3(pro) fuson protein expressed in *Pichia* was purified by using a modification of the HiTrap chelating column metal affinity chromtography. Post-translational modification of the protein by N-linked glycosylation was probed by using endoglycosidase H. NS3(pro) was incubated in 10x denaturation buffer (5% SDS, 10% β-mercaptoethanol) at 95°C for 10 min. 1/10 volume of 10x G5 buffer (sodium citrate, pH 5.5) and 100 U of endoglycosidase H (New England Biolab)

were added. The reaction was incubated for 18 hrs and analyzed on 15% SDS-PAGE.

Protease assay with fluorescent tripepetides labeled with AMC

Samples containing NS3 protease were assayed for enzymatic activity by using commercially available peptides with dibasic residues carrying a fluorescent 4-methylcoumaryl-7-amide (AMC) reporter group. tBoc-gly-arg-arg-AMC was obtained from Sigma Chemistry. Protein samples (0.1 - 10 μg total protein) were incubated in microtiter plates at room temperature for various periods of time. Reaction mixtures contained in 100 μl final volume: assay buffer (50 mM Tris-Cl, pH 8.0, 25 mM NaCl), and 2.5 μM GRR-AMC. Reactions were continuously monitored in 10 min - intervalls by a Labsystems Fluoroscan II fluorometer using a filter pair of fixed wavelengths set to λ_{EX} = 355 nm and λ_{EM} = 460 nm. Control reactions contained trypsin (Sigma Chemistry) at 1 μg ml^{-1}. A linear calibration curve for AMC concentrations between 0.625 μM and 5 μM was generated by using a stock solution (10 mM) prepared in MeOH.

Protease assay using HPLC with dansylated peptide substrates

Peptide Substrates. 12mer peptide substrates encompassing the native dengue virus type 2 strain 16681 polyprotein 2A/B, 2B/3, 3/4A and 4/B cleavage sites containing a N-terminal dansyl group (Dns) were commercially synthesized by Biosynthesis Inc., Lewisville, USA., and purified by reversed-phase HPLC. A 6mer cleavage product peptide, Dns-FAAGRK, of the 3/4A site was used as retention reference for HPLC analysis. The monobasic peptide Dns-FAAGRASLTLNL containing a K→A substitution at the P1 position served as negative control in cleavage assays with DEN NS3 protease. Peptides were solubilized in water or absolute MeOH.

HPLC NS3 Protease Assays. Reaction mixtures for the detection of NS3 proteolytic activity were composed of 100 mM Tris-HCl, pH 8.0, containing 10 μM of peptide substrate and 2-4 μg of purified NS2B-NS3(pro) protein in a final volume of 40 μl. Cleavage controls contained trypsin (Sigma Chemistry, St. Louis, USA,) at a concentration of 5 μg ml^{-1}. Samples were incubated at room temperature (25°C) for various periods of time and reactions were quenched by the addition of acetonitrile to give a final concentration of 20%. Samples were loaded on a reversed-phase HPLC (Waters) and separated on a C-18 phenomenex column (250 mm x 4.6 mm, Pharmacia) equilibrated with 0.1% (vol./vol.) TFA/water. Reaction products were separated by a linear gradient of 2% min^{-1} of 0.1% TFA/acetonitrile. Products were analyzed by UV-detection at

214 and 280 nm and by a fluorescence monitor (Waters 474) set to excitation wavelength λ = 340 nm and emission wavelength λ = 510 nm. The extend of hydrolysis was estimated from the decrease in area of the substrate peak. Injection of standard amounts of substrate was used to calibrate peak areas.

RESULTS AND DISCUSSION

Cloning, expression and purification of DEN virus proteins. The initial objective of this work was to clone and overexpress the components of the NS2B/NS3 protease complex and to develop and establish methods for the assay of NS3 proteolytic activity. We expected based on analogous findings with the hepatitis C virus (HCV) protease, that the unliganded form of the NS3 protease and the NS2B/NS3 heterodimeric complex would exhibit different specificities for polyprotein-derived substrates and that the activity of the protease towards 'native' substrates would be enhanced in the presence of the NS2B cofactor. We have therefore decided to clone, express and purify the NS3 protease and the NS2B cofactor separately in order to generate a 2-component system suitable for an *in vitro* activation assay, and we have cloned the full-length NS2B/NS3 and NS2B/NS3(pro) proteins as well as the enzymatically inactive S135A mutant of NS3 [(NS2B/NS3(pro)M]. The latter contains only 184 amino acid residues of the N-terminal protease domain, which have been shown to be sufficient for NS3 serine protease activity.

A methodological uniform and relatively straightforward approach was employed to construct expression clones for the target proteins in *E. coli*. Plasmid clones of dengue virus serotype 2 strain 16681 were used as templates for the PCR amplification of the desired sequences. The 2.3 kb sequence of NS2B/NS3 was amplified by using a dengue virus genomic cDNA plasmid [**21**], which was kindly provided by Dr. Siritorn Butrapet, Center for Vaccine Development, Mahidol University. Dengue cDNA cloned in pBluescript was kindly donated by Dr. Nopporn Sitthisombut, Dept. of Microbiology, Chiang Mai University, and were used to amplify by PCR sequences for NS2B and NS3(pro). After cloning and sequence analysis of the recombinant constructs in pUC18 vector, the gene encoding NS2B/NS3f (full-length) was subcloned in the expression vector pTrcHisA, and the sequences encoding NS2B, NS3(pro) and NS2B/NS3(pro) were subcloned in pTrcHisB, were they were fused to a 3.5 kb hexahistidine affinity purification tag. Expression of the recombinant proteins was tested in several *E. coli* host strains including TOP10, JM109, C41(DE3)

[**22**]and BL21(DE3)pLysS. Although expression of the DEN proteins was observed with all host strains, expression levels were clearly strain-dependent and were relatively higher in the ompT protease strains BL21 and C41 and polypeptides of the expected sizes were produced predominantly as inclusion bodies. In *E. coli* BL21, reduced growth rates were observed with the recombinant plasmid indicative of toxic effects of the heterologous proteins to *E. coli* host. Interactions of the hydrophobic sequences of the viral proteins with membranes of the host have been observed in case of the Japanese encephalitis virus NS2B protein, which is closely related to dengue virus [**23**]. The NS2B/NS3 full-length protein was obtained from *E. coli* JM 109, whereas NS2B and NS3(pro) were produced in *E. coli* C41. Expression of DEN proteins produced in recombinant *E. coli* was analyzed by SDS-PAGE and subsequent Western blotting by using Ni-NTA conjugated to alkaline phosphatase as probe for the presence of the His6x tag. The DEN proteins exhibited strong cross-reactivity with the probe for the affinity tag as well as with antibodies directed against the polyHis sequence. The Western analysis allowed a relatively fast and specific determination of the target proteins on SDS-PAGE gels and a semi-quantitative estimation of expression yields.

A purification protocol for the DEN proteins was developed, which was based on detergent extraction of the inclusion bodies, solubilization in urea or guanidinium hydrochloride, chromatographic purification by immobilized Ni-metal chelate chromatography and refolding by stepwise dialysis. Contrary to our expectations, no autocleavage of the NS2B/NS3 fusion protein was observed at the 2B/3 site in *E. coli*. We therefore assumed that the protein eventually may have not folded to a native conformation which was competent for autoprocessing at the polyprotein junction. Detergent-extracted inclusion bodies were solubilized in 8 M urea and the the material was purified on ProBond™ Ni-metal chelate affinity columns by washing with buffers containing 50 and 100 mM imidazole, respectively. The target proteins were eluted in the presence of 500 mM imidazole and the elution profiles were similar for all DEN proteins, indicating that the separation was largely based on the presence of only the His-tag and not on the pI or molecular weight of the individual proteins. After Ni-metal chelate chromatography, the samples were desalted by using a PD10 desalting column and samples were subjected to SDS-PAGE. Based on band intensities in Coomassie-stained gels, estimated yields were 0.5, 6.0 and 1.0 mg l^{-1} from a bacterial culture grown to OD_{600} = 1.0 for NS2B, NS3(pro) and NS2B/NS3, respectively. All proteins were recovered at a level of purity greater than 95% as judged by densitometry of SDS-PAGE gels. NS2B/NS3f was recovered as a homogenous protein of 88 kDa molecular weight and cleavage

products were not detectable after purification. The polyHis tag could not be cleaved off the purified DEN 2 fusion proteins with enterokinase and the recombinant NS2B(H)/NS3(pro) protein has been shown to undergo autoproteolytic processing after purification and refolding from *E. coli* [**20**]. This observation indicated, that the proteins recovered from our purification method may not have folded to the correct native conformation. The development of a suitable refolding method for the DEN proteins was a relatively tedious task in the absence of active enzyme useful for comparative enzyme assays. Among several methods which were explored, the step-dialysis in buffers containing decreased concentrations of urea yielded finally active enzyme, which was demonstrated by using an assay with synthetic peptide substrates (*see below*).

Expression and purification of 'CSP'. In order to generate a sustrate for the NS3 protease, we have initially produced the polyprotein fragment designated 'CSP' (cleavage site probe), which is a 220 residues part of the native DEN polyprotein which contains the NS2A/NS2B and NS2B/NS3 polyprotein cleavage sites. This protein would represent a 'native' substrate for the DEN NS3 protease in contrast to short, synthetic peptides since its folding resembles the original polyprotein substrate. Moreover, the CSP gene sequence would be amenable to modification by site-directed mutagenesis thereby generating derivatives of the cleavage site sequences useful for a mechanistic analysis for the NS3 protease. Analogous to the NS2B/NS3 proteins, the sequence encoding CSP was obtained by PCR and the protein was expressed from pTrcHisB. Expression of CSP in *E. coli* yielded a strong band at 28 kDa associated with the inclusion fraction and cross-reacting with Ni-NTA / AP. The protein was purified by affinity chromatography and subsequently labeled with fluorescein isothiocyanate. The fluorescent cleavage products would allow the identification of NS3 activity against a background of unlabeled proteins, for example in complex samples like DEN - infected cells.

Peptide-based enzymatic assays for NS3. In parallel to the purification and refolding experiments, an assay system had to be set up to monitor NS3 protease enzyme activity recovered from the purification procedures. The first report on NS3 *in vitro* determination was published by *Yusof et al.* [20]. The authors had used commercially available short tripepetides, which were labeled with an AMC fluorescent reporter group. Their publication had demonstrated that catalytic efficiency of NS3 was drastically enhanced (3.300 - 7.600fold) in the presence of the NS2B cofactor and that the NS2B protein was indispensable for cleavage of a native polyprotein substrate. However, cleavage of a small protease model substrate, BAPA, was more efficient by NS3(pro) alone when compared to the NS2B/NS3(pro) co-complex. These findings have provided

evidence to suggets the existence of 2 discrete conformations of the DEN protease with different substrate specificites, making both forms of the enzyme a target for inhibitor design. We have tested the substrate G-R-R-AMC with preparations of our enzyme and could - within the experimental error - reproduce the kinetic values obtained by *Yusof et al.*. We have assayed the fusion protein with the peptide substrate G-R-R-AMC and the K_m - values obtained for this substrate are comparable (approx. 190 μM).The fusion protein containing the 40 residues - hydrophilic activation domain of NS2B with the protease domain [NS2B(H)/NS3(pro)] was constructed by SOE-PCR. The protein we have generated is identical to the fusion protein used by *Yusof et al.* and this protein undergoes autoproteolytic cleavage upon purification and refolding from overexpressing *E. coli*, thereby independently confirming the initial findingsreported in the literature.

Short tripepetides would not necessarily represent the best substrates for the NS3 protease, since presence of the NS2B cofactor in the liganded form of the enzyme presumably extends the enzyme-substrate interactions beyond the P2 and P2' residues. Peptides with a longer sequence would therefore satisfy all enzyme-substrate interactions in the NS2B/NS3 co-complex. We have therefore decided to establish a proteolysis assay which is based on synthetic 12mer peptide substrates mimicking native polyprotein sequences. Discontinuous, HPLC - assays using peptide substrates are a standard technique in the field. A number of virus proteases have been assayed by using HPLC - analysis of cleavage peptides [**24**, **25**, **26**]. In order to facilitate detection, the peptides were N-terminally labeled with a dansyl (Dns) - reporter group which allows detection by UV and very sensitive detection by fluorescence. Purified NS2B-NS3(pro) was incubated with the cleavage site peptides corresponding to the 2A/B, 2B/3, 3/4A and 4B/5 cleavage sites of the DEN 2 polyprotein which were separated from the reaction products by reversed-phase HPLC using a C_{18} - column. By using fluorescence detection, a decrease of the substrate peak and a concomitant increase of the product signal was observed upon incubation of the substrates with NS2B-NS3(pro). In all cases, only a single peak of product appeared after incubation of NS2B-NS3(pro) with the substrates, whereas the control reactions with trypsin produced multiple signals according to the presence of at least two basic residues in the substrates. A 6mer cleavage product peptide FAAGRK of the 3/4A polyprotein site was synthesized and used as retention time marker for HPLC analysis. The cleavage peptide produced in NS3-assays containing the 3/4A substrate co-eluted with this reference peptide in the HPLC separation. Incubation of the NS2B-NS3(pro) protein with a control peptide, F-A-A-G-R-A-S-L-T-L-N-L, containing a K→A substitution at the P1 position of the 3/4A cleavage