Proteolytic Processing of the Nonstructural Proteins of Dengue 2 Virus

Thesis by

Frank Preugschat

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1990

(Submitted May 23, 1990)

Dedication

To the memory of my father.

Acknowledgements

First and foremost, many thanks to Jim and Ellen Strauss for their kindness, patience and guidance. To those people who read theses acknowledgements, I simply say that Jim and Ellen are two of the best kept secrets of the graduate program. I thank my thesis committee members, Drs. Charles Brokaw, Norman Davidson, Elliot Meyerowitz and Ellen Rothenberg for their feedback and advice. Special thanks to Norman Davidson for highly recommending Jim's lab as a place for graduate study.

When you spend about twelve years of your life in higher education, there are a lot of people that you need to thank for helping you maintain perspective. I'd like to thank everyone collectively and extend special thanks to: Mike Smith and Imre Kovesdi at Simon Fraser for turning me on to molecular biology. Norma Foster for introducing me to my better half. Dave Ruff and Jamie Lee for sharing the graduate experience during the early years.

I've learned a lot of different things from a number of people that have passed through the Strauss lab and I extend my appreciation to all of you. Richard Kuhn kept the lab up and running and was a constant source of energy, ideas and enthusiasm. Bert Niesters provided some good company and sage advice. Chen-Wen Yao played some mean ping pong, and generated useful antibody reagents. Reef Hardy and Raoul de Groot provided insight with their work on the Sindbis nonstructural proteinase. Edith Lenches provided technical support when I needed it most, and Zhang Hong did a lot to keep things running smoothly. Kangsheng Wang, Nathalie Pardignon, Young and Chang Hahn, Yukio Shirako and Randy Levinson provided some good company and different perspectives on science. Special thanks to the staff of the animal care, monoclonal

antibody, microchemical, and protein sequencing facilities for all their assistance in carrying out experiments.

Finally to my family, thanks for your love and support over all these years.

Abstract

The genomes of many positive stranded RNA viruses and of all retroviruses are translated as large polyproteins which are proteolytically processed by cellular and viral proteases. Viral proteases are structurally related to two families of cellular proteases, the pepsin-like and trypsin-like proteases. This thesis describes the proteolytic processing of several nonstructural proteins of dengue 2 virus, a representative member of the *Flaviviridae,* and describes methods for transcribing full-length genomic RNA of dengue 2 virus. Chapter 1 describes the *in vitro* processing of the nonstructural proteins NS2A, NS2B and NS3. Chapter 2 describes a system that allows identification of residues within the protease that are directly or indirectly involved with substrate recognition. Chapter 3 describes methods to produce genome length dengue 2 RNA from cDNA templates.

The nonstructural protein NS3 is structurally related to viral trypsin-like proteases from the alpha-, picorna-, poty-, and pestiviruses. The hypothesis that the flavivirus nonstructural protein NS3 is a viral proteinase that generates the termini of several nonstructural proteins was tested using an efficient *in vitro* expression system and antisera specific for the nonstructural proteins NS2B and NS3. A series of cDNA constructs was transcribed using T7 RNA polymerase and the RNA translated in reticulocyte lysates. Proteolytic processing occurred *in vitro* to generate NS2B and NS3. The amino termini of NS2B and NS3 produced *in vitro* were found to be the same as the termini of NS2B and NS3 isolated from infected cells. Deletion analysis of cDNA constructs localized the protease domain necessary and sufficient for correct cleavage to the first 184 amino acids of NS3. Kinetic analysis of processing events *in vitro* and experiments to examine the sensitivity of processing to dilution suggested that an

intramolecular cleavage between NS2A and NS2B preceded an intramolecular cleavage between NS2B and NS3. The data from these expression experiments confirm that NS3 is the viral proteinase responsible for cleavage events generating the amino termini of NS2B and NS3 and presumably for cleavages generating the termini of NS4A and NS5 as well.

Biochemical and genetic experiments using viral proteinases have defined the sequence requirements for cleavage site recognition, but have not identified residues within proteinases that interact with substrates. A biochemical assay was developed that could identify residues which were important for substrate recognition. Chimeric proteases between yellow fever and dengue 2 were constructed that allowed mapping of regions involved in substrate recognition, and site directed mutagenesis was used to modulate processing efficiency.

Expression *in vitro* revealed that the dengue protease domain efficiently processes the yellow fever polyprotein between NS2A and NS2B and between NS2B and NS3, but that the reciprocal construct is inactive. The dengue protease processes yellow fever cleavage sites more efficiently than dengue cleavage sites, suggesting that suboptimal cleavage efficiency may be used to increase levels of processing intermediates *in vivo*. By mutagenizing the putative substrate binding pocket it was possible to change the substrate specificity of the yellow fever protease; changing a minimum of three amino acids in the yellow fever protease enabled it to recognize dengue cleavage sites. This system allows identification of residues which are directly or indirectly involved with enzyme-substrate interaction, does not require a crystal structure, and can define the substrate preferences of individual members of a viral proteinase family.

Full-length cDNA clones, from which infectious RNA can be transcribed, have been developed for a number of positive strand RNA viruses, including the flavivirus type virus, yellow fever. The technology necessary to transcribe genomic RNA of dengue 2 virus was developed in order to better understand the molecular biology of the dengue subgroup. A 5' structural region clone was engineered to transcribe authentic dengue RNA that contains an additional 1 or 2 residues at the 5' end. A 3' nonstructural region clone was engineered to allow production of run off transcripts, and to allow directional ligation with the 5' structural region clone. *In vitro* ligation and transcription produces full-length genomic RNA which is noninfectious when transfected into mammalian tissue culture cells. Alternative methods for constructing cDNA clones and recovering live dengue virus are discussed.

viii

Table of Contents

Page

Introduction 1

Chapter 1 *In Vitro* Processing of Dengue 2 Nonstructural Proteins NS2A, NS2B and NS3. 29

Chapter 2 Studying Enzyme-Substrate Interaction Using Chimeric Proteinases: Identification of an Intragenic Locus Important for Substrate Recognition. 79

Chapter 3 *In Vitro* Transcription of Genomic RNA from Dengue 2 cDNA Templates. 115

Introduction

Historical and medical overview: The dengue viruses comprise an antigenic subgroup of the flaviviruses which are responsible for worldwide epidemics of increasing proportion (9, 10, 36). The first documented outbreak of dengue-like illness in the Americas occurred in Philadelphia in 1780, and major epidemics of dengue-like illnesses have been reported throughout the tropical and subtropical world over the last two centuries (11, 29, 34, 35, 68). In 1903, dengue-like illness was shown to be transmissible by mosquitoes and was identified as a filterable agent (virus) in 1926. Dengue virus was isolated from the blood of infected soldiers in 1944 and was the second flavivirus to be isolated (the type virus, yellow fever, was isolated in 1926). Viral isolates were serologically distinguishable and were designated as the prototypic dengue 2 (New Guinea C) and dengue 1 (Hawaii) strains. Dengue 3 and 4 serotypes were isolated in Manila, in 1956, during an epidemic of dengue hemorrhagic fever (67).

The three primary hosts of dengue viruses are man, lower primates and mosquitoes. Transmission to man by mosquito occurs during a blood meal taken by an infected female of the genus *Aedes*. Classical dengue fever ensues suddenly after a 3- to 15-day incubation period. Chills, headache, postorbital and lower back pain are usually observed and the most severe manifestations are accompanied by extreme joint pain and prostration. The fever and other symptoms persist for two to three days, then decrease and are followed by a second attack of fever, which is usually accompanied by a rash. The presence of fever, headache and rash constitute what is known as the the dengue triad, and are characteristic for positive diagnosis (69).

Simple uncomplicated dengue fever is rarely fatal, but a secondary manifestation of shock (dengue shock syndrome or DSS) or hemorrhagic fever (DHF) is more serious and can be fatal (5, 8, 16). The factors contributing to DSS and DHF are not well understood, but it is widely held that secondary infection with a different serotype of dengue virus is necessary. An immune response directed against one serotype will cross-react extensively with the other three serotypes but will not provide cross-protection. It has been proposed that subneutralizing titers of cross-reacting antibodies aid the establishment of infection by a secondary serotype (39, 40).

Dengue hemorrhagic fever is the second leading cause of death attributed to infectious disease in southeast Asia and efforts to identify molecular markers of virulence and biochemical targets for antiviral therapy have intensified as the world-wide risk has increased. The incidence of dengue infection, dengue shock syndrome and dengue hemorrhagic fever has been increasing rapidly over the last few decades and with over 1.5 billion people at risk, dengue represents a serious health concern for people who live in tropical or subtropical areas. Unfortunately there is no animal model for either classical dengue fever or the more severe hemorrhagic fever and the presence of four different, but related serotypes has hampered development of an efficacious vaccine. The techniques of molecular biology have provided a much better understanding of dengue gene expression and have offered new hope for the development of effective antivirals and vaccines (2, 39, 77).

Virion structure and gene expression: The dengue virion is composed of three protein species (C, M and E), and a single RNA molecule (80). The major structural protein of the virion is a glycosylated, membrane bound envelope protein that is capable of eliciting neutralizing antibodies and is responsible for hemagglutination and cell fusion acitivities (74, 78, 83). The membrane or M protein is a small hydrophobic, integral membrane protein whose role in virion structure is not well understood. A small, basic capsid protein complexes the genomic RNA with roughly icosahedral symmetry. The 5' end of the positive stranded genomic RNA has a type 1 cap and the 3' end is not polyadenylated (14, 38, 90).

Dengue viruses are capable of entering cells through simple fusion with cellular membranes, although receptors may be utilized to enter specialized cells (41, 82). Once internalized, the genomic RNA is released into the cytoplasm and viral gene expression begins. Viral gene expression does not shut down host cell synthesis and induces the proliferation of Golgi-associated membranous vesicles (79). Genomic RNA is copied from a full-length minus strand template and is the sole detectable viral messenger RNA (15, 81). After a latent period of eight to twelve hours, mature virions are assembled and released from the surface of infected cells by an unknown mechanism (42).

The genomic organization and polyprotein processing of the dengue 2 PR159 S1 vaccine strain are detailed in Figure 1 (3, 4, 27, 28, 38, 71). A single large polyprotein is translated from an approximately ten kilobase long open reading frame. The structural proteins are translated before the nonstructural proteins and all proteins are translated

Fig. 1. Genome organization and expression of dengue 2 virus. The upper half of the figure details the genome organization of the PR159 S1 vaccine strain. The lower half of the figure depicts the processing events generating the structural and nonstructural proteins. Virion structural proteins are designated by dark cross-hatched boxes and nonstructural proteins are designated by open boxes. Cleavages after basic residues are designated by downward facing arrowheads. Host cell signalase cleavages are designated by open ball and stick icons and potential viral encoded signalase-like cleavages are designated by a shaded ball and stick icon. A late, membrane associated cleavage generating the membrane structural protein and mature capsid protein are designated by a ball and arrowhead.

DENGUE 2 GENOME
(10,712nt)
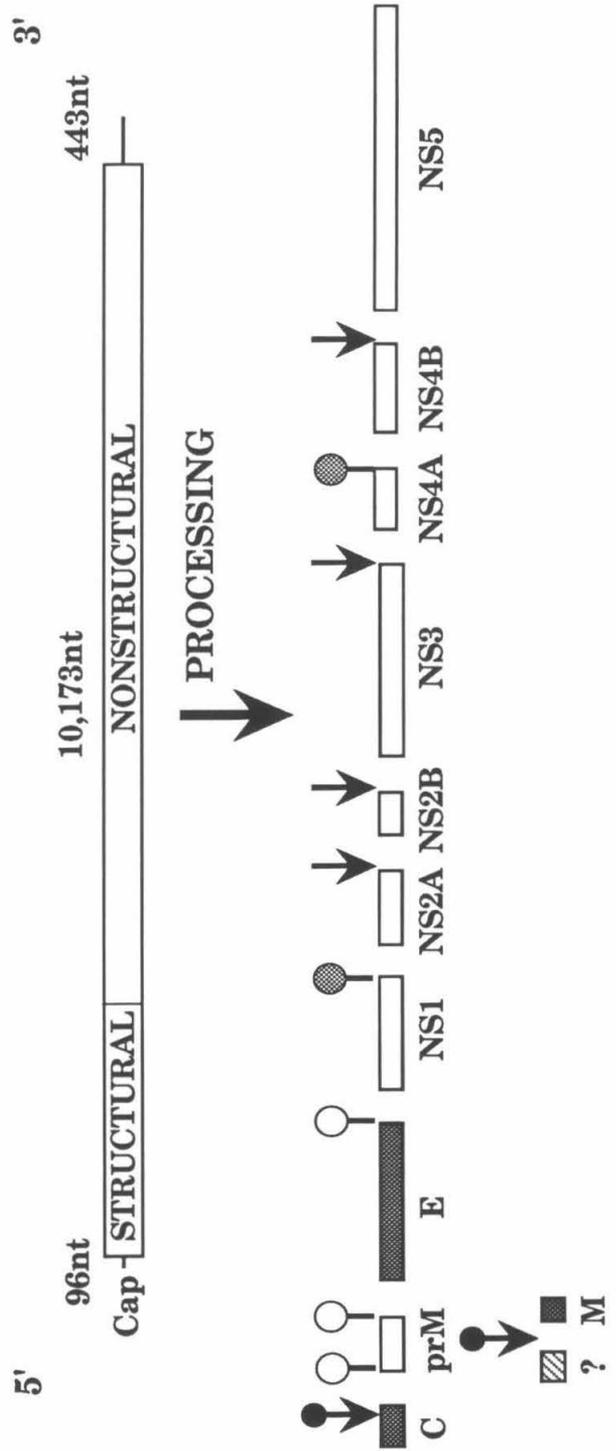
in stoichiometric amounts. The dengue polyprotein is co- and posttranslationally processed by cellular and viral proteinases into ten identifiable gene products (73).

Membrane associated processing: Historically flaviviral polyprotein processing has been difficult to study because of fast processing kinetics *in vivo*, and *in vitro* expression experiments have been hampered by poor fidelity and low efficiency. Recently several groups have utilized improved expression systems to analyze signalase mediated cleavage events (31, 51, 56, 65, 66) and have identified the flaviviral nonstructural proteinase (Chapter 1).

Cleavages maturing the dengue polyprotein can be classified into two main groups, membrane associated and cytosolic cleavages. Membrane associated cleavages can be further classified into three subgroups; cotranslational host cell signalase cleavages, posttranslational host cell or viral signalase-like cleavages, and late cleavages associated with virion assembly. Several groups have studied the processing events generating the structural proteins and have concluded that host cell signalases process the structural and some of the nonstructural proteins (51, 56, 66). Host cell signalases cleave co-translationally between the capsid protein and prM, between prM and E, and between E and NS1 to generate the amino termini of prM, E and NS1. Closely spaced and alternating hydrophobic stop transfer and signal sequences lie between prM and E, and between E and NS1, to direct the processed proteins into the lumen of the endoplasmic reticulum.

The cleavage event which generates the carboxy terminus of the nonstructural protein, NS1 and the amino terminus of NS2A is not well understood. Residues preceeding the 1/2A cleavage site are not preferred by signalase and cleavage is not always cotranslational (59, 85). NS2A sequences are required *in cis* for processing to occur at the 1/2A boundary, and it has been suggested that NS2A may be a viral proteinase (31). The amino terminus of NS4B may also be generated by a posttranslational host cell or viral signalase-like cleavage. A polyprotein consistent in molecular weight and immunoreactivity with NS4A4B can be detected in infected cells suggesting that the signalase-like cleavage is not cotranslational and it is possible that the amino termini of NS2A and NS4B are generated by the same proteinase (13).

The virion structural proteins C and M are processed late in the assembly process by unidentified proteinases. Processing occurs between two basic amino acids and a small side chain amino acid, which is similar to the processing of the nonstructural proteins (see below). prM is initially glycosylated in the Golgi aparatus and then moves to the cell surface where it forms a heterodimer with the envelope protein (89). A late cleavage event removes the glycosylated segment of prM to generate mature M protein (72). The cleavage of prM is sensitive to inhibition by acidotropic amines and it has been hypothesized that prM could be an autoprotease (60, 64). Intracellular forms of the capsid protein possess a C terminal membrane spanning segment that is cleaved prior to assembly of the virion, by an unknown mechanism (56, 75). Cleavage occurs after several basic amino acids and is presumed to be mediated by the same proteinase which processes the nonstructural proteins (64).

Processing in the cytosol: Processing of viral proteins in the cytosol is usually mediated by a viral proteinase. Many positive strand RNA viruses encode one or more proteinases which process the structural and nonstructural proteins (44, 48, 87). Viral proteinases exhibit strict substrate specificity and this selectivity is used to regulate the production of viral proteins from a polyprotein precursor. Processing intermediates may have functions distinct from the end products and proteinases provide a mechanism for fine tuning viral gene regulation (21, 92).

The amino termini of NS2B, NS3, NS4A and NS5 are generated in the cytosol, by cleavage between an arginine or lysine, and a serine, glycine or alanine. In an effort to identify the flaviviral nonstructural proteinase, two laboratories proposed models which identified the amino terminus of NS3 to be a trypsin-like protease, based upon limited sequence and potential secondary structure similarity with known cellular and viral proteinases (7, 32). Sequence alignments and structural predictions of the model are presented in Figure 2.

A 180 amino acid domain at the amino terminus of the nonstructural protein, NS3, can be aligned using a relaxed sequence and secondary structure profile to known cellular and viral trypsin-like proteinases. The classification of NS3 as a trypsin-like serine proteinase identified it as a member of a large family of viral and cellular proteinases that use either a catalytic serine or cysteine at the active site. The trypsin-like serine proteinase family includes proteinases from the alpha-, flavi-, picorna-, poty-, and pestiviruses. The sequence/structure model predicts that amino acids 51, 75 and 135 of NS3 are elements of a classic histidine,

aspartate and serine catalytic triad. Five amino acids within the putative substrate binding pocket are predicted to be in direct contact with bound substrate, and in particular aspartate 129 is proposed to interact directly with arginine or lysine residues within cleavage sites.

The protease model was tested by *in vitro* experiments which examined cleavages generating the amino termini of NS2B and NS3. Monospecific antisera reactive with NS2B and NS3 were raised in rabbits and were used to study processing in an *in vitro* expression system. Deletion analysis of the protease domain localized the sequences necessary and sufficient for cleavage at the 2A/2B and 2B/3 junctions to the first 184 amino acids of NS3. Cleavages generating NS2B and NS3 occur intramolecularly and with a defined order and complete fidelity *in vitro*. The implications of these results with respect to the nature and order of flaviviral polyprotein cleavage are discussed in Chapter 1.

Figure 2. Sequence and structural alignment of cell and viral proteinases: The large cellular proteases and the amino terminal protease domains of flaviviruses (NS3) are aligned according to the protease model to emphasize sequence and structural conservation. Dots represent matches and dashes represent insertions. The cellular proteases are aligned relative to elastase and the flavivirus NS3 protease domains are aligned relative to dengue 2. The catalytic residues are numbered using chymotrypsin (above) and dengue NS3 (below) numbering schemes, and are boxed and designated using an asterisk. Core beta strands that are conserved among the trypsin-like serine proteinases are indicated by wavy lines below the sequence alignment. Residues which are predicted to be in direct contact with bound substrate are designated by an accent symbol (^). Cellular proteinase sequences are referenced in (7). Virus abbreviations and sequences used for viral proteinase alignments are as follows: DEN2, dengue 2 PR159 S1 (38); DEN3, dengue 3 (57); DEN4, dengue 4 (49); MVE, Murray Valley encephalitis (19); KUN, Kunjin (17); SLE, Saint Louis encephalitis (84); WN, West Nile (12); YF, yellow fever Asibi strain (37); JE, Japanese encephalitis Beijing 1 strain (43) and TBE, tick borne encephalitis (50).

BOX 1 BOX 2 BOX 3 SBP

CELLULAR PROTEASES

ELASTASE
S. GRISEUS TRP
TRYPSIN
CHYMOTRYPSIN
S. AUREUS V8

FLAVIVIRUS NS3

DEN2
DEN3
DEN4
MVE
KUN
WN
YF
JE
TBE

BOX 1

57
*

```
QNWVMTAAHCVDRE---LTFRVVVG
.DI.L....SGSGNNTSITATG.
SQ..VS....KS----GIQ.RL.
E..V...GVT----TSDV..A.
KDTLL.NK.V.--ATHGDPHALKA

EGTFHTMWHVT-----RGAVLMHKG
.V.-------..TNH....
.V.-------.S.IC.ET
.V..L..T.----.AI.SGE
.V..L..T.----K.A..SGE
.V..L..T.----K.A..SGE
.G.V..-------.F.VRN.
.NV..L..T.----.AI.SGE
K.VL...-------.A.SIDD
51
```

BOX 2

102
*

```
GYDIALLRLAQ
.K.W...K..
NN..M.IK.KS
NN..T.K.ST
EG.L.IVKFSP

KKDLISYGGGW
.........
RN.M......
.E..VT...P
.E.RLC...P
.E.RLC...P
.E..VA...S
.E.R.A...P
RE..VVC...A
75
```

BOX 3

195
*

```
GVRSGCQGDSGGPLHCLVNG-QYAVHGVTSFVSR
.GVDT....MFRKD.ADEWIQV.IV.WGYG
.GKDS.M....VV.SG----KLQ.IV.WG.G
SGV.S.M...LV.KK..-AWTLV.IV.WG.S
-DL.TTG.N..S.VFNEK.--E.I.IHWGGVP

VSLDFSPGTSGSPIVDKKG----KVVGLYGNGVV
AIL.K....INRE.......
.T...K....INR.......I
.E...YPI...NSN......EII
.E...YPT...N......D.I
.T..PT....N......D.I
.A..YPS...NRN......E.I
.E..YPR...L.SN......DII
IPI.LVK....LNAQ....V.
135
```

SBP

> W

DEN2
DEN3
DEN4
MVE I
KUN I
WN I
YF I
JE .....E
TBE .....LK

The nature of substrate recognition: Viral proteinases have been chemically synthesized or purified from bacterial and eukaryotic systems, and biochemical assays have been developed to investigate enzyme-substrate interactions (47, 55, 70, 91). Purified proteinases have been used to cleave synthetic or natural substrates (46, 47, 58), and site-directed mutagenesis has provided a genetic approach for defining the sequence and position requirements of cleavage (24, 25). Biochemical and genetic studies using retro-, picorna-, and potyviral proteinases have shown that the substrate requirements of viral proteinases can vary widely and the molecular basis of viral proteinase specificity is not understood. In general, it is clear that the sequence surrounding the scissile bond (P1 and P'1 residues) and the local structural context of a viral polyprotein are important determinants of proteinase specificity.

The HIV (human immunodeficiency virus) and RSV (Rous sarcoma virus) proteinases are small aspartic proteinases whose crystal structures and biochemical properties have been intensively studied (46, 52, 53). The P1 and P'1 positions of retroviral cleavage sites are not conserved between retroviruses or even within a retroviral polyprotein, yet HIV proteinase can discriminate between HIV and RSV synthetic peptide substrates (47). This suggests that sequence (however poorly conserved) as well as structural context are important for substrate recognition. Molecular modeling studies using the HIV crystal structure have provided clues as to how retroviral proteinases can recognize a wide variety of cleavage sites and still maintain high selectivity (86).

The TEV (tobacco etch virus) 49kDa proteinase is a member of a large trypsin-like proteinase family (6). Site-directed mutagenesis and

biochemical inhibition studies have provided evidence that the active site contains a catalytic cysteine residue (26). The 49kDa proteinase processes five sites within the TEV polyprotein and it is unique among the viral proteinases in that it recognizes an extended seven amino acid cleavage site. The P6, P3 and P1 positions within TEV cleavage sites are conserved and are absolutely required for cleavage, while nonconserved residues modulate cleavage efficiency (24). A small cleavage site cassette can be moved to a new position within a polyprotein and will function as a substrate, suggesting that sequence and not structure dominates substrate recognition (25).

The poliovirus 3C proteinase is a trypsin-like proteinase (with a catalytic cysteine residue) which cleaves nine out of thirteen potential glutamine-glycine pairs within the poliovirus polyprotein (54). Purified 3C will also cleave between glutamine-alanine, glutamine-serine, glutamine-leucine and glutamate-glycine pairs, suggesting that structural determinants play an important role in substrate recognition (55, 93). The role of structural determinants in enzyme-substrate interactions has been investigated using site-directed mutagenesis and chimeric proteinases. Site-directed mutagenesis of the 3C/D cleavage site provided evidence that 3CD and not 3C is the proteinase responsible for cleaving the structural (P1) polyprotein and that processing intermediates have substrate specificities different than products (92). Experiments with chimeric polio-, rhino-, and coxsackie virus 3C and 3CD proteinases provided additional evidence that determinants of structural polyprotein processing are different than those of nonstructural polyprotein processing (22).

All of the genetic and biochemical experiments discussed above have concentrated on defining what constitutes a cleavage site and have not identified residues within proteinases which are important for substrate recognition. If the three-dimensional structure of the proteinase is known, then residues which are in direct or close contact with substrates can be changed by site-directed mutagenesis and the effect of each individual mutation can be assayed for using biochemical methods (30, 33, 88). Molecular modeling techniques would be potentially useful in predicting which amino acid substitutions are allowable, but the ultimate proof would lie with a biochemical test of each individual mutation (18).

RNA-dependent RNA polymerases lack the ability to proofread the RNA that they synthesize, therefore proteins encoded by RNA viruses undergo rapid evolutionary change (23, 61, 76). Selective pressures operate to conserve protein function but residues which are not absolutely required to maintain structure or catalytic function are freely substituted. Viral proteinases which are released from polyproteins by autocatalytic cleavage are both enzymes and substrates and therefore the substrate binding pocket of the proteinase and cleavage sites must coevolve. A change in the substrate binding pocket of the proteinase must be counterbalanced by a change in a cleavage site to maintain proper order and kinetics of polyprotein processing.

Even though the yellow fever and dengue protease domains have diverged significantly, we took advantage of conserved restriction sites within the NS3 gene to create chimeric protease domains. We asked whether dengue and yellow fever protease domains could process heterologous cleavage sites. Using a series of chimeric cDNA constructs

we localized a region within the protease domain that played an important role in substrate recognition. The polymerase chain reaction (PCR) was used to interchange residues between the dengue and yellow fever protease domains, and the effects of single and multiple mutations were assayed for using an *in vitro* expression system. We concluded that residues 141 to 147 of dengue NS3, corresponding to residues 144 to 150 of yellow fever NS3, define an intragenic locus that is an important determinant of substrate specificity for the flaviviral proteinases. The techniques described in Chapter 2 can be applied to map enzyme-substrate interactions within other viral proteinases.

Infectious clones and reverse genetics: The application of cDNA technology has revolutionized the study of RNA virus genetics. cDNA clones of RNA viruses can be mutagenized and the biological effects of individual mutations can be studied *in vitro* or *in vivo*. Infectious RNA can be transcribed from cDNA clones of a number of positive strand RNA viruses, including the flavivirus type virus, yellow fever (1, 20, 45, 62, 63). In order to better understand the molecular biology of the dengue subgroup we have developed methods to transcribe full-length genomic RNA of dengue 2 virus. Chapter 3 details the techniques and reagents that are necessary to produce full-length genomic RNA.

References Cited

1. **Ahlquist, P., R. French, M. Janda, and L. S. Loesch-Fries.** 1984. Multicomponent RNA plant virus infection derived from cloned viral cDNA. Proc. Natl. Acad. Sci. USA **81:**7066-7070.

2. **Bancroft, W. H.** 1987. Current status of dengue vaccines and prospects for the future. Puerto Rico Health Sci. J. **6:**23-6.

3. **Bancroft, W. H., R. M. Scott, K. H. Eckels, C. H. Hoke, T. E. Simms, K. D. T. Jesrani, P. L. Summers, D. R. Dubois, D. Tsoulos, and P. K. Russell.** 1984. Dengue virus type 2 vaccine: Reactogenicity and immunogenicity in soldiers. J. Inf. Dis. **149:**1005-1010.

4. **Bancroft, W. H., F. H. Top, K. H. Eckels, J. H. Anderson, J. M. McCown, and P. K. Russell.** 1981. Dengue-2 vaccine: Virological, immunological, and clinical responses of six yellow fever-immune recipients. Infect. Immun. **31:**698-703.

5. **Barnes, W. J. S., and L. Rosen.** 1974. Fatal hemorrhagic disease and shock associated with primary dengue infection on a Pacific island. Am. J. Trop. Med. Hyg. **23:**495-506.

6. **Bazan, J. F., and R. J. Fletterick.** 1988. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. Proc. Natl. Acad. Sci. USA **85:**7872-7876.

7. **Bazan, J. F., and R. J. Fletterick.** 1989. Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. Virology **171:**637-639.

8. **Bhamarapravati, N.** 1981. Pathology and pathogenesis of DHF, p. 207-214. In Dengue Hemorrhagic Fever. S. Hotta (ed.), International Center for Medical Research, Kobe University School of Medicine, Kobe, Japan.

9. **Calisher, C. H., N. Karabatsos, J. M. Dalrymple, R. E. Shope, J. S. Porterfield, E. G. Westaway, and W. E. Brandt.** 1989. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. J. Gen. Virol. **70:**37-43.

10. **Calisher, C. H., R. E. Shope, W. Brandt Casals, J., N. Karabatsos, F. A. Murphy, R. B. Tesh, and M. E. Wiebe.** 1980. Proposed antigenic classification of registered arboviruses. I. Togaviridae, *Alphavirus*. Intervirology **14:**229-232.

11. **Carey, D. E.** 1971. Chikungunya and dengue: A case of mistaken identity? J. Hist. Med. **26:**243-262.

12. **Castle, E., U. Leidner, T. Nowak, G. Wengler, and G. Wengler.** 1986. Primary structure of the West Nile flavivirus genome region coding for all the nonstructural proteins. Virology **149:**10-26.

13. **Chambers, T. J., D. W. McCourt, and C. M. Rice.** 1989. Yellow fever virus proteins NS2a, NS2b and NS4b: Identification and partial N-terminal amino acid sequence analysis. Virology **169:**100-109.

14. **Cleaves, G. R., and D. T. Dubin.** 1979. Methylation status of intracellular dengue type 2 40S RNA. Virology **96:**158-165.

15. **Cleaves, G. R., T. E. Ryan, and R. W. Schlesinger.** 1981. Identification and characterization of type 2 dengue virus replicative intermediate form RNAs. Virology **111:**73-83.

16. **Cohen, S. N., and S. B. Halstead.** 1966. Shock associated with dengue infection. I. Clinical and physiologic manifestations of dengue hemorrhagic fever in Thailand, 1964. J. Pediatr. **68:**448-456.

17. **Coia, G., M. D. Parker, G. Speight, M. E. Byrne, and E. G. Westaway.** 1988. Nucleotide and complete amino acid sequences of Kunjin virus:

Definitive gene order and characteristics of the virus-specified proteins. J. Gen. Virol. **69**:1-21.

18. **Craik, C. S., S. Roczniak, S. Sprang, R. Fletterick, and W. J. Rutter.** 1987. Redesigning trypsin via genetic engineering. J. Cell. Biochem. **33**:199-211.

19. **Dalgarno, L., D. W. Trent, J. H. Strauss, and C. M. Rice.** 1986. Partial nucleotide sequence of Murray Valley encephalitis virus: Comparison of the encoded polypeptides with yellow fever virus structural and nonstructural proteins. J. Mol. Biol. **187**:309-323.

20. **Dasmahapatra, B., R. Dasgupta, K. Saunders, B. Selling, T. Gallagher, and P. Kaesberg.** 1986. Infectious RNA derived by transcription from cloned cDNA copies of the genomic RNA of an insect virus. Proc. Natl. Acad. Sci. USA. **83**:63-66.

21. **de Groot, R. J., W. R. Hardy, Y. Shirako, and J. H. Strauss.** 1990. Cleavage site preferences of Sindbis virus polyproteins containing the nonstructural proteinase: Evidence for temporal regulation of polyprotein processing *in vivo*. EMBO, J. in press.

22. **Dewalt, P. G., M. A. Lawson, R. J. Colonno, and B. L. Semler.** 1989. Chimeric picornavirus polyproteins demonstrate a common 3C proteinase substrate specificity. J. Virol. **63**:3444-3452.

23. **Doolittle, R. F., D.-F. Feng, M. S. Johnson, and M. A. McClure** 1989. Origins and evolutionary relationships of retroviruses. Quarterly Review of Biology **64**:1-30.

24. **Dougherty, W. D., and T. D. Parks.** 1989. Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage

sequence in determining the reaction profile at two tobacco etch virus cleavage sites in cell free systems. Virology **172**:145-155.

25. **Dougherty, W. G., S. M. Cary, and T. D. Parks.** 1989. Molecular genetic analysis of a plant virus polyprotein cleavage site: A model. Virology **171**:356-364.

26. **Dougherty, W. G., T. D. Parks, S. M. Cary, J. F. Bazan, and R. J. Fletterick.** 1989. Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. Virology **172**:302-310.

27. **Eckels, K. H., W. E. Brandt, V. R. Harrusin, J. M. McCown, and P. K. Russell.** 1976. Isolation of a temperature-sensitive dengue-2 virus under conditions suitable for vaccine development. Infect. Immun. **14**:1221-1227.

28. **Eckels, K. H., V. R. Harrison, P. L. Summers, and P. K. Russell.** 1980. Dengue-2 vaccine: Preparation from a small-plaque virus clone. Infect. Immun. **27**:175-180.

29. **Ehrenkranz, N., A. Ventura, R. Cuadrado, W. Pond, and J. Porter.** 1971. Pandemic dengue in Caribbean countries and the southern United States--Past, present, and potential problems. New Engl. J. Med. **285**:1460-1469.

30. **Estell, D. A., T. P. Graycar, J. V. Miller, D. B. Powers, J. P. Burnier, P. G. Ng, and J. A. Wells.** 1986. Probing steric and hydrophobic effects on enzyme substrate interactions by protein engineering. Science **233**:659-663.

31. **Falgout, B., R. Chanock, and C. J. Lai.** 1989. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2A. J. Virol. **63**:1852-1860.

32. **Gorbalenya, A. E., A. P. Donchenko, E. V. Koonin, and V. M. Blinov.** 1989. N-Terminal domains of putative helicases of Flavi and Pestiviruses may be serine proteases. Nucl. Acids. Res. **17**:3889-3897.

33. **Graf, L., C. S. Craik, A. Patthy, S. Roczniak, R. J. Fletterick, and W. J. Rutter.** 1987. Selective alteration of substrate specificity by replacement of aspartic acid 189 with lysine in the binding pocket of trypsin. Biochemistry **26**:2616-2623.

34. **Groot, H., A. Morales, M. Romero, H. Vidales, C. Lesmes-Donaldson, G. Marquez, D. Escobar de Calvache, and M. Saenz.** Recent outbreaks of dengue in Colombia, pp. 31-39. In Dengue in the Caribbean. Pan American Health Organization, Washington D.C (ed.), 1977.

35. **Gubler, D. J.** 1987. Dengue and dengue hemorrhagic fever in the Americas. Puerto Rico Health Sci. J. **6**:107-11.

36. **Gubler, D. J.** 1989. *Aedes aegypti* and *Aedes aegypti*-borne disease control in the 1990s: Top down or bottom up. Am. J. Trop. Med. Hyg. **40**:571-578.

37. **Hahn, C. S., J. M. Dalrymple, J. H. Strauss, and C. M. Rice.** 1987. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. Proc. Natl. Acad. Sci. **84**:2019-2023.

38. **Hahn, Y. S., R. Galler, T. Hunkapiller, J. Dalrymple, J. H. Strauss, and E. G. Strauss.** 1988. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. Virology **162**:167-180.

39. **Halstead, S. B.** 1988. Pathogenesis of dengue: Challenges to molecular biology. Science **239**:476-81.

40. **Halstead, S. B., J. S. Chow, and N. J. Marchette.** 1973. Immunological enhancement of dengue virus replication. Nature (London) New Biol. **243:**24-26.

41. **Hase, T., P. L. Summers, and K. H. Eckels.** 1989. Flavivirus entry into cultured mosquito cells and human peripheral blood monocytes. Arch. Virol. **104:**129-43.

42. **Hase, T., P. L. Summers, K. H. Eckels, and W. B. Baze.** 1987. An electron and immunoelectron microscopic study of dengue-2 virus infection of cultured mosquito cells: Maturation events. Arch. Virol. **92:**273-91.

43. **Hashimoto, H., A. Nomoto, K. Watanabe, T. Mori, T. Takezawa, C. Aizawa, T. Takegami, and K. Hiramatsu.** 1988. Molecular cloning and complete nucleotide sequence of the genome of Japanese encephalitis virus Beijing-1 strain. Virus Genes **1:**305-317.

44. **Hellen, C. T., H. Kraeusslich, and E. Wimmer.** 1989. Proteolytic processing of polyproteins in the replication of RNA viruses. Biochemistry. **28:**9881-9890.

45. **Kaplan, G., J. Lubinski, A. Dasgupta, and V. R. Racaniello.** 1985. *In vitro* synthesis of infectious poliovirus RNA. Proc. Natl. Acad. Sci. USA **82:**8424-8428.

46. **Kotler, M., R. A. Katz, W. Danho, J. Leis, and A. M. Skalka.** 1988. Synthetic peptides as substrates and inhibitors of a retroviral protease. Proc. Natl. Acad. Sci. USA. **85:**4185-4189.

47. **Kraeusslich, H. G., R. H. Ingraham, M. T. Skoog, E. Wimmer, P. V. Pallai, and C. A. Carter.** 1989. Activity of purified biosynthetic proteinase

of human immunodeficiency virus on natural and synthetic peptides. Proc. Natl. Acad. Sci. USA **86:**807-811.

48.   **Kraeusslich, H. G., and E. Wimmer.** 1988. Viral proteinases. Ann. Rev. Biochem. **57:**701-754.

49.   **Mackow, E., Y. Makino, B. Zhao, Y.-M. Zhang, L. Markoff, A. Buckler-White, M. Guiler, R. M. Chanock, and C.-J. Lai.** 1987. The nucleotide sequence of dengue type 4 virus: Analysis of genes coding for nonstructural proteins. Virology **159:**217-228.

50.   **Mandl, C. W., F. X. Heinz, E. Stoeckl, and C. Kunz.** 1989. Genome sequence of tick-borne encephalitis virus (western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. Virology **173:**291-301.

51.   **Markoff, L.** 1989. *In vitro* processing of dengue virus structural proteins: Cleavage of the pre-membrane protein. J. Virol. **63:**3345-3352.

52.   **Miller, M., M. Jaskolski, J. K. M. Rao, J. Leis, and A. Wlodawer.** 1989. Crystal structure of a retroviral protease proves relationship to aspartic protease family. Nature **337:**576-579.

53.   **Navia, M. A., P. M. D. Fitzgerald, B. M. McKeever, C. T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer.** 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature **337:**615-620.

54.   **Nicklin, M. J. H., H. Toyoda, M. G. Murray, and E. Wimmer.** 1986. Proteolytic processing in the replication of polio and related viruses. Biotechnology **4:**33-42.

55. **Nicklin, M. J. H., H. Toyoda, M. G. Murray, and E. Wimmer.** 1988. Poliovirus 3C: Large-scale expression, purification and specific cleavage activity on natural and synthetic substrates *in vitro*. J. Virol. **62:**4586-4594.

56. **Nowak, T., P. M. Farber, G. Wengler, and G. Wengler.** 1989. Analyses of the terminal sequences of West Nile virus structural proteins and of the *in vitro* translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. Virology **169:**365-376.

57. **Osatomi, K., and H. Sumiyoshi.** 1990. Complete nucleotide sequence of dengue type 3 virus genome RNA. Virology. in press.

58. **Pallai, P. V., F. Burkhardt, M. Skoog, K. Schreiner, P. Bax, K. A. Cohen, G. Hansen, D. E. H. Palladino, K. S. Harris, M. J. Nicklin, and E. Wimmer.** 1989. Cleavage of synthetic peptides by purified poliovirus 3C proteinase. J. Biol. Chem. **264:**9738-9741.

59. **Perlman, D., and H. O. Halvorson.** 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. **167:**391-409.

60. **Randolph, V. B., G. Winkler, and V. Stollar.** 1990. Acidotropic amines inhibit proteolytic processing of flavivirus prM protein. Virology **174:**450-458.

61. **Reanney, D. C.** 1982. The evolution of RNA viruses. Ann. Revs. Microbiol. **36:**47-73.

62. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers.** 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by *in vitro* ligation. New Biol. **1:**285-296.

63. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: Mapping of lethal mutations, rescue of a temperature-sensitive marker, and *in vitro* mutagenesis to generate defined mutants. J. Virol. **61:**3809-3819.

64. **Rice, C. R., E. G. Strauss, and J. H. Strauss.** Structure of the flavivirus genome, pp. 279-327. In The Togaviridae and Flaviviridae. S. Schlesinger and M. J. Schlesinger (ed.), Plenum Publishing Corp., New York, 1986.

65. **Ruiz-Linares, A., M. Bouloy, M. Girard, and A. Cahour.** 1989. Modulations of the *in vitro* translational efficiencies of yellow fever virus mRNAs: Interactions between coding and noncoding regions. Nucl. Acids. Res.

66. **Ruiz-Linares, A., A. Cahour, P. Despres, M. Girard, and M. Bouloy.** 1989. Processing of the yellow fever virus polyprotein: Role of cellular proteases in the maturation of the structural proteins. J. Virol. **63:**4199-4209.

67. **Sabin, A. B.** 1952. Research on dengue during World War II. Am. J. Trop. Med. Hyg. **1:**30-50.

68. **Schatzmayer, H. G., R. M. R. Nogueira, and A. P. A. Travassos da Rosa.** 1986. An outbreak of dengue virus at Rio de Janeiro (1986). Mem. Inst. Oswaldo Cruz **81:**245-246.

69. **Schlesinger, R. W.** Dengue viruses, S. Hotta (ed.), Virology Monographs. Springer-Verlag, New York, pp. 1-132, 1977.

70. **Schneider, J., and S. B. H. Kent.** 1988. Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. Cell **54**:363-368.

71. **Scott, R. M., K. H. Eckels, W. H. Bancroft, P. L. Summers, J. M. McCown, J. H. Anderson, and P. K. Russell.** 1983. Dengue 2 vaccine: Dose response in volunteers in relation to yellow fever status. J. Inf. Dis. **148**:1055-1060.

72. **Shapiro, D., E. E. Brandt, and P. K. Russell.** 1972. Change involving a viral membrane glycoprotein during morphogenesis of group B arboviruses. Virology **50**:906-911.

73. **Smith, G. W., and P. W. Wright.** 1985. Synthesis of proteins and glycoproteins in dengue type 2 virus-infected Vero and *Aedes albopictus* cells. J. Gen. Virol. **66**:559-571.

74. **Smith, T. J., W. E. Brandt, J. L. Swanson, J. M. McCown, and E. L. Buesher.** 1970. Physical and biological properties of dengue 2 virus and associated antigens. J. Virol. **5**:524-532.

75. **Speight, G., and E. G. Westaway.** 1989. Carboxy terminal analysis of nine proteins specified by the flavivirus Kunjin: Evidence that only the intracellular core protein is truncated. J. Gen. Virol. **70**:2209-2214.

76. **Steinhauer, D. A., and J. J. Holland.** 1987. Rapid evolution of RNA viruses. Ann. Revs. Microbiol. **41**:409-433.

77. **Stephenson, J. R.** 1988. Flavivirus vaccines. Vaccine **6**:471-80.

78. **Stevens, T. M., R. W. Schlesinger, and I. T. Schulze.** 1965. Sudies on the nature of dengue viruses. I. Correlation of particle density, infectivity and RNA content of type 2 virus. Virology **27**:103-112.

79. **Stohlman, S. A., C. L. Wisseman, O. R. Eylar, and D. J. Silverman.** 1975. Dengue virus induced modifications of host cell membranes. J. Virol. **16:**1017-1026.

80. **Stollar, V.** 1969. Studies on the nature of dengue viruses IV: The structural proteins of type 2 dengue virus. Virology **39:**426-438.

81. **Stollar, V., R. W. Schlesinger, and T. M. Stevens.** 1967. Studies on the nature of dengue viruses III: RNA synthesis in cells infected with type 2 dengue virus. Virology **33:**650-658.

82. **Summers, P. L., W. H. Cohen, M. M. Ruiz, T. Hase, and K. H. Eckels.** 1989. Flaviviruses can mediate fusion from without in *Aedes albopictus* mosquito cell cultures. Virus Res. **12:**383-392.

83. **Sweet, B. H., and A. B. Sabin.** 1954. Properties and antigenic relationships of hemagglutinins associated with dengue viruses. J. Immunol. **73:**363-373.

84. **Trent, D. W., R. M. Kinney, B. J. Johnson, A. V. Vorndam, J. A. Grant, V. Deubel, C. M. Rice, and C. Hahn.** 1987. Partial nucleotide sequence of St. Louis encephalitis virus RNA: Structural proteins, NS1, NS2A, and NS2B. Virology **156:**293-304.

85. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucl. Acids Res. **14:**4683-4690.

86. **Weber, I. T., M. Miller, M. Jaskolski, J. Leis, A. M. Skalka, and A. Wlodawer.** 1989. Molecular modeling of the HIV-1 protease and its substrate binding site. Science **243:**928-931.

87. **Wellink, J., and A. van Kammen.** 1988. Proteases involved in the processing of viral polyproteins. Arch. Virol. **98:**1-26.

88. **Wells, J. A., B. C. Cunningham, T. P. Graycar, and D. A. Estell.** 1987. Recruitment of substrate specificity properties from one enzyme into a related one by protein engineering. Proc. Natl. Acad. Sci. USA **84:**5167-5171.

89. **Wengler, G., and G. Wengler.** 1989. Cell-associated West Nile flavivirus is covered with E and pre-M protein heterodimers which are destroyed and reorganized by proteolytic cleavage during virus release. J. Virol. **63:**2521-2526.

90. **Wengler, G., G. Wengler, and J. Gross.** 1978. Studies on virus specific nucleic acids synthesized in vertebrate and mosquito cells infected with flaviviruses. Virology. **89:**423-437.

91. **Yoshinaka, Y., I. Katoh, T. D. Copeland, G. W. Smythers, and S. Oroszlan.** 1986. Bovine leukemia virus protease: Purification, chemical analysis, and *in vitro* processing of gag precursor polyproteins. J. Virol. **57:**826-832.

92. **Ypma-Wong, M. F., P. G. Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166:**265-270.

93. **Ypma-Wong, M. F., D. J. Filman, J. M. Hogle, and B. L. Semler.** 1988. Structural domains of the poliovirus polyprotein are major determinants for proteolytic cleavage at Gln-Gly pairs. J. Biol. Chem. **263:**17846-17856.

Chapter 1

In Vitro Processing of Dengue 2 Nonstructural Proteins NS2A, NS2B and NS3.

This chapter was prepared for submission.

In Vitro Processing of Dengue 2 Nonstructural Proteins NS2A, NS2B and NS3.

Frank Preugschat, Chen-Wen Yao[1], and James H. Strauss*.

Division of Biology, 156-29
California Institute of Technology
Pasadena, CA 91125.

[1]*Present Address: Department of Molecular Biology, Institute of Preventive Medicine, National Defense Medical Center, Taipei, Taiwan.*

Running title: Flavivirus nonstructural protein processing.

* Author to whom correspondence should be addressed.

Abstract

We have tested the hypothesis that the flavivirus nonstructural protein NS3 is a viral proteinase that generates the termini of several nonstructural proteins, using an efficient *in vitro* expression system and monospecific antisera directed against the nonstructural proteins NS2B and NS3. A series of cDNA constructs was transcribed using T7 RNA polymerase and the RNA translated in reticulocyte lysates. The resulting protein patterns indicated that proteolytic processing occurred *in vitro* to generate NS2B and NS3. The amino termini of NS2B and NS3 produced *in vitro* were found to be the same as the termini of NS2B and NS3 isolated from infected cells. Deletion analysis of cDNA constructs localized the protease domain within NS3 to the first 184 amino acids. Kinetic analysis of processing events *in vitro* and experiments to examine the sensitivity of processing to dilution suggested that an intramolecular cleavage between NS2A and NS2B preceded an intramolecular cleavage between NS2B and NS3. The data from these expression experiments confirm that NS3 is the viral proteinase responsible for cleavage events generating the amino termini of NS2B and NS3 and presumably for cleavages generating the termini of NS4A and NS5 as well.

Introduction

The dengue viruses belong to the *Flaviviridae*, a family of approximately 70 viruses, most of which are arthropod-borne, which can be grouped into 8 antigenic complexes (6, 7). There are four serotypes of dengue, and all are involved in worldwide epidemics of increasing proportion (17, 23, 44). Our laboratory has been working with the PR159 S1 (candidate vaccine) strain of dengue 2 virus (1, 2, 16, 26, 45). The PR159 genome is a positive stranded RNA molecule 10,712 nucleotides in length that is capped at the 5' terminus with a type 1 cap and has at the 3' terminus a uridine residue (11, 26, 52). Translation begins at nucleotide 97 and continues for 10,173 bases to produce a 3391 amino acid polyprotein that is both co- and posttranslationally processed into at least 10 different polypeptides (46).

Sequence analysis of flavivirus proteins and *in vitro* expression experiments have implicated host cell signalases in the amino terminal processing of prM, E, NS1 and possibly NS2A and NS4B (4, 9, 12, 18, 22, 34, 36, 43, 53). These host cell signalases cleave on the carboxy terminal side of hydrophobic leader sequences and function in the lumen of the endoplasmic reticulum (50). On the other hand, the cleavages that generate the amino termini of NS2B, NS3, NS4A and NS5, which follow two basic amino acids and occur in the cytosol, are believed to be due to a viral proteinase (38, 39).

The short half-lives of polyprotein precursors *in vivo* combined with the poor fidelity of *in vitro* expression systems has made the study of flaviviral polyprotein processing difficult. Early experiments that translated genomic RNA *in vitro* did not produce detectable amounts of

nonstructural proteins and correct processing of structural proteins occurred only in the presence of exogenously supplied microsomal membranes (49, 51). *In vivo* pulse chase experiments that utilized amino acid analogs or starved cells prior to amino acid labeling were successful at detecting dengue-specific higher molecular weight proteins, but the lack of specific immune reagents and the inability to chase these putative precursors limited the experimental analysis (10, 37). Recent studies utilizing improved expression systems have made progress in analyzing signalase mediated cleavage events but have not addressed the cleavages that occur after dibasic amino acids (18, 34, 36, 43).

Recently two laboratories have proposed models predicting that the amino terminus of NS3 is a trypsin-like protease (3, 19), based upon limited sequence similarity to cellular serine proteases and more extensive similarity to the protease domain of the capsid protein of the alphavirus Sindbis virus (24, 25). These molecular modeling studies predict that His-51, Asp-75, and Ser-135 of NS3 form a classic serine protease catalytic triad, and that the entire protease domain lies within the first 180 amino acids of this protein. The proteolytic activity of this NS3 enzyme is thought to be responsible for cleavages generating the amino termini of NS2B, NS3, NS4A, and NS5, and possibly the carboxy terminus of the capsid protein. To test this model, we have developed an efficient *in vitro* expression system for NS2B and NS3 and generated specific immune reagents reactive with these proteins. In this system, the processing events that generate the amino termini of NS2B and NS3 occur faithfully *in vitro*, mediated by the nonstructural protein NS3. By deletion analysis the protease domain has been mapped to within the first 184 amino acids of

NS3, and from the kinetics of cleavage we propose that an intramolecular cleavage between NS2A and NS2B precedes an intramolecular cleavage between NS2B and NS3.

Materials and Methods

Construction of *trpE* fusions: cDNA clones of the PR159 strain of dengue 2 virus were used for all plasmid constructions (26). All plasmids were constructed using standard recombinant DNA techniques and were purified on CsCl density gradients (33). To create a gene fusion between *trpE* and dengue NS2B, the vector pATH3 (48) was digested with *Bam*HI, the 3' recessed end was filled in using *Escherichia coli* DNA polymerase I Klenow fragment, and then digested with *Cla*I. A *Stu*I to *Hpa*II fragment (from cDNA clone 2, referred to here as pDN2) containing nucleotides 4077 to 4463 of the dengue genome was inserted to yield the construct pTNS2.

To create a gene fusion with NS3, the vector pATH3 was digested with *Eco*RI, Klenow treated, and digested with *Sal*I. An *Asp*718 (Klenow treated) to *Sal*I fragment (from cDNA clone pDN2) containing nucleotides 4497 to 4944 of the dengue genome was inserted to produce pTNS3. Recombinant clones were screened by analysis of protein expression patterns, restriction endonuclease digestion, and DNA sequencing of miniprep DNAs.

Expression of fusion proteins and immunizations: *trpE* fusion proteins were prepared from large-scale induced cultures essentially as described (27) with the following modifications. All *trpE* plasmids were propagated in bacterial strain XL1 (Stratagene). After induction, pelleted cells were twice frozen and thawed in 50 mM Tris (pH 7.5), 0.5 mM EDTA, 300 mM NaCl (TEN) buffer. To promote efficient lysis cells were treated for 15 minutes in a sonicating water bath (Bransonic 12) at 4°C. Insoluble inclusion bodies were purified by two successive pelletings through a 25%

sucrose, 1 mM EDTA cushion and were solubilized directly into sodium dodecylsulfate (SDS) loading buffer (32) for electrophoresis in preparative 10% polyacrylamide SDS gels (SDS PAGE). Proteins were visualized by staining with 250 mM KCl at 4°C. Gels were homogenized in sterile phosphate buffered saline (pH 7.5) and mixed with MPL adjuvant (RIBI Immunochem) prior to injection. New Zealand white rabbits were injected intramuscularly and subcutaneously on a 3 to 4 week schedule using 100 to 150 μg of fusion protein per injection session and were bled 10 to 14 days after boosting. Serum aliquots were stored at -20°C until further use.

Cells and virus stocks: Stocks of the PR159 dengue 2 S1 isolate (1, 2, 45) were prepared on *Aedes albopictus* C6/36 cells (13). C6/36 cells were propagated at 30°C in Dulbecco's modified Eagle medium containing 10% fetal calf serum and supplemented with nonessential amino acids. BHK-21 clone 15 cells, obtained from Dr. Joel Dalrymple, USAMRIID, Bethesda MD, were used to plaque purify the virus and to prepare infected cell lysates for immunoprecipitation. BHK cells were propagated at 37°C in minimum essential medium supplemented with nonessential amino acids and containing 5% fetal calf serum. Plaque assays were performed at 34°C on BHK cells using a 1% low melting temperature agarose (Seakem) overlay containing minimum essential medium and 5% fetal calf serum. Plaques were readily visible after 4 days by direct visualization or after staining the monolayer with neutral red.

Labeling of infected cells: Equivalent conditions were used to prepare labeled viral proteins for immunoprecipitation assays and amino

terminal sequencing. BHK cells were infected with a multiplicity of 5 in minimum essential medium. At 30 hours post-infection cells were labeled in minimum essential medium containing 2% dialyzed fetal calf serum, 1/40 the normal concentration of methionine and 75 µCi/ml of [^{35}S]methionine (Amersham) for 12 hours. For preparation of leucine-labeled proteins, cells were incubated in minimum essential medium containing 2% dialyzed serum, 1/100 the normal amount of leucine and 75 µCi/ml of [^{3}H]leucine (New England Nuclear) for 12 hours. Cell lysates were prepared by solubilizing the monolayers in a denaturing lysis buffer containing 0.5% SDS, 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 20 µg/ml PMSF (Sigma), 2 µg/ml aprotinin (Boehringer Mannheim), and 2 µg/ml leupeptin (Sigma).

Translation constructs: Vector p5'L213 was constructed to facilitate high level expression of dengue sequences in an *in vitro* translation system. Briefly an *Eco*RI to *Pst*I fragment encoding nucleotides 16 to 213 of the dengue 2 genome (from pDNC3) was cloned into *Eco*RI and *Pst*I digested pGEM1 (Promega) to yield pGem16/213. A double stranded DNA oligonucleotide was synthesized that contained a T7 RNA polymerase promoter immediately adjacent to the first 55 nucleotides of the dengue genome (the sequence of the first 15 nucleotides was assumed to be the same as for the dengue 2 Jamaica strain (14), and was inserted into *Sac*I/*Pvu*II digested pGem16/213, yielding vector p5'213. p5'213 and pGEM4 were digested with *Pst*I and *Nhe*I and polylinker sequences of pGem4 (a *Pst*I to *Nhe*I fragment) were inserted to yield vector p5'L213.

Vector p5'L213 contains several unique restriction sites that allow fusion of heterologous sequences with the dengue capsid gene reading frame.

Vector p5'L213 was digested with *Pst*I, the 3' overhang made blunt by treatment with T4 DNA polymerase, and digested with *Eco*RI. An *Ase*I (Klenow treated) to *Xho*I digested fragment containing nucleotides 3778 to 5427 of the dengue genome, from cDNA clone pDN2, and a *Xho*I to *Eco*RI fragment containing nucleotides 5428 to 6351 of the dengue genome, from cDNA clone pDN5, were inserted to create pT10. This construct contains the 5' untranslated region of the dengue genome and the nucleotides encoding the first 37 amino acids of the dengue capsid protein fused in frame to the 101st codon of NS2A, followed by the sequences encoding the remainder of NS2A (amino acids 101-218), the complete coding sequences of NS2B (amino acids 1-129), and the first 610 amino acids of NS3 (the carboxy terminal 8 amino acids of NS3, amino acids 611-618, are not present).

Proteins produced *in vitro* with the same termini as *in vivo* have been designated by the prefix NS whereas proteins with an altered structure have been designated by the prefix P and the use of a prime superscript. For example, pT10-programmed translations yield the precursor P2A2B3, the processing intermediate P2B3, and the products P2A, NS2B and P3. Precursors, intermediates and products generated by the proteolytically active deletion constructs pT11 and pT12 described below, are designated P2A2B3', P2A2B3", P2B3', P2B3", P2A, NS2B, P3' and P3", respectively.

Plasmid pT10 was used to produce deletion clones to map the boundaries of the protease domain, using *Xho*I, *Asu*II, *Sal*I and *Kpn*I

(*Asp* 718) restriction sites within NS2B and NS3. In each case pT10 was digested with EcoRI and the second enzyme (*Xho*I, *Asu*II etc.), the 3' recessed ends were filled in with Klenow, and the plasmid recircularized using T4 DNA ligase to create the deleted constructs designated pT11, pT12, pT13, and pT14 respectively.

Two additional deletion clones were constructed that together removed almost all of the NS2A and NS2B coding sequences represented in pT10. pT11 was digested with *Kpn*I, blunted with T4 DNA polymerase, then digested with *Sna*BI and cyclized to create pT15. pT12 was digested with *Pvu*II and *Sna*BI, then cyclized to create pT16.

***In vitro* transcription and translation:** All plasmids used for *in vitro* transcription and translation were purified by CsCl density gradient centrifugation. After linearization with appropriate restriction enzymes, templates were digested with proteinase K and phenol extracted. Ethanol precipitated templates were resuspended in diethylpyrocarbonate treated, ribonuclease free water at a concentration of 0.5 µg/ml. Transcription reactions containing 0.5 mM rNTPs (Pharmacia), 0.5 mM cap analog m^{7}G(5')ppp(5')G (New England Biolabs), 50 ng/µl of template, and 2 units/µl of T7 RNA polymerase (Pharmacia) in 1X transcription buffer (41) were incubated at 37°C for 45 minutes. Quality of synthesized RNA was routinely checked by gel electrophoresis prior to *in vitro* translation, and the mass of RNA synthesized was quantitated by inclusion of trace amounts of [^{3}H]rGTP in the transcription mix. RNA yields were approximately 5 to 10 µg of RNA per µg of input template. Micrococcal nuclease treated rabbit reticulocyte lysates containing 0.025 Eq/µl

microsomal membranes (Promega) were programmed with *in vitro* transcribed RNA at a concentration of 5 to 10 μg/ml. All translations were carried out at 30°C for 90 minutes unless otherwise noted. Products of pT10-programmed translations were excised from SDS PAGE gels and the radioactivity was quantitated by liquid scintillation counting.

Immunoprecipitation: The amount of antiserum necessary to quantitatively immunoprecipitate radiolabeled viral proteins from cell lysates and *in vitro* translation mixes was empirically determined. Labeled cell lysates were diluted 5 fold with radioimmune precipitation (RIPA) buffer (1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl, 50 mM Tris HCl pH 7.5) prior to immunoprecipitation, and protein A agarose beads (Boehringer Mannheim) were used at a 5-fold binding excess over immune sera (30). After a 1-hour incubation at room temperature, the immune complexes were washed twice with a RIPA buffer containing 1 mg/ml bovine serum albumin and once with a RIPA buffer lacking albumin. For immunoprecipitation of *in vitro* translations, reticulocyte lysates were diluted 25-fold in denaturing lysis buffer and heated to 70°C for 5 minutes prior to dilution 5-fold with a RIPA buffer. Immunoprecipitated proteins were eluted by boiling in an SDS loading buffer. Samples were analyzed by 15% SDS PAGE and gels were fluorographed at -80°C (8).

Sequencing of viral proteins: Radiolabeled NS2B was prepared from infected BHK cell lysates by preparative immunoprecipitation. 2×10^7 cells labeled with [^{35}S]methionine or 6×10^7 cells labeled with [^{3}H]leucine were

lysed in 4 ml of denaturing lysis buffer, then diluted in a RIPA buffer for immunoprecipitation. Immunoprecipitates were washed three times with RIPA and twice with 10 mM ammonium bicarbonate buffer (pH 7.0) prior to being eluted into 250 µl of 0.1% trifluoroacetic acid and blotted onto glass fiber discs (29). *In vitro* processed NS2B was isolated by preparative immunoprecipitation of denatured reticulocyte lysates. 50 µl of a pT11-programmed [^{35}S]methionine- or 200 µl of a [^{3}H]leucine-labeled translation mix were denatured and diluted in RIPA buffer prior to immunoprecipitation. For sequencing of P2B3' precursors, 50 µl of a pT11 [^{35}S]methionine-labeled translation was fractionated by SDS PAGE, then electroblotted onto Immobilon PVDF (Millipore) membranes using a carbonate-based buffer system (15, 35).

For sequencing of P3' produced *in vitro*, 200 µl of a pT11-programmed [^{3}H]leucine-labeled or 300 µl of [^{3}H]valine-labeled translation mix was fractionated by SDS PAGE, then electroblotted onto Immobilon membranes. Immobilon membranes were autoradiographed to locate [^{35}S]methionine-labeled P3' markers and adjacent lanes containing [^{3}H]leucine or [^{3}H]valine-labeled P3' were excised. Samples labeled *in vivo* or *in vitro* were sequenced on an Applied Biosystems 477A gas phase sequenator without PTH derivatization using a customized ATZ-1 program. Radioactivity in the eluate of each cycle was determined in a liquid scintillation counter and ^{3}H counting efficiency was determined to be 40% using calibration standards.

Results

Production of fusion proteins and specificity of antisera. We have produced monospecific immune reagents reactive with NS2B and NS3 by constructing plasmids (pTNS2 and pTNS3) which express these moieties as fusion proteins with *trpE*. The locations in the dengue genome of the inserts used are shown in Fig. 1, which also illustrates schematically the posttranslational processing of the dengue polyprotein to give the mature structural and nonstructural proteins. When bacterial cultures harboring either pTNS3 or pTNS2 were induced and insoluble inclusion bodies were purified and analyzed by SDS PAGE, fusion proteins of 51 and 54 kDa, respectively, were found (Fig. 2A), indicating that the fusion proteins were stable when overexpressed. Fusion proteins were purified by preparative SDS PAGE and used to inject two rabbits each. High titer antiserum was usually obtained with three injections.

Denatured lysates of dengue infected cells were analyzed using antisera directed against pTNS2 and pTNS3 fusion proteins. Serum from a rabbit injected with the pTNS2 fusion protein immunoprecipitated a protein of approximately 14 kDa, identified as NS2B from its molecular size, immunoreactivity, and N-terminal sequence (see below), from infected but not from mock infected cell lysates (Fig. 2B), and the serum was therefore designated αNS2B serum. NS2B was immunoprecipitated from denatured but not from nondenatured lysates, implying that its immunoreactive epitopes are linear in nature or are masked in nondenatured lysates (data not shown). Since pTNS2 contains the last 19 amino acids of NS2A as well as the first 111 amino acids of NS2B (Table 1), antiserum raised against the encoded fusion protein could theoretically

Fig. 1. Genome organization and expression of dengue 2 virus. A schematic diagram aligning the *trpE* cDNA fusions to their corresponding regions of the dengue genome is shown. The upper third of the figure details the genome organization of the PR159 strain. The middle third of the figure depicts the processing events generating the structural and nonstructural proteins. Virion structural proteins are designated by dark crosshatched boxes and nonstructural proteins are designated by open boxes. Cleavages after dibasic residues are designated by downward facing arrowheads. Host cell signalase cleavages are designated by open ball and stick icons and a potential viral encoded signalase-like cleavage is designated by a shaded ball and stick icon. A late, cell membrane-associated cleavage generating the membrane structural protein is designated by a ball and arrowhead. The boundaries of the dengue inserts in the pTNS2 and pTNS3 cDNA fusions are mapped on the expanded view of NS2A, NS2B and NS3 shown in the lower third of the figure. Full details of constructs pTNS2 and pTNS3 are given in Table 1.

DENGUE 2 GENOME
(10,712nt)
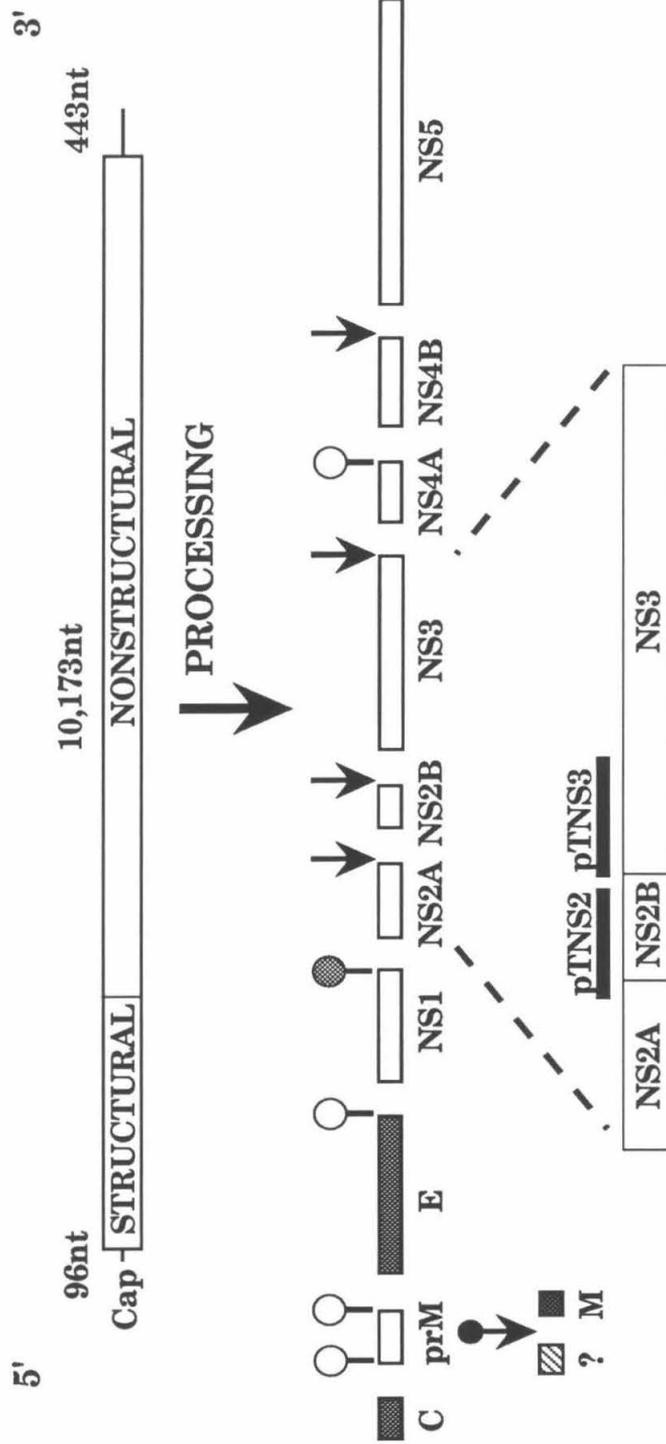
have reacted with NS2A sequences. However, no NS2A protein (predicted molecular weight, 24kDa) was detected when either denatured or nondenatured lysates were immunoprecipitated with αNS2B (Fig. 2B).

Serum from a rabbit injected with the pTNS3 fusion protein precipitated a polypeptide of 69 kDa, identified as NS3, from infected cell lysates (Fig. 2B). In addition three larger proteins were present in significant amounts in the immunoprecipitates, of sizes 85, 93, and 130 kDa. All four species were also immunoprecipitated by a polyvalent anti-dengue dengue mouse hyperimmune ascites fluid (Fig. 2B, lane HIS). The 93 kDa protein appears to be NS5, and the reason it should precipitate with the αNS3 antiserum is obscure. The size and reactivity of the 85 kDa protein are consistent with it being polyprotein NS34A. The identity of the 130 kDa species is unknown; it may be a polyprotein containing NS3.

Since pTNS3 encodes the last 8 amino acids of NS2B as well as the first 141 amino acids of NS3 (Table 1), antiserum raised against the encoded fusion protein could in theory also precipitate proteins containing NS2B sequences. αNS3 serum immunoprecipitated small amounts of NS2B from nondenatured lysates (data not shown) but not from denatured lysates (Fig. 2B). Thus with denatured lysates, the αNS2B and αNS3 sera described here are essentially monospecific, and denatured lysates were used for all the subsequent immunoprecipitations shown.

***In vitro* translation and processing kinetics.** Processing of flavivirus nonstructural proteins occurs very rapidly *in vivo*, and viralpolyproteins can only be detected in pulse chase experiments under

Fig. 2. *trpE*/dengue fusion proteins and characterization of antisera. Panel A: A Coomassie blue stained SDS PAGE gel with samples of purified inclusion bodies from induced cultures of XL1 alone and XL1 cultures containing plasmids pTNS3 and pTNS2, respectively. The positions of protein standards (Biorad) are shown to the left with the molecular weights in kDa indicated. Panel B: Immunoprecipitates of [^{35}S] methionine-labeled infected and uninfected BHK cells analyzed by PAGE. 1 X 10^5 infected (I and PI) or mock infected cells (U) were labeled for 12 hours beginning at 30 hours post-infection. Lysates were immunoprecipitated with a polyvalent mouse hyperimmune ascites (HIS) or monospecific αNS2B and αNS3 immune (I) and preimmune (PI) sera. Locations of bands containing the viral proteins E, NS1, NS2B, NS3, NS5 and the putative precursor NS34A are indicated. The molecular weights of ^{14}C-labeled protein markers (Amersham) are shown.

A.

B.

Table 1: *trpE*/dengue cDNA fusions.

Construct	Vector [a]	Genes	Insert [b]	A.A. [b]	Protein [c]
pTNS3	pATH3 Eco(K)/Sal	NS2B3	4497-4944 Asp(K)/Sal	122-129/1-141	NS3
pTNS2	pATH3 Bam(K)/Cla	NS2A2B	4077-4463 Stu/HpaII	200-218/1-111	NS2B

Footnotes

[a] Vector and insert DNAs were prepared as described in materials and methods. 5' overhangs that were blunted by Klenow treatment are designated with a (K).

[b] Nucleotides are numbered relative to the dengue genomic sequence and amino acid numbers of individual gene products are listed.

[c] Dengue specific protein recognized by each antiserum under denaturing conditions.

special labeling conditions (10). However, it seemed likely that these processing events could be followed upon translation *in vitro*. Our initial expression experiments used RNA transcripts of dengue 2 cDNA clones that were not engineered to have specific 5' leader sequences or initiation codons in an appropriate context. Only low levels of protein synthesis were obtained *in vitro* and variable amounts of specific (i.e., at the 5' methionine in the RNA) versus nonspecific (internal) initiation of translation was observed. Levels of expression were increased approximately 5-fold by inserting the dengue 5' nontranslated region upstream of the region to be translated. The efficiency of translation was increased an additional 2- to 3-fold whenthe reading frame of interest was fused to that of the dengue capsid protein, such that the mRNA translated had the authentic dengue 2 leader and the initiation codon in the normal dengue context (data not shown). The yellow fever 5' nontranslated region has also been shown to enhance expression of heterologous coding sequences by reticulocyte lysates (42).

The first construct tested for protease activity was pT10. RNA transcribed from this construct with T7 RNA polymerase is translated into a polyprotein (P2A2B3) containing the first 37 amino acids of the capsid protein, the C-terminal 118 amino acids of NS2A, all of NS2B, and most of NS3 (Fig. 3 and Table 2). Protein patterns resulting from different times of translation of this RNA in reticulocyte lysates are shown in Fig. 4A. A protein with the molecular weight predicted for the slightly truncated NS3 (P3) was produced (Fig. 4A), showing that proteolytic processing had occurred in this *in vitro* system. Proteins with molecular weights

Fig. 3. Schematic diagram of dengue 2 protein and nucleic acid sequences expressed in the translation vector pT10. T7 RNA polymerase transcribes a chimeric RNA which contains nucleotides 1-213 fused to nucleotides 3778-6351 of the dengue genome. An initiator methionine codon at position 96 is designated by a solid diamond. Upon translation the encoded polyprotein contains amino acids 1-37 of the dengue capsid proten, 101-218 of NS2A, all of NS2B and the first 610 amino acids of NS3. The predicted protease domain of NS3 is indicated by the solid black region.

pT10
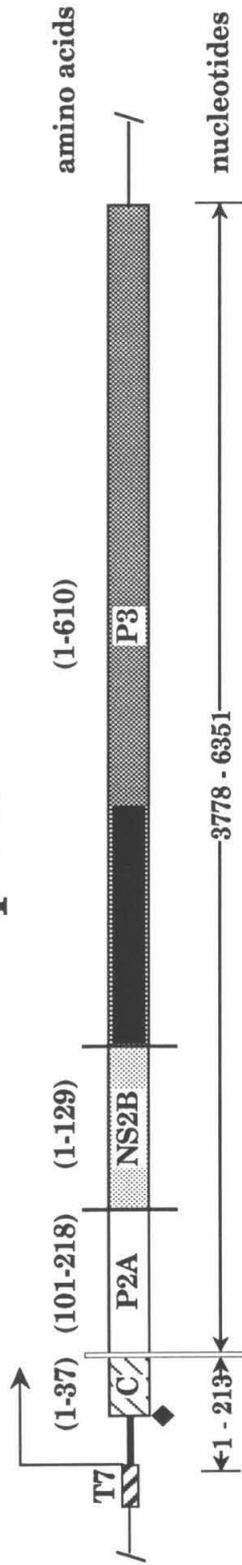
predicted for the full-length (unprocessed) translate P2A2B3 and for the processing intermediate P2B3 were also observed (Fig. 4A).

In order to study the kinetics of polyprotein processing, bands P2A2B3, P2B3 and P3 in Fig. 4A were excised from the dried gel, solubilized, and assayed for radioactivity (Fig. 4B). The resulting kinetics of accumulation and processing exhibited two phases. All three species accumulated with similar half-times during the first 30 minutes. After 30 minutes of translation (post synthesis phase) there was no further incorporation of label into newly synthesized protein. The label in the full-length P2A2B3 band remained constant between 30 and 90 minutes, indicating that these molecules were stable and were not processed, whereas P2B3 decayed with a half-life of 42 minutes and P3 accumulated at the same rate suggesting a precursor product relationship between these two species. The fact that P2A2B3 appeared to be processed rapidly in the first 30 min but to be stable thereafter suggests that folding of this protein does not occur optimally during translation in reticulocyte lysates, and that those molecules that fold correctly during the initial synthesis phase are processed rapidly while misfolded molecules are processed slowly or not at all. The fraction of P2A2B3 that was processed varied from experiment to experiment, being sensitive to the lot of reticulocyte lysate used and to other unknown variables.

Since the pT10 construct contains two cleavage sites, different processing intermediates would be predicted depending upon the temporal order in which the cleavages occur. If the cleavage between P2A and NS2B preceded cleavage between NS2B and P3, intermediate P2B3 would be generated. Conversely, if cleavage between NS2B and P3 preceded

cleavage between P2A and NS2B, P2A2B rather than P2B3 would be produced. The potential precursor P2A2B was not observed by SDS PAGE (Fig. 4A) nor by immunoprecipitation with αNS2B serum (data not shown), but P2B3 was readily detected as described above, suggesting that cleavage at the 2A/2B junction occurred first. The amounts and processing kinetics of P2B3 were also consistent with the view that it is the major intermediate in the processing pathway.

Cleavage of the dengue polyprotein in *cis*. Intermolecular (*trans*) cleavages exhibit second-order kinetics and are concentration-dependent, whereas intramolecular (*cis*) cleavages are first-order and concentration-independent. To determine whether the cleavages at the dengue 2A/2B and 2B/3 sites were concentration-dependent, we examined the protein products obtained upon translating differing amounts of input RNA,which effectively changes the concentration of the proteinase (Fig. 4C). At the highest concentrations of RNA tested the components of the translation system were limiting (28) and processing appeared to proceed more rapidly. At lower concentrations of RNA, total incorporation appeared to increase between 25 and 90 minutes of translation and proportionately more label was present in a 46 kDa reticulocyte-specific band. The ratio of P2A2B3 to P3 after 90 minutes of translation at 30°C appeared to be independent of the concentration of input RNA (Fig. 4C), however, this suggests that cleavage occurs in *cis*, although extremely efficient *trans* cleavage by the dengue proteinase cannot be ruled out.

We have also used a second approach to examine whether the dengue polyprotein can be cleaved in *trans*. Polyproteins produced by

54

Fig. 4. Processing of proteins during *in vitro* translation of construct pT10. Panel A: aliquots of a pT10-programmed translation were removed at the designated times after initiation of translation and analyzed by SDS PAGE. Precursors P2A2B3 and P2B3 as well as mature P3 are indicated. Panel B: bands containing P2A2B3, P2B3, and P3 were excised from gels like that shown in A and quantitated as described in Materials and Methods. Panel C: aliquots of RNA transcribed from pT10 were serially diluted and translated *in vitro*. Aliquots of the translation mixtures were analyzed by PAGE. In order to equalize for [^{35}S] methionine incorporation, varying amounts of lysate were loaded in each lane. Lanes 1-4 contain 1 µl of lysate, lanes 5, 6, 8 and 10 contain 2 µl of lysate and lanes 7, 9 and 11 contain 4 µl of lysate. Mass of RNA per ml and the length of the translation incubation are indicated. Locations of bands containing nonstructural protein precursors P2A2B3 and P2B3 as well as mature P3 are indicated.
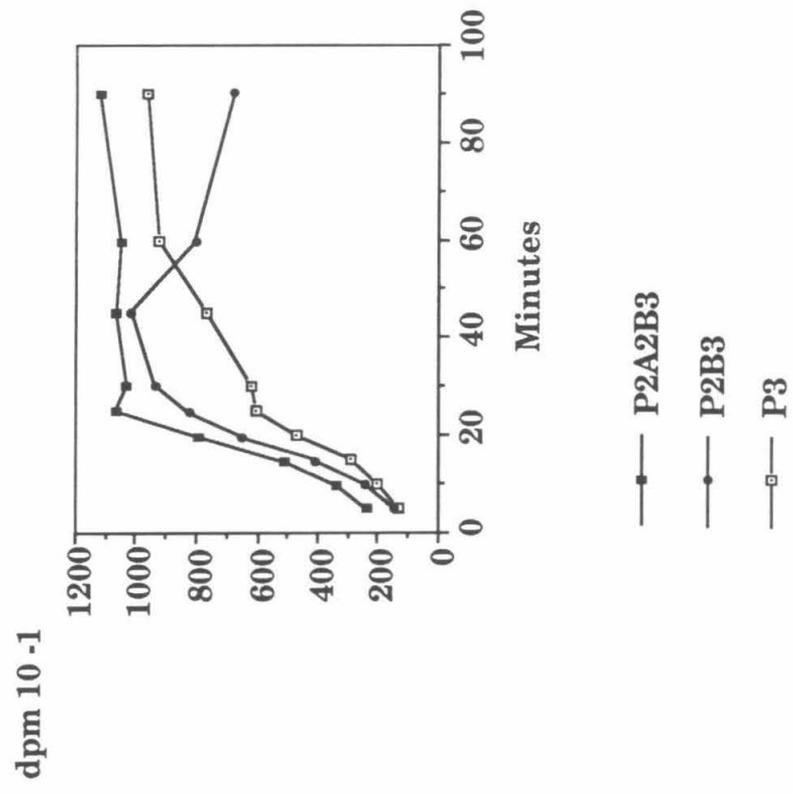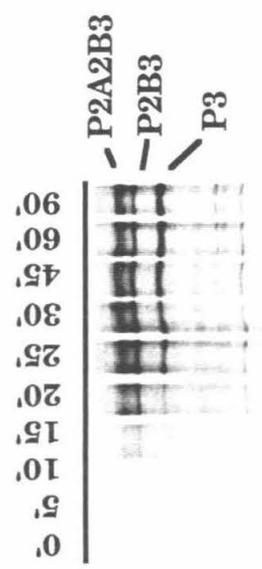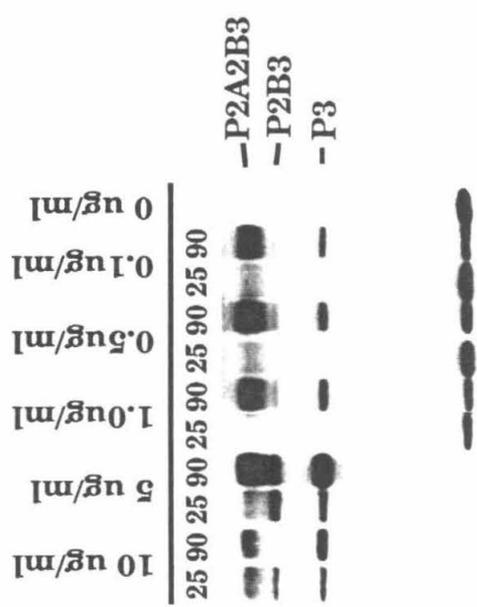
pT13- and pT14-programmed translations contain an inactive protease but the potential cleavage sites are intact (see below). These polyproteins were used as substrates for the active proteinase produced by a pT10-programmed translation, and no detectable cleavage of these polyproteins occurred (data not shown). The lack of detectable *trans* cleavage in these experiments is consistent with the results of the kinetic and dilution sensitivity experiments which suggested that cleavage occurred at the 2A/2B and 2B/3 junctions only in *cis*.

Deletion mapping of the dengue protease. Based upon molecular modeling studies, both Bazan and Fletterich (3) and Gorbalenya et al. (20) have predicted the precise boundaries of the flavivirus protease domain. This predicted domain is illustrated schematically in Fig. 5 in relation to the dengue polyprotein. The entire domain is hypothesized to span the first 180 amino acids of NS3 (black shading) and to contain 4 subregions or boxes of homology with serine proteases, the first three of which contain the three elements of the catalytic triad and the fourth is involved in substrate binding. To test these predictions, we constructed a series of deletion constructs that together span the entire length of NS3 and that delete large regions of P2A and NS2B (Fig. 5 and Table 2); translation mixes were programmed with RNA transcribed from these constructs, and the resulting protein products immunoprecipitated with αNS2B and αNS3 sera and analyzed by SDS PAGE (Fig. 6).

Constructs pT10, pT11, and pT12 contain all of the proposed protease domain (see Fig. 5), and from these constructs polypeptides with the molecular weights and immunoreactivities of mature NS2B and of the

Fig. 5. Schematic map of deletion constructs used to map protease activity. The relative positions of the four homology boxes proposed by Bazan and Fletterick (3), and Gorbalenya et. al. (20) are shown. The proposed catalytic histidine, aspartate and serine residues are listed within boxes 1, 2, and 3 and indicated by vertical lines on the schematic. SBP refers to the proposed substrate binding pocket. Gene products, P2A and NS2B are designated by open and lightly shaded boxes respectively. The putative protease and helicase domains of NS3 are designated by solid and deeply shaded boxes respectively. The restriction sites used to generate deletion constructs are indicated as follows; A for *Ase*I, P for *Pvu*II, Sn for *Sna*BI, K for *Kpn*I (*Asp*718), D for *Dra*III, S for *Sal*I, As for *Asu*II, X for *Xho*I and E for *Eco*RI. The structures of deletion clones are mapped below. Cleavages at the 2A/2B or 2B/3 junctions which occur are indicated by arrows. A detailed description of each construct is given in Table 2.
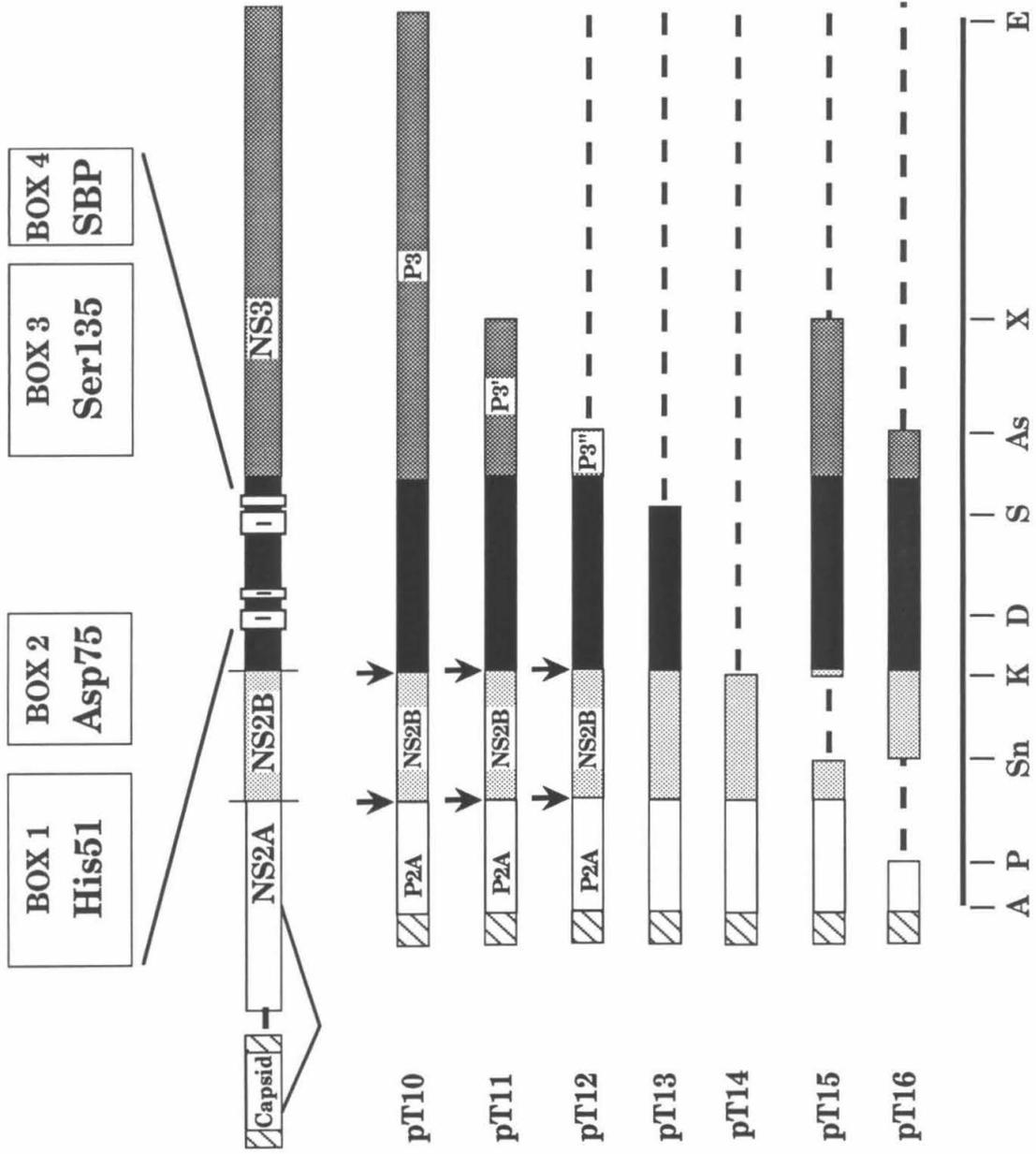
Table 2: Deletion constructs.

Construct	Position of deleted amino acids [a]	
pT10	NS3	611-618
pT11	NS3	303-618
pT12	NS3	184-618
pT13	NS3	141-618
pT14	NS2B/NS3	122-129/1-618
pT15	NS2B/NS3	43-121, 303-618
pT16	NS2A/NS2B/NS3	146-218/1-42/184-618

Footnotes

[a] Genes affected by deletion. The amino acids deleted in individual gene products are listed. When more than one gene product contains a deletion, amino acid numbers are separated by a "/" . When separate deletions are present in the same gene product, the amino acid numbers are separated by a comma.

truncated forms of NS3 (P3, P3' and P3") were detected (Fig. 6A). In contrast pT13 programmed translations did not produce any (truncated) NS3 or NS2B (Fig. 6B), indicating that this deletion abolished activity, presumably by invading the protease domain (see Fig. 5). Thus the protease domain encompasses a maximum of 183 amino acids at the N-terminus of NS3. Polypeptides consistent in molecular weight and immune reactivity with a P2B3-processing intermediate could also be detected in pT10-, pT11- and pT12-programmed translations, but not pT13-programmed translations (Fig. 6).

Translation of RNA transcribed from pT14 yielded a polypeptide which was consistent in molecular weight and immunoreactivity with that predicted for P2A2B'. Processing did not occur at the 2A/2B site. The location of the band and the amount expressed from pT14 were such that if this precursor had been present upon translation of RNA from constructs such as pT10, it would have been readily detected. This supports the earlier conclusions that P2A2B was never generated during processing *in vitro*.

To examine the amino terminal boundary of the protease domain, that is whether sequences in P2A or NS2B are required for proteolytic activity, we analyzed the *in vitro* translation patterns produced from two deletion constructs, pT15 and pT16, which together delete almost all of the P2A and NS2B sequences. When pT15-programmed translations were fractionated using αNS2B only low levels of immunoreactive species were observed upon overexposure of autoradiographs (Fig. 6B). The molecular weights of these species were not consistent with correct processing at the 2A/2B or 2B/3 cleavage sites. The low level of immunoreactivity suggests

Fig. 6. Immunoprecipitations of *in vitro* translations programmed with RNA transcribed from NS3 deletion constructs.

Panel A: aliquots of reticulocyte lysates were denatured and immunoprecipitated with αNS2B (2) or αNS3 (3) immune sera, immune complexes were dissolved in SDS-containing loading buffer and analyzed by PAGE. The construct from which RNA was transcribed is indicated above each lane. Protein species which were identified as P2A2B3", P2B3", P3, P3', P3" and NS2B by immunoreactivity and molecular weight are indicated. The molecular weights of ^{14}C-labeled protein standards (Amersham) are indicated.

Panel B: immunoprecipitations of *in vitro* translations from proteolytically inactive templates. Unprocessed precursors P2A2B3'" and P2A2B' are designated as are the positions of molecular weight markers.
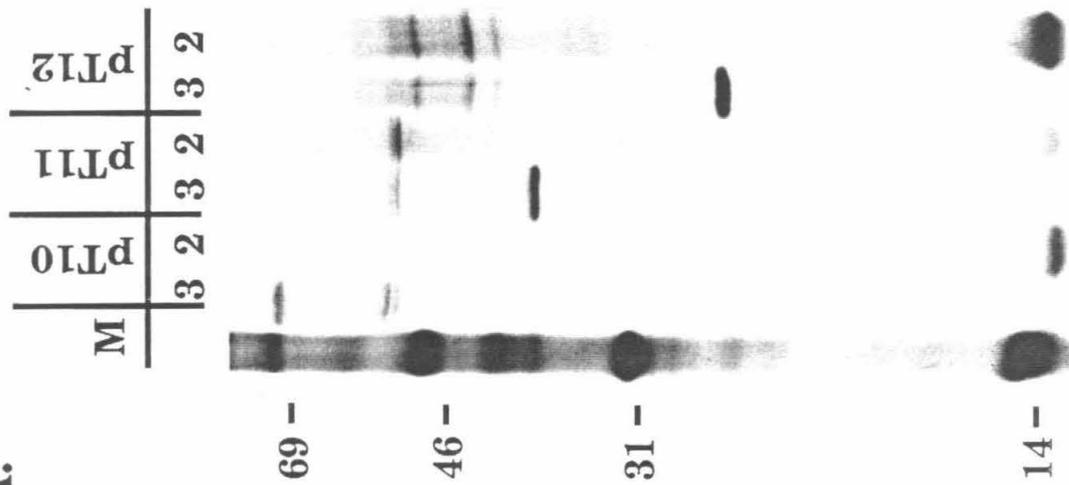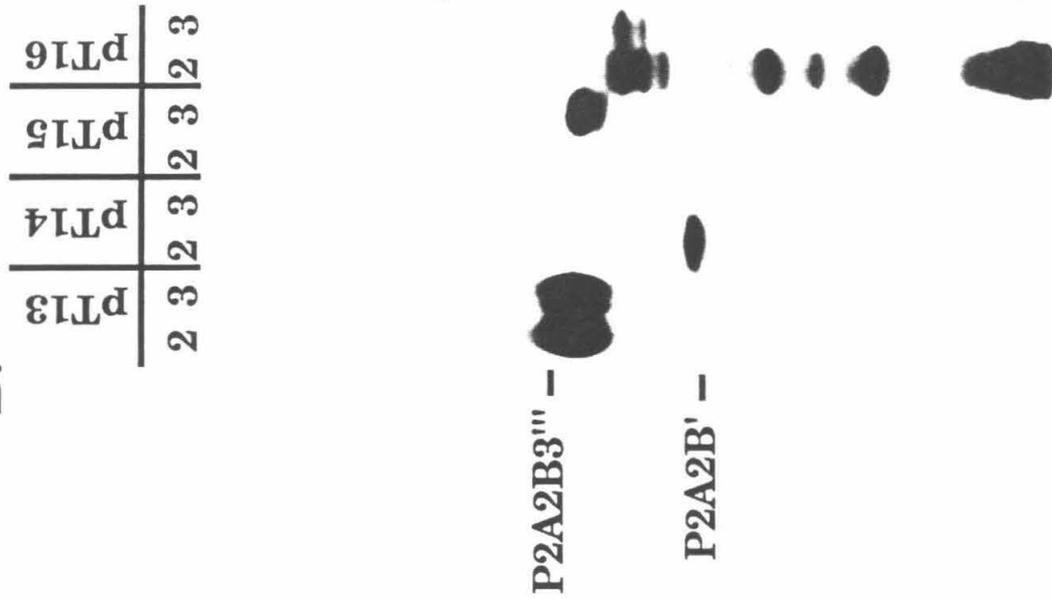
that the major antigenic epitopes recognized by αNS2B have been deleted in this construct. Upon precipitation with αNS3, a small amount of aberrantly processed P3' was detected. When pT16-programmed translations were immunoprecipitated and analyzed by SDS PAGE, both αNS2B and αNS3 immunoreactive species of aberrant molecular weight could be detected. Since the 2A/2B cleavage site had been removed by this deletion only cleavage at the 2B/3 boundary could be examined. Thus it is unclear whether sequences in P2A or NS2B are required for proteolytic activity or whether the deletions induce misfolding of the molecule such that the correct cleavage sites are not recognized.

Sequencing of viral proteins. In order to confirm that the cleavage events observed *in vitro* occur at the same sites as utilized *in vivo*, and to confirm the identities of the proteins produced *in vitro*, we compared the amino terminal amino acid sequences of NS2B and NS3 produced *in vitro* with those of proteins isolated from infected cell lysates, and obtained the N-terminal sequence of P2B3'. NS2B from dengue 2-infected cells labeled with [^{35}S]methionine or [^{3}H]leucine was isolated by preparative immunoaffinity chromatography. Approximately 85% of the label in the purified NS2B used for sequencing consisted of mature NS2B, as determined by electrophoresis. The results clearly showed that serine 1345 of the dengue polyprotein is the amino terminal residue of NS2B produced *in vivo* (Fig. 7A) This assignment is in agreement with data for other sequenced NS2B species (9, 47, 53), based upon sequence homology.

In order to sequence NS2B produced *in vitro*, pT11-programmed translations were immunoprecipitated with αNS2B. Approximately 85%

of the NS2B immunoprecipitated was fully cleaved, and approximately 10% of the label was in P2B3', which is coterminal with mature NS2B and contributes to the observed NS2B signal (see below). The amino terminus of *in vitro* produced NS2B was indistinguishable from that found *in vivo*. (Fig. 7B).

Sequencing of P3' labeled *in vitro* in pT11-programmed translations identified the amino terminal residue of P3' as Ala-1476 of the dengue 2 polyprotein (Fig. 7C). Peaks of [^{3}H]leucine at cycles 4 and 18 and of [^{3}H]valine at cycles 3, 7, and 13 make this assignment unambiguous. This result agrees perfectly with the sequence of NS3 isolated from dengue 2 infected cells (4).

The P3' results are also of interest because the cleavage which produces the amino terminus of dengue 2 NS3 is unique in that it appears to occur after a single basic amino acid; the cleavages to produce NS3 in all other sequenced flaviviruses occur after dibasic residues. The fact that the amino terminus of P3' produced *in vitro* is the same as that of NS3 isolated from infected cells suggests that the dengue 2B/3 cleavage site is in fact different, and that the observed terminus of NS3 does not arise from cleavage at an upstream site characterized by dibasic residues followed by amino-terminal nibbling (4).

In order to further define the specificity of cleavage events and definitively identify processing intermediates, the putative P2B3' was isolated from pT11-programmed translations by preparative SDS-PAGE and sequenced. Peaks of ^{35}S were observed at positions 9 and 13, aligning this species perfectly with the amino terminus of NS2B and confirming the

Fig. 7. Amino terminal sequencing of viral proteins. Data in each panel are plotted as radioactivity per Edman degradation cycle versus cycle number. Panel A: *in vivo* labeled NS2B; panel B: *in vitro* labeled NS2B; panel C: *in vitro* labeled P3'; panel D: precursor P2B3'. The amino terminal sequence of NS2B, NS3 and NS2B3 deduced from the nucleotide sequence of the RNA genome is indicated below each sequence panel with the leucine, valine and methionine residues identified by Edman degradation highlighted in bold, capital letters. The burst of [^{35}S] dpm observed in the first sequencing cycle of immunoaffinity purified NS2B (both *in vivo* and *in vitro*) is not consistent with the deduced sequence of NS2B and is presumably contributed by contaminants which coprecipitate with NS2B. The single letter amino acid code used is as follows: A, alanine; C, cysteine; D, aspartate; E, glutamate; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan and Y, tyrosine.
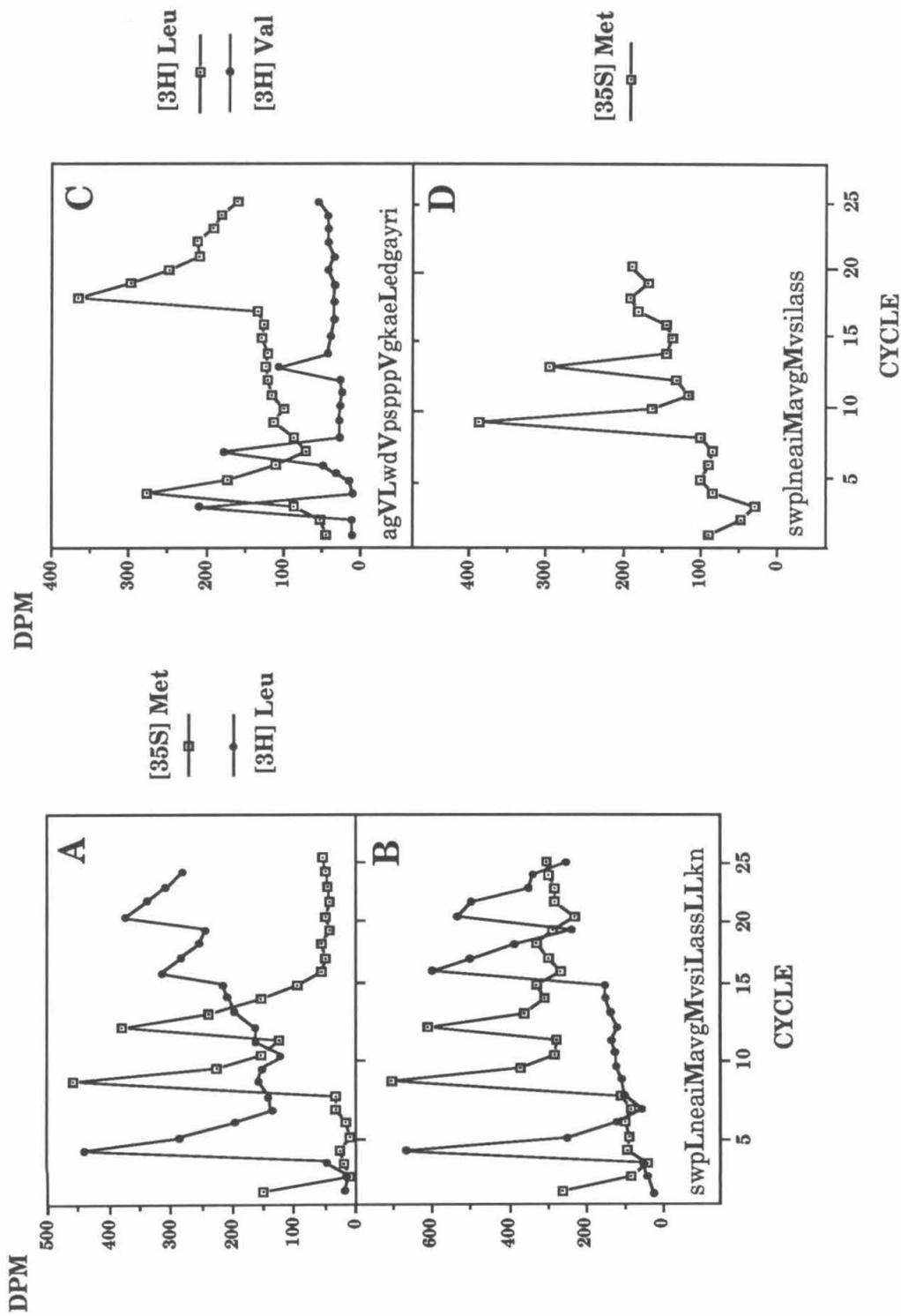
previous identification based on gel mobility and immunoreactivity (Fig. 7D).

Discussion

Cleavage of the dengue polyprotein *in vitro*. The kinetic experiments indicated that *in vitro* cleavage at the 2A/2B junction precedes that at the 2B/3 junction. These junctions were inactive as substrates in a *trans* cleavage assay and cleavage was insensitive to dilution, implying that cleavage at these two junctions occurs by an intramolecular mechanism. It is formally possible that NS3 is extremely active in *trans* such that sensitivity to dilution could not be demonstrated and that the negative results of the *trans* processing experiments resulted from the perturbed tertiary structure of substrates, although we have used similar methods to demonstrate *trans* cleavage of the Sindbis virus nonstructural polyprotein (28). Differences in the rate of translation and in the association of membranes with processing intermediates could also influence the observed rate of cleavage *in vivo*.

Cleavage at the 2A/2B and 2B/3 junctions *in vitro* proceeds with the same specificity as found *in vivo*. Translation products are cleaved at the correct sites of cleavage and nonspecific cleavage is rarely observed. While the kinetics of cleavage *in vitro* are much slower than those observed *in vivo*, it is highly probable that the fidelity and order of *in vitro* cleavage are a true reflection of *in vivo* processing. The use of improved expression systems and *in vitro* mutagenesis of cleavage sites will be useful in establishing whether the order of cleavage observed is obligatory or whether it simply represents the kinetically favored order.

Testing the flavivirus protease model. Positive identification of NS3 as a flavivirus proteinase will stimulate research on the nature of the catalytic and substrate recognition residues involved in proteolysis. The specific predictions as to which histidine, aspartate and serine residues form the catalytic triad are testable through site specific mutagenesis, and mutants which are catalytically active *in vitro* can be tested for biological activity *in vivo* through the use of infectious cDNA clones (40). NS3 shares the greatest amount of sequence similarity with the Sindbis virus capsid protein protease domain and site-directed mutagenesis of the NS3 protease domain will be greatly assisted by molecular modeling studies once the coordinates of the Sindbis capsid protein become available (5).

Four specific predictions of the protease model were addressed by experiments presented here (3, 20). Our experimental results are consistent with the predictions that the protease domain consisted of approximately 180 amino acids at the amino terminus of NS3 and could retain function when severed from the remainder of NS3, which forms a putative helicase domain. However our results are not consistent with the prediction that an internal conserved cleavage site in NS3 is utilized to separate the helicase and protease domains, or the prediction that cleavage of the flavivirus polyprotein by NS3 is an ATP-dependent process. The antiserum used in our experiments was directed specifically against the protease domain, and therefore any internal cleavage of NS3 *in vitro* or *in vivo* should have been detected. Secondly, proteolysis occurred even in those cases where the helicase domain had been deleted, making it unlikely that hydrolysis of ATP is absolutely required for proteolysis,

although the possibility that it has an effect *in vivo* has not been eliminated.

The presence of protease and helicase domains within a single viral protein is a common structural motif found in alpha-, pesti-, poty-, and coronavirus proteins as well as flaviviruses (21). The primary function of a viral proteinase is to posttranslationally regulate the production of individual gene products from a polyprotein precursor. The viral proteinase can also produce processing intermediates which may themselves be functional components of the viral life cycle. The choice between emphasizing proteinase or helicase function would depend on the specific needs of the virus at that point in the life cycle. It appears likely that the main function of mature NS3 is as a helicase, since the initial cleavages releasing NS2B and NS3 most likely occur in *cis* and mature NS3 does not appear to work efficiently in *trans*.

Implications for dengue polyprotein processing. A number of viral proteinases function both in *cis* and in *trans* to cleave polyproteins *in vivo* (31). Like other positive strand RNA viral proteinases, NS3 may also function both in *cis* and in *trans*. The amino terminus of NS4B is believed to be generated by a signalase cleavage event which presumably occurs co-translationally upon insertion of the hydrophobic tail of NS4A into the lumen of the endoplasmic reticulum (9, 47). Once this event occurs the dengue polyprotein backbone is severed and by definition any processing events that occur downstream of this scission must occur in *trans*, and if NS3 is the proteinase responsible for cleavage at the 4B/5 boundary then it must cleave in *trans*.

Processing events consistent with scission at the 3/4A junction are observable *in vitro* (F. Preugschat unpublished data) and putative NS34A intermediates can be detected *in vivo* (Fig. 2B). This implies that the kinetics of 3/4A cleavage are slower than the kinetics of 2A/2B and 2B/3 cleavage. It is possible that NS34A is a precursor that is restricted to the plane of the endoplasmic reticulum by a carboxy terminal membrane spanning segment, and that this is the form of proteinase that is responsible for *trans* cleavage at the 4B/5 junction. By restricting the proteinase to the plane of the endoplasmic reticulum, the concentration-dependence of the *trans* cleavage step would be reduced and result in rapid kinetics of cleavage in infected cells.

The flavivirus capsid protein undergoes a complex two-step maturation process in infected cells. As the carboxy terminus of the capsid protein is translated it acts as a membrane insertion sequence for a cotranslational signalase mediated cleavage that generates the amino terminus of the prM protein. Nascent, intracellular capsid protein possesses a C-terminal membrane-spanning segment that is cleaved prior to assembly of the virion (36). This cleavage event occurs on the carboxy terminal side of several basic residues and could be potentially mediated in *trans* by NS3. The initial signalase mediated cleavage can be observed using an *in vitro* translation system but the secondary maturation cleavage that removes the membrane insertion sequence has not been observed *in vitro* (36, 43). By using an *in vitro* expression system designed to analyze cleavages at the 3/4A, 4B/5 and C/prM junctions, it should possible to reconstruct the flaviviral protein processing pathway using the experimental approaches described in this paper.

Acknowledgements

References Cited

1.   **Bancroft, W. H., R. M. Scott, K. H. Eckels, C. H. Hoke, T. E. Simms, K. D. T. Jesrani, P. L. Summers, D. R. Dubois, D. Tsoulos, and P. K. Russell.** 1984. Dengue virus type 2 vaccine: Reactogenicity and immunogenicity in soldiers. J. Inf. Dis. **149:**1005-1010.

2.   **Bancroft, W. H., F. H. Top, K. H. Eckels, J. H. Anderson, J. M. McCown, and P. K. Russell.** 1981. Dengue-2 vaccine: Virological, immunological, and clinical responses of six yellow fever-immune recipients. Infect. Immun. **31:**698-703.

3.   **Bazan, J. F., and R. J. Fletterick.** 1989. Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. Virology **171:**637-639.

4.   **Biedrzycka, A., M. R. Cauchi, A. Bartholomeusz, J. J. Gorman, and P. J. Wright.** 1987. Characterization of protease cleavage sites involved in the formation of envelope glycoprotein and three nonstructural proteins of dengue virus type 2, New Guinea C strain. J. Gen. Virol. **68:**1317-1326.

5.   **Boege, U., M. Cygler, G. Wengler, P. Dunras, and J. Tsao.** 1989. Sindbis virus core protein crystals. J. Mol. Biol. **208:**79-82.

6. **Calisher, C. H., N. Karabatsos, J. M. Dalrymple, R. E. Shope, J. S. Porterfield, E. G. Westaway, and W. E. Brandt.** 1989. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. J. Gen. Virol. **70:**37-43.

7. **Calisher, C. H., R. E. Shope, W. Brandt Casals, J., N. Karabatsos, F. A. Murphy, R. B. Tesh, and M. E. Wiebe.** 1980. Proposed antigenic classification of registered arboviruses. I. Togaviridae, *Alphavirus.* Intervirology **14:**229-232.

8. **Chamberlin, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. Anal. Biochem. **98**:132-135.

9. **Chambers, T. J., D. W. McCourt, and C. M. Rice.** 1989. Yellow fever virus proteins NS2a, NS2b and NS4b: Identification and partial N-terminal amino acid sequence analysis. Virology **169**:100-109.

10. **Cleaves, G. R.** 1985. Identification of dengue type 2 virus-specific high molecular weight proteins in virus-infected BHK cells. J. Gen. Virol. **66**:2767-2771.

11. **Cleaves, G. R., and D. T. Dubin.** 1979. Methylation status of intracellular dengue type 2 40S RNA. Virology **96**:158-165.

12. **Coia, G., M. D. Parker, G. Speight, M. E. Byrne, and E. G. Westaway.** 1988. Nucleotide and complete amino acid sequences of Kunjin virus: Definitive gene order and characteristics of the virus-specified proteins. J. Gen. Virol. **69**:1-21.

13. **Condreay, L. D., R. H. Adams, J. Edwards, and D. T. Brown.** 1988. Effect of actinomycin D and cycloheximide on replication of Sindbis virus in *Aedes albopictus* (mosquito) cells. J. Virol. **62**:2629-2635.

14. **Deubel, V., R. M. Kinney, and D. W. Trent.** 1988. Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins of dengue type 2 virus, Jamaica genotype: Comparative analysis of the full-length genome. Virology **165**:234-244.

15. **Dunn, S. D.** 1986. Effects of the modification of transfer buffer composition and the renaturation of proteins in gels on the recognition of proteins on western blots by monoclonal antibodies. Anal. Biochem. **157**:144-153.

16. **Eckels, K. H., V. R. Harrison, P. L. Summers, and P. K. Russell.** 1980. Dengue-2 vaccine: Preparation from a small-plaque virus clone. Infect. Immun. **27:**175-180.

17. **Ehrenkranz, N., A. Ventura, R. Cuadrado, W. Pond, and J. Porter.** 1971. Pandemic dengue in Caribbean countries and the southern United States--Past, present, and potential problems. New Engl. J. Med. **285:**1460-1469.

18. **Falgout, B., R. Chanock, and C. J. Lai.** 1989. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2A. J. Virol. **63:**1852-1860.

19. **Gorbalenya, A. E., A. P. Donchenko, V. M. Blinov, and E. V. Koonin.** 1989. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases: A distinct protein superfamily with a common structural fold. FEBS Letters **243:**103-114.

20. **Gorbalenya, A. E., A. P. Donchenko, E. V. Koonin, and V. M. Blinov.** 1989. N-Terminal domains of putative helicases of Flavi and Pestiviruses may be serine proteases. Nucl. Acids. Res. **17:**3889-3897.

21. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1988. A conserved NTP-binding motif in putative helicases. Nature **333:**22.

22. **Gruenberg, A., W. S. Woo, A. Biedrzycka, and P. J. Wright.** 1988. Partial nucleotide sequence and deduced amino acid sequence of the structural proteins of dengue virus type 2, New Guinea C and PUO-218 strains. J. Gen. Virol. **69:**1391-1398.

23.   **Gubler, D. J.** 1989. *Aedes aegypti* and *Aedes aegypti*-borne disease control in the 1990s: Top down or bottom up. Am. J. Trop. Med. Hyg. **40:**571-578.

24.   **Hahn, C. S., E. G. Strauss, and J. H. Strauss.** 1985. Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid autoprotease. Proc. Natl. Acad. Sci. **82:**4648-4652.

25.   **Hahn, C. S., and J. H. Strauss.** 1990. Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. J. Virol. in press.

26.   **Hahn, Y. S., R. Galler, T. Hunkapiller, J. Dalrymple, J. H. Strauss, and E. G. Strauss.** 1988. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. Virology **162:**167-180.

27.   **Hardy, W. R., and J. H. Strauss.** 1988. Processing of the nonstructural polyproteins of Sindbis virus: Study of the kinetics *in vivo* using monospecific antibodies. J. Virol. **62:**998-1007.

28.   **Hardy, W. R., and J. H. Strauss.** 1989. Processing the nonstructural proteins of Sindbis virus: Nonstructural proteinase is in the C-terminal half of nsP2 and functions both in *cis* and *trans*. J. Virol. **63:**4653-4664.

29.   **Hewick, R. M., M. W. Hunkapiller, L. E. Hood, and W. J. Dreyer.** 1981. A gas-liquid-solid phase peptide and protein sequenator. J.Biol.Chem **256:**7990-7997.

30.   **Kelley, P. M., and M. J. Schlesinger.** 1982. Antibodies to two major chicken heat shock proteins cross-react with similar proteins in widely divergent species. Mol. Cell. Biol. **2:**267-274.

31. **Kraeusslich, H. G., and E. Wimmer.** 1988. Viral proteinases. Ann. Rev. Biochem. **57:**701-754.

32. **Laemmli, U. K.** 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature **227:**680-685.

33. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

34. **Markoff, L.** 1989. *In vitro* processing of dengue virus structural proteins: Cleavage of the pre-membrane protein. J. Virol. **63:**3345-3352.

35. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262:**10035-10038.

36. **Nowak, T., P. M. Farber, G. Wengler, and G. Wengler.** 1989. Analyses of the terminal sequences of West Nile virus structural proteins and of the *in vitro* translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. Virology **169:**365-376.

37. **Ozden, S., and B. Poirier.** 1985. Dengue virus-induced polypeptide synthesis. Arch. of. Virol. **85:**129-137.

38. **Rice, C., M, E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss.** 1985. Nucleotide sequence of yellow fever virus: Implications for flavivirus gene expression and evolution. Science **229:**726-733.

39. **Rice, C. M., R. Aebersold, D. B. Teplow, J. Pata, J. R. Bell, A. V. Vorndam, D. W. Trent, M. W. Brandriss, J. J. Schlesinger, and J. H. Strauss.** 1986. Partial N-terminal amino acid sequences of three nonstructural proteins of two flaviviruses. Virology **151:**1-9.

40. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers.** 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by *in vitro* ligation. New Biol. **1**:285-296.

41. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: Mapping of lethal mutations, rescue of a temperature-sensitive marker, and *in vitro* mutagenesis to generate defined mutants. J. Virol. **61**:3809-3819.

42. **Ruiz-Linares, A., M. Bouloy, M. Girard, and A. Cahour.** 1989. Modulations of the *in vitro* translational efficiencies of yellow fever virus mRNAs: Interactions between coding and noncoding regions. Nucl. Acids. Res. **17**:2463-2476.

43. **Ruiz-Linares, A., A. Cahour, P. Despres, M. Girard, and M. Bouloy.** 1989. Processing of the yellow fever virus polyprotein: Role of cellular proteases in the maturation of the structural proteins. J. Virol. **63**:4199-4209.

44. **Schatzmayer, H. G., R. M. R. Nogueira, and A. P. A. Travassos da Rosa.** 1986. An outbreak of dengue virus at Rio de Janeiro (1986). Mem. Inst. Oswaldo Cruz **81**:245-246.

45. **Scott, R. M., K. H. Eckels, W. H. Bancroft, P. L. Summers, J. M. McCown, J. H. Anderson, and P. K. Russell.** 1983. Dengue 2 vaccine: Dose response in volunteers in relation to yellow fever status. J. Inf. Dis. **148**:1055-1060.

46. **Smith, G. W., and P. W. Wright.** 1985. Synthesis of proteins and glycoproteins in dengue type 2 virus-infected Vero and *Aedes* albopictus cells. J. Gen. Virol. **66**:559-571.

47. **Speight, G., G. Coia, M. D. Parker, and E. G. Westaway.** 1988. Gene mapping and positive identification of the nonstructural proteins NS2A, NS2B, NS3, NS4B and NS5 of the flavivirus Kunjin and their cleavage sites. J. Gen. Virol. **69:**23-34.

48. **Spindler, K. R., D. S. E. Rosser, and A. J. Berk.** 1984. Analysis of adenovirus transforming proteins from early regions 1A and 1B with antisera to inducible fusion agents produced in *Escherichia coli.* J. Virol. **49:**132-141.

49. **Svitkin, Y. V., V. N. Lyapustin, V. A. Lashkevich, and V. I. Agol.** 1984. Differences between translation products of Tick-borne encephalitis virus RNA in cell-free systems from Krebs-2 cells and rabbit reticulocytes: Involvement of membranes in the processing of nascent precursors of flavivirus structural proteins. Virology. **135:**536-541.

50. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucl. Acids. Res. **14:**4683-4690.

51. **Wengler, G., M. Beato, and G. Wengler.** 1979. *In vitro* translation of 42S virus-specific RNA from cells infected with the flavivirus West Nile virus. Virology. **96:**516-529.

52. **Wengler, G., G. Wengler, and J. Gross.** 1978. Studies on virus-specific nucleic acids synthesized in vertebrate and mosquito cells infected with flaviviruses. Virology. **89:**423-437.

53. **Wright, P. S., M. R. Cauchi, and M. L. Ng.** 1989. Definition of the carboxy termini of the three glycoproteins specified by dengue virus type 2. Virology. **171:**61-67.

Chapter 2

Studying Enzyme-Substrate Interaction Using Chimeric Proteinases: Identification of an Intragenic Locus Important for Substrate Recognition.

This chapter was prepared for submission.

Studying Enzyme-Substrate Interaction Using Chimeric Proteinases: Identification of an Intragenic Locus Important for Substrate Recognition.

Frank Preugschat and James H. Strauss*.

Division of Biology 156-29.
California Institute of Technology
Pasadena, CA. 91125.

Running Title: Chimeric Flavivirus Nonstructural Proteinases.

*Author to whom correspondence should be addressed.

Abstract

Flaviviruses encode their structural and nonstructural proteins as a single long polyprotein which is cotranslationally and posttranslationally processed by both cellular enzymes and a viral-encoded protease. The viral protease is located within the N-terminal 184 amino acids of the NS3 protein and efficiently cleaves the sites separating protein NS2A from NS2B and NS2B from NS3 itself, acting primarily autocatalytically. To study the nature and specificity of the flaviviral protease, we have constructed recombinant polyproteins in which part or all of the protease domains have been exchanged between the yellow fever and dengue genomes. Expression *in vitro* revealed that the dengue protease domain efficiently processes the yellow fever polyprotein between NS2A and NS2B and between NS2B and NS3, but that the reciprocal construct is inactive. The dengue protease processes heterologous cleavage sites more efficiently than homologous cleavage sites, suggesting that suboptimal cleavage efficiency may regulate the production of processing intermediates *in vivo*. By mutagenesis within the putative substrate binding pocket it was possible to change the substrate specificity of the yellow fever protease; changing a minimum of three amino acids in the yellow fever protease enabled it to recognize dengue cleavage sites. This system allows identification of residues which are directly or indirectly involved with enzyme-substrate interaction, that does not require a crystal structure, and can define the substrate preferences of individual members of a viral proteinase family.

Introduction

The *Flaviviridae* are a family of positive stranded RNA viruses whose genomes contain a single large open reading frame that is translated into a polyprotein. Co- and posttranslational processing of the polyprotein by viral and host cell proteinases produces ten individual proteins (32). The structural proteins are cotranslationally processed in the lumen of the endoplasmic reticulum by host cell signalases (21, 27, 33). The nonstructural proteins are processed in the cytosol by a virally encoded trypsin-like serine proteinase (NS3), as well as in the lumen of the endoplasmic reticulum by host cell or virally encoded signalases (10) (Chapter 1). NS3 cleaves between a lysine or arginine and a glycine, serine, or alanine to generate the amino termini of NS2B, NS3, NS4A, and, presumably, NS5. NS3 processes the 2A/2B cleavage site before processing the 2B/3 cleavage site, and the protease domain of NS3 is functionally contained within the amino terminal 184 amino acids (Chapter 1).

Sequence comparisons of viral proteinases and their cleavage sites are instructive for what they tell us about the evolution of proteinases and cleavage sites. Some regions of the NS3 protease domain are well conserved, but other regions are quite divergent. Presumably the rapidly evolving regions of the protease domain are not absolutely required to maintain structure or catalytic function, or alternatively some of these changes may be important for substrate recognition. Sites cleaved by NS3 are not well conserved, which suggests that a limited subset of amino acids and an unknown structure define a cleavage site. The rapid evolution of viral proteinases and cleavage sites suggests that viral proteinases should process homologous cleavage sites with greater

efficiency than heterologous cleavage sites and recent studies using chimeric picornavirus 3C proteinases and cassette mutagenesis to investigate enzyme-substrate interactions have provided support for this view (7, 8).

Viral proteinases are both enzymes and substrates, and the proteinase and cleavage sites coevolve to maintain the proper order and kinetics of polyprotein processing. A change in the substrate binding pocket must be counterbalanced by a compensating change in a cleavage site and *vice versa*. Viral proteinases regulate the production of processing intermediates by generating modified proteinases with altered substrate specificities and by varying the sequence of target cleavage sites to modulate the efficiency of cleavage (6, 41). The delicately regulated production of processing intermediates must be preserved during coevolution of proteinases and their cleavage sites.

A number of viral proteinases have been purified and their biochemical and physical properties have been extensively studied (24, 25, 26, 34, 40). Biochemical and genetic experiments to examine enzyme-substrate interactions have concentrated on defining what constitutes a cleavage site and have not identified residues within a proteinase that interact with substrates (16, 17, 29). We were interested in developing a biochemical assay that could identify residues which were important for substrate recognition. Chimeric protease domains (between yellow fever and dengue 2) were constructed that allowed mapping of regions involved in substrate recognition, and site-directed mutagenesis was used to modulate processing efficiency.

Materials and Methods.

Plasmid constructions: All plasmids were constructed using cDNA clones of the PR159 strain of dengue 2 and the Asibi strain of yellow fever virus, using standard recombinant DNA techniques (11, 12, 20). The plasmids pDD1 (which was named pT11) and p5'L213 have been previously described (Chapter 1). Recombinant polyproteins were created taking advantage of a *Dra*III site conserved between dengue and yellow fever, or using a *Hpa*I site in yellow fever that corresponds to a *Sal*I site in dengue.

Nomenclature: All constructs beginning with pD contain the cleavage sites of dengue, and all constructs beginning with pY contain the cleavage sites of yellow fever. The next letter specifies whether part or all of the protease domain of dengue (D) or yellow fever (Y) is present in the construct. Proteins produced by *in vitro* translation have been designated by the prefix P and the use of a prime superscript. For example, pDD1-programmed translations yield the precursor P2A2B3', the processing intermediate P2B3', and the product P3' (Chapter 1). Substrate binding pocket mutants of pDD1, pDY1 and pYD1 are designated using a decimal and a number. Substrate binding pocket residues are defined using a single letter to specify the wild-type amino acid, followed by a number specifying position within NS3, followed by a letter specifying the mutant amino acid.

pDD1 was digested with *Dra*III and *Nhe*I, and a *Dra*III to *Nhe*I fragment containing nucleotides 4720 to 5463 of the yellow fever genome (from cDNA clone 29) (11) was inserted to yield pDY1. *In vitro* transcription and translation of RNA from pDY1 produces a chimeric

polyprotein that contains the dengue 2A/2B and 2B/3 cleavage sites, the first 49 amino acids of dengue NS3, and the catalytic triad and substrate binding pocket of yellow fever.

pDY1 was digested with *Hpa*I and *Nhe*I, and a *Hinc*II to *Nhe*I fragment containing nucleotides 4942 to 5427 of the dengue genome (from plasmid pDD1) was inserted to yield pDY2. *In vitro* transcription and translation of RNA from pDY2 produces a chimeric polyprotein that contains the dengue 2A/2B and 2B/3 cleavage sites, the catalytic triad of yellow fever and the substrate binding pocket of dengue.

Yellow fever cDNA clone 29 was partially digested with *Hpa*I and completely digested with *Nhe*I to generate a fragment containing nucleotides 5000 to 5463 of the yellow fever genome. Vector pDD1 was completely digested with *Nhe*I and partially digested with *Hinc*II, and was ligated to the fragment described above to yield pDY3. *In vitro* transcription and translation of RNA from pDY3 produces a chimeric polyprotein that contains the dengue catalytic triad and the substrate binding pocket of yellow fever.

Yellow fever cDNA clone 29 was digested with *Ava*I, the 3' recessed end was filled in using *Escherichia coli* DNA polymerase I Klenow fragment, and then digested with *Bam*HI to generate a fragment containing nucleotides 3801 to 5463 of the yellow fever genome. Vector p5'L213 was digested with *Pst*I, the 3' overhang was made blunt by treatment with T4 DNA polymerase and then digested with *Bam*HI. Vector p5'L213 was ligated to the fragment described above to yield pYY1. *In vitro* transcription and translation of RNA from pYY1 produces a polyprotein that contains the first 37 amino acids of the dengue capsid

protein, the last 127 amino acids of yellow fever NS2A, all of yellow fever NS2B, and the first 299 amino acids of yellow fever NS3.

pYY1 was partially digested with *Dra*III, and completely digested with *Nhe*I, and a *Dra*III to *Nhe*I fragment containing nucleotides 4665 to 5427 of the dengue genome (from plasmid pDD1) was inserted to yield pYD1. *In vitro* transcription and translation of RNA from pYD1 produces a chimeric polyprotein that contains the yellow fever 2A/2B and 2B/3 cleavage sites, the first 50 amino acids of yellow fever NS3, and the catalytic triad and substrate binding pocket of dengue.

PCR mutagenesis: The polymerase chain reaction (PCR) was used to amplify cDNA containing nucleotides 4940 to 5071 of the dengue genome and nucleotides 5000 to 5112 of the yellow fever genome. Substrate binding pocket mutants were created using individual mutagenic oligonucleotide primers and an oligonucleotide primer of constant sequence to synthesize the plus and minus strands, respectively. pDD1 and pYY1, at 100 ng per ml, were used as templates for PCR amplification. Mutagenic and constant primers were used at a concentration of 1 nanomole per ml. A standard PCR cycle consisted of a 1 minute, 94° C denaturation step, followed by a 1 minute 50° C hybridization step, and a 2 minute 72° C polymerization step. After 15 cycles, PCR reaction mixes were phenol extracted, ethanol precipitated and resuspended in restriction endonuclease digestion buffer.

All PCR products and vectors used for cloning were isolated from low melting temperature agarose gels (Seakem) prior to ligation. Mutagenized dengue PCR products were digested with *Sal*I and *Asu*II

and ligated to *Sal*I and *Asu*II digested pYD1. Mutagenized yellow fever PCR products were digested with *Hinc*II and *Sac*I and ligated to *Hpa*I (partially digested) and *Sac*I digested pDY1. Recombinant mutagenized plasmids were analyzed by DNA sequencing and by *in vitro* transcription and translation of encoded RNAs (see below).

In vitro **transcriptions and translations:** *Nhe*I linearized templates were digested with proteinase K, phenol extracted and ethanol precipitated. Templates were resuspended in diethylpyrocarbonate treated, ribonuclease free water at a concentration of 0.5 μg/ml. Transcription reactions containing 0.5 mM rNTPs (Pharmacia), 0.5 mM cap analog $m^7G(5')ppp(5')G$ (New England Biolabs), 50 ng/μl of template, and 2 units/μl of T7 RNA polymerase (Pharmacia) in 1X transcription buffer were incubated at 37°C for 45 minutes (31). Micrococcal nuclease treated rabbit reticulocyte lysates containing 0.025 Eq/μl microsomal membranes (Promega) were programmed with *in vitro* transcribed RNA at a concentration of 5 to 10 μg/ml (5). All translations were carried out at 30°C for 90 minutes unless otherwise noted.

Results

***In vitro* processing of chimeric proteinases:** The high error frequency of RNA-dependent RNA polymerases is reflected in the divergence rates of proteins encoded by RNA viruses (9, 14, 35). The protease domain of NS3 and the small hydrophobic proteins NS2A and NS2B are 55%, 19%,and 33% conserved respectively between dengue 2 and yellow fever. The aligned amino acid sequences of the protease domains found in the N-termini of the NS3's of DEN and YF are shown in Fig. 1A. and sequence for several flaviviruses of the amino acids surrounding two of the cleavage sites processed by this protease are displayed in Fig. 1B. It is evident that the cleavage sites and some portions of the polyprotein are less well conserved than the central core of the protease. NS3 cleaves between a basic residue (usually arginine) and a glycine, serine or alanine, but helix breaking and acidic residues (boxed in Fig. 1B) may also be important for substrate recognition.

We have previously described the construction of a expression plasmid, pDD1; *in vitro* transcription and translation of RNA from pDD1 produces protein products that are cleaved at the 2A/2B and 2B/3 sites (Chapter 1). For this project we also constructed plasmid pYY1, in which the leader sequence and the domain encoding the first 37 amino acids of the capsid protein of dengue virus were fused in frame to the sequences encoding the C-terminal region of NS2A, all of NS2B, and most of NS3 from yellow fever virus. These constructs are illustrated in Figure 2A. *In vitro* transcription and translation of RNA from pYY1 resulted in correct processing of the yellow fever proteins at the 2A/2B and 2B/3 sites (Fig 2B). The yellow fever protease is more efficient than the dengue protease at

Fig 1: Sequence comparisons of enzymes and substrates.

Panel A: The amino acid sequences of the N-terminal domains of the NS3 proteins of dengue 2 PR159 S1 (DEN2) and yellow fever Asibi (YF) have been aligned. Residues are numbered from the N-terminus of the protein. The three enzymatic domains and the substrate binding pocket as postulated by Bazan and Fletterick (1989) are boxed, and the amino acids which form the putative catalytic triad (His, Asp, and Ser) are shaded. Proposed contact residues are circled. The *Dra*III and *SalI/Hpa*I restriction sites which were used to make chimeric proteases are indicated. The single letter amino acid code is used throughout: A, alanine; C, cysteine; D, aspartate; E, glutamate; F, phenyalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan and Y, tyrosine.

Panel B: Aligned 2A/2B and 2B/3 cleavage sites for several flaviviruses. Cleavage usually occurs between two basic residues (shaded) and a glycine, serine, or alanine (shaded). Conserved prolines and acidic residues downstream of the cleavage site are boxed. Virus abbreviations and sequences used are: DEN2, dengue 2 PR159 S1 (12); DEN3, dengue 3 (28); DEN4, dengue 4 (18); MVE, Murray Valley encephalitis (4); KUN, Kunjin (3); SLE, Saint Louis encephalitis (36); WN, West Nile (2); YF, yellow fever Asibi strain (11); JE, Japanese encephalitis virus Beijing 1 strain (13) and TBE, tick borne encephalitis (19).

A.

DEN2 AG–VLWDVPSPPPVGKAE–LEDGAYRIKQKGILGYSQIGA
YF   S.D....I.T.KIIEEC.H....I.G.F.STF...A..R.V

DEN2 GVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLIS
YF   ..AQG.V......F.VRN.KLI...S..E..VA

DEN2 YGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFRT
YF   ....S....R.DGE....LI.AV..VVN.....S...KV

DEN2 –NTGTIGAVSLOOSPGTSGSPIVDKKGKVVKLOOGVVTR
YF   R.G.E....A.OOPS..NRN..E.I..OO.ILVG

DEN2 SGAYVSAIAQTEKSIEDNPEIEDDIFRKRRLTIMDLHPGA
YF   DNSF....S..VKE.GKE.LQEIPTMLKKGMTTV.DFHP

B.

	NS2A	NS2B		
DEN 2	TSKKR	SWPLNE	AIMA	
DEN 3	.L.R.	...G.	GV..	
DEN 4	GASR.	.G.	G...	
MVE	PN..	.AT.	VLT.	
WN	PNR.	.AT.	VMT.	
KUN	PNR.	.AT.	VMT.	
SLE	PNG.	.AS.	VLTG	
JE	PN..	.AT.	FLS.	
YF	IFGR.	I.V.	.LA.	
TBE	HRGR.	F––S.	PLTV	

	NS2B	NS3	
DEN2	VKKQR	AG–VLWDVPS	
DEN3	KQT.	S....	
DEN4	.T	..–A..	
MVE	KYTK.	G..–.F.	
WN	QYTK.	G..I..T	
KUN	QYTK.	G..I..T	
SLE	KHSK.	G..–A..	
JE	KTTK.	G..–.F..	
YF	RGAR.	S.D...I.T	
TBE	RSSR.	SDLVFSGQGG	

processing homologous 2A/2B and 2B/3 cleavage sites *in vitro*.

In order to investigate whether the two NS3 proteases, which have diverged significantly, could process heterologous cleavage sites which have also diverged we constructed plasmids expressing chimeric polyproteins of yellow fever and dengue 2 (Fig. 2A). Complete details of the NS3 amino acids represented in these constructs are listed in Table 1.

The first constructs tested were two reciprocal chimeras, pDY1 and pYD1, that were made using a conserved *Dra*III restriction site (which lies just upstream of the putative catalytic histidine residue, Fig. 1A). In essence, the 2A/2B and 2B/3 sites in each chimera are derived from one virus and the enzymatic domains from the other, so that we can examine the activity of each protease on a heterologous substrate. Since *in vitro* transcription and translation of pDY1 produced an unprocessed full-length translation product (Fig. 2B), we concluded that the the yellow fever protease could not cleave dengue sites. (The full-length polyprotein contains sequences within P2A and NS2B that cause the polyprotein to migrate as a heterogeneous smear in SDS PAGE gels.) In contrast, *in vitro* transcription and translation of pYD1 produces fully processed P3', suggesting that the 2A/2B and 2B/3 cleavage sites were correctly processed. Small amounts of higher molecular weight P3' polypeptides are sometimes observed, suggesting that cryptic cleavage sites within P2A or NS2B of yellow fever were recognized with a low frequency. The processing of the precursors in the translation mix of the chimera pYD1 is more complete than processing by the parental dengue construct (compare lane pYD1 with lane pDD1), suggesting that the dengue protease domain

Fig. 2: Structure and phenotypes of chimeric proteinases.

Panel A: Parental and chimeric proteinase constructs are displayed. Processing at the 2A/2B and 2B/3 cleavage sites is indicated by downward facing arrowheads. Dengue and yellow fever catalytic histidine, aspartate, and serine residues are indicated by vertical bars and asterisks within open boxes, respectively. The dengue and yellow fever substrate binding pockets are indicated by open and solid boxes, respectively. Dengue capsid protein, P2A, NS2B and P3' sequences are designated by light diagonally shaded, open, lightly shaded, and solid boxes, respectively. Yellow fever P2A, NS2B, and P3' sequences are designated by lightly shaded, heavy diagonally shaded, and open boxes, respectively.

Panel B: An autoradiograph of a SDS PAGE gel displaying protein patterns from *in vitro* translation mixes programmed with RNA transcribed from constructs in panel A. Unprocessed precursor P2A2B3' species (lanes pDY1, 2, 3 and pDD1), intermediate P2B3' species (lanes pDY2 and pDD1), product P3' species (lanes pDY3, pDD1, pYY1 and pYD1), and a 46 kDa reticulocyte lysate specific-band are indicated.
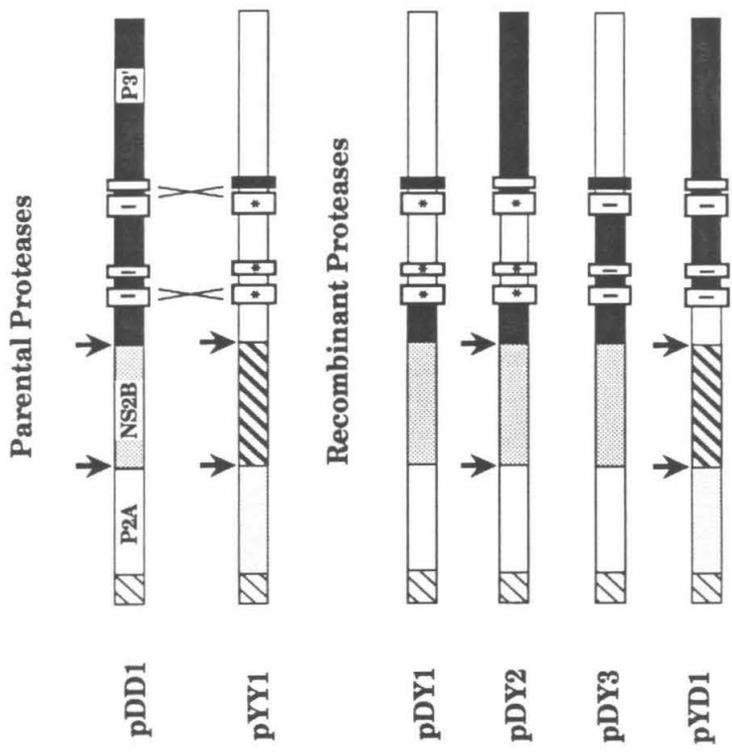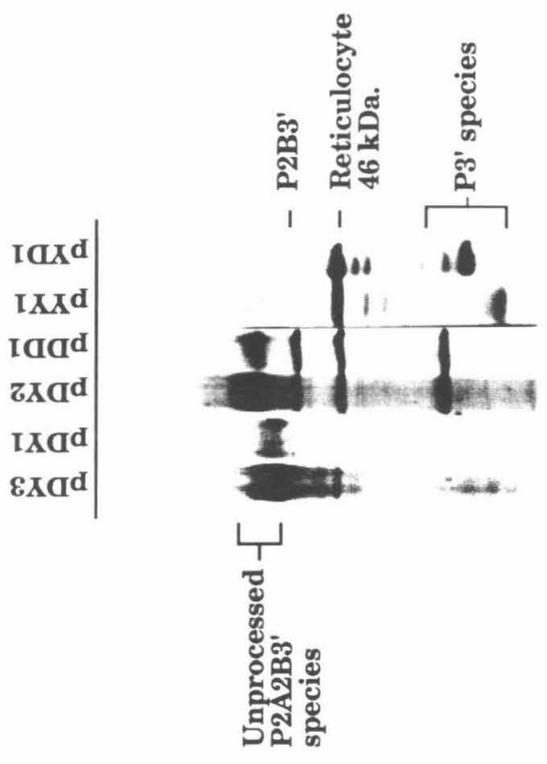
Table 1: Translation constructs.

Construct	C.S.[a]	Catalytic[b]	SBP[c]	DEN2[d]	YF
pDD1	DEN2	DEN2	DEN2	1-303	NA
pYY1	YF	YF	YF	NA	1-299
pDY1	DEN2	YF	YF	1-49	52-299
pDY2	DEN2	YF	DEN2	1-49, 141-303	52-143
pDY3	DEN2	DEN2	YF	1-140	144-301
pYD1	YF	DEN2	DEN2	50-303	1-51

Footnotes

[a, b, c] The catalytic triad, substrate binding pocket (SBP), and 2A/2B, 2B/3 cleavage site (C.S.) residues represented in each construct.

[d] Dengue 2 (DEN2) and yellow fever virus (YF) NS3 amino acids represented in each construct.

processes heterologous cleavage sites more efficiently than homologous cleavage sites.

The yellow fever substrate binding pocket of pDY1 was replaced with the dengue substrate binding pocket in plasmid pDY2. This construct contains the catalytic triad of yellow fever combined with the substrate binding pocket and cleavage sites of dengue 2. *In vitro* transcription and translation of pDY2 produces the processing intermediate P2B3' and the expected P3' product. The precursor (P2A2B3' and P2B3') to product (P3') ratio suggest that the efficiency of polyprotein processing is lower than that of the parental construct, pDD1 (compare lane pDY2 and lane pDD1) This result suggested that the yellow fever substrate binding pocket does not efficiently recognize dengue cleavage sites. In plasmid pDY3 the substrate binding pocket of pDD1 was replaced with the yellow fever substrate binding pocket. Translation of pDY3 RNA produced an unprocessed full-length polyprotein (Fig. 2B). These results clearly showed that residues 141 to 184 of the dengue protease domain are important for dengue cleavage site recognition.

Mutagenesis of the substrate binding pocket: Bazan and Fletterick recently proposed a model for the flavivirus proteinase that is based upon structural and functional analogies with the trypsin-like serine proteinases (1). The model predicts that five conserved residues are in direct contact with bound substrate (circled in Fig.1A). Three of these are located within the proposed substrate binding pocket; the other two are located just upstream of the putative catalytic serine. The nonreciprocal processing of dengue and yellow fever 2A/2B and 2B/3 cleavage sites

suggested that other nonconserved residues were also important for substrate recognition.

The chimeric proteinase expression experiments pinpointed amino acids 141 to 184 of the dengue protease domain as being important for substrate recognition. Residues 155 to 161 and 170 to 184 have diverged significantly between yellow fever and dengue whereas residues 141 to 154 which are predicted to form part of a loop and beta strand structure surrounding the proposed catalytic serine (1) contain interspersed conserved and nonconserved residues (Fig. 1A). Five amino acid differences between dengue and yellow fever virus, three of which are nonconservative, are found within this loop and beta strand. To test the importance of these residues in substrate recognition, PCR mutagenesis was used to interchange four of the dengue, and five of the yellow fever residues. D141 of dengue was not changed to the corresponding asparagine residue of yellow fever because of cloning considerations. In order to ascertain which amino acids played a dominant role in substrate recognition, sets of conservative and nonconservative mutations were made, and nonconservative mutations were made individually. The strategy used to generate mutants is schematically presented in Fig. 3 and full details of the templates and primers used are presented in Table 2.

pDD1 was mutagenized to change K142, K143, K145 and V147 to the corresponding yellow fever amino acids (R, N, E, and I respectively), to produce plasmid pDD1.1. The precursor (P2A2B3' and P2B3') to product (P3') ratio of a pDD1.1-programmed translation was greater than a parental pDD1-programmed translation, suggesting a lower cleavage efficiency (Fig 4A).

Fig. 3: PCR mutagenesis strategy used to generate substrate binding pocket mutants.

Panel A: Dengue and yellow fever NS3 protease domains are displayed as in Fig. 2. The relative positions of *Sal*I, *Asu*II, *Hpa*I/*Hinc*II and *Sac*I restriction sites used for cloning are shown.

Panel B: Four and five residues respectively were changed individually and in sets within the dengue and yellow fever substrate binding pockets to generate eleven different mutants. Sequences of wild-type and mutant substrate binding pockets are shown (mutant residues are shaded). Full details of the primers used are described in Table 2.

A. DENGUE 2 PROTEASE
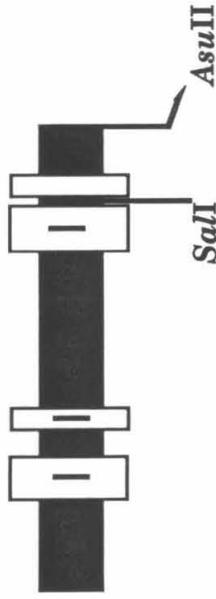
YELLOW FEVER PROTEASE
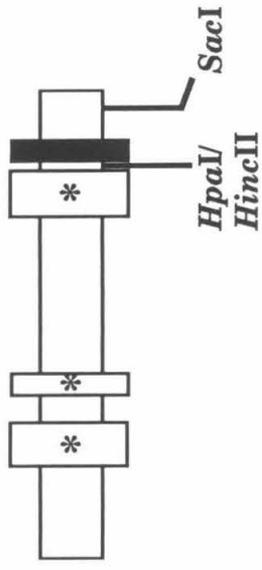
B. DENGUE 2 SBP

YELLOW FEVER SBP

PCR MUTAGENESIS

Table 2: Substrate binding pocket mutagenesis primers.

Parent	Progeny		Oligonucleotide [a]
pYD1	pYD1.1	D1	CAATCGTCGACAGAAACGGAGAAGTTATAGGTCTCTA
pYD1	pYD1.2	D2	ATCGTCGACAGAAAAGGAAAAGTTATAG
pYD1	pYD1.3	D3	ATCGTCGACAAAAACGGAGAAG
pYD1	pYD1.4	D4	ATCGTCGACAAAAACGGA
pYD1	pYD1.5	D5	ATCGTCGACAAAAAAGGAGAA
pDY1	pDY1.1	Y1	TCTCCTATTGTTGACAAGAAAGGAAAGGTGGTTGGGCTGTA
pDY1	pDY1.2	Y2	TATTGTTGACAGGAAAGGAAAGGT
pDY1	pDY1.3	Y3	ATTGTTAACAAGAACGGAGAGGTGGTTG
pDY1	pDY1.4	Y4	TATTGTTGACAGGAAC
pDY1	pDY1.5	Y5	ATTGTTAACAGGAAAGGA
pDY1	pDY1.6	Y6	ATTGTTAACAGGAACGGAAAG
		DM	GTGGAGATCCATGATAG
		YM	TCTCTTGGAGCTCCTCC

Footnotes

[a] The name and sequence (5' to 3') of oligonucleotide primers used to generate substrate binding pocket mutants are listed in upper and lower case respectively. DM and YM are minus strand primers that were used to synthesize dengue and yellow fever mutant SBPs, respectively. Primers Y2, Y4 and Y5 were hybridized at 40° C (see Materials and Methods).

pDY1 was mutagenized to change N144, R145, N146, E148 and I150 to the corresponding dengue amino acids (D, K, K, K, and V respectively) to produce plasmid pDY1.1. Changing these amino acids allowed the yellow fever protease domain to process dengue cleavage sites (Fig. 4A). The processing efficiencies of pDY1.1- and pYD1-programmed translations were very similar (see below). The presence of fully cleaved P3' suggested that the 2A/2B and 2B/3 cleavage sites had been correctly processed.

pYD1 was mutagenized to change K142, K143, K145 and V147 to the corresponding yellow fever amino acids (R, N, E, and I respectively), to produce plasmid pYD1.1. Changing these amino acids appeared to change the efficiency of processing and eliminated the production of several minor P3' species. The precursor (P2A2B3' and P2B3') to product (P3') ratios of pYD1.1- and pDD1.1-programmed translations were very similar. Changes in the electrophoretic mobilities of P3' and P2B3' translated from pYD1.1 are presumably due to the amino acid substitutions in this mutant.

Since processing efficiency was clearly affected by multiple mutations in the dengue and yellow fever substrate binding pockets, we wondered whether the effects observed could be attributed to a single amino acid or to some subset of amino acid changes. One, two or three amino acids in pDY1, pYD1 and pDD1 (data not shown) were changed and their affects on processing were assayed *in vitro*.

pDY1 mutagenesis: Three amino acid substitutions in mutant pDY1.2, (N144D, N146K and E148K) were sufficient to allow the yellow fever

protease domain to process dengue cleavage sites. The processing patterns produced by pDY1.2- and pDY1.1- programmed translations were very similar, suggesting that these three residues played a dominant role in changing the substrate specificity of the yellow fever protease domain (Fig 4B). The yellow fever protease domain was not activated when each of the nonconservative amino acid changes were made individually in mutants pDY1.4 (N144D), pDY1.5 (N146K), pDY1.6 (E148K). Two amino acids changes in mutant pDY1.3 (R145K and I150V) did not change the specificity of the yellow fever protease domain. In summary a minimum of three nonconservative amino acid changes were necessary in order to activate the yellow fever protease domain towards processing dengue cleavage sites. Changing these residues individually or making conservative amino acids changes in tandem had no effect.

pYD1 mutagenesis: pYD1.2- (K142R and V147I), pYD1.3- (K143N and K145E), pYD1.4- (K143N) and pYD1.5- (K145E) and pYD1-programmed translations patterns were very similar. The processing efficiency was only significantly altered when all four amino acids (pYD1.1) were changed simultaneously. Similar results were observed when the equivalent mutations were introduced into a pDD1 background (F. Preugschat, unpublished data).

Fig.4: *In vitro* translation of substrate binding pocket mutants. Mutagenized plasmids were linearized, transcribed and translated in reticulocyte lysates as described in Materials and Methods.

Panel A: pDD1, pDD1.1, pDY1, pDY1.1, pYD1, and pYD1.1 programmed *in vitro* translations were resolved on an SDS PAGE gel and autoradiographed. Unprocessed P2A2B3' species (lanes pDD1.1, pYD1.1 and pDY1), intermediate P2B3' species (lanes pDD1.1, pDD1 and pYD1.1) processed P3' species (lanes pDD1.1, pDD1, pYD1.1, pYD1 and pDY1.1) and a 46kDA reticulocyte lysate-specific band are indicated. The left pDD1.1 lane, and lanes pYD1.1 and pDY1, have been overexposed to emphasize the reduced cleavage efficiency exhibited by these constructs. Panel B: The left side of panel B displays pDY1.1, pDY1.2, pDY1.3, pDY1.4, pDY1.5 and pDY1.6 programmed *in vitro* translations that were resolved on an SDS PAGE gel and autoradiographed. The right side of panel B displays pYD1.1 pYD1.2, pYD1.3, pYD1.4 and pYD1.5 programmed *in vitro* translations that were resolved on an SDS PAGE gel and autoradiographed. Lane B is a blank, unprogrammed reticulocyte lysate *in vitro* translation. Unprocessed P2A2B3' species (lanes pDY1.3, 1.4, 1.5, 1.6 and pYD1.1), processed P3' species (lanes pDY1.1, 1.2, pYD1.2, 1.3, 1.4 and 1.5), and a 46kDA reticulocyte lysate-specific band are indicated.

A.
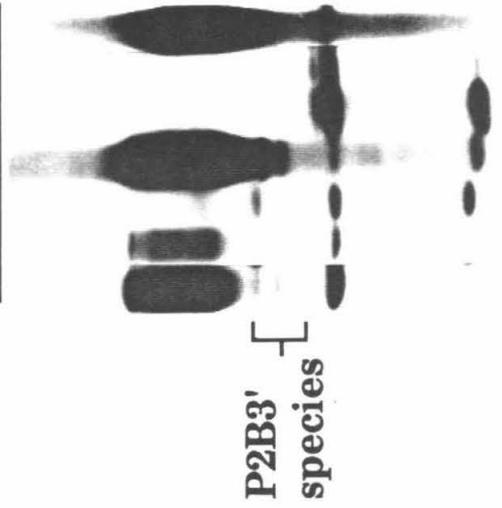
B.
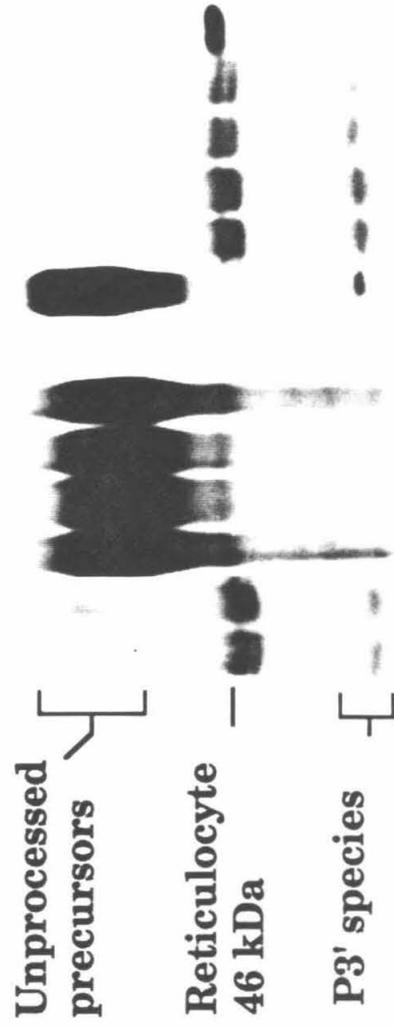
Discussion

Structure-function relationships: Molecular chimeras have been used to exchange antigen binding loops, DNA recognition helices, or whole domains involved in ligand binding and have demonstrated that mini-domains of large proteins can function independently of the whole (15, 22, 38, 39). Experiments with bacterial subtilisins proved that the substrate specificity of two serine proteases (which are 69% conserved, but which have virtually superimposable three-dimensional structures), could be interconverted by mutagenizing selected amino acids within the substrate binding pocket (37). We have used chimeric protease domains which express a portion or all of the protease domain of one virus, and the cleavage sites of another to map residues which are important for substrate recognition. Mutagenesis of substrate binding pocket residues enabled us to modulate the cleavage efficiency of an active protease, and to exchange the substrate specificity of one protease with that of another.

The cleavage generating the amino terminus of dengue NS3 is unique to the dengue virus subgroup. Cleavage occurs between a single arginine and an alanine or serine, whereas other flaviviruses generate the amino terminus of NS3 by cleaving betweeen two basic amino acids and a serine, or glycine (Fig. 1B). The yellow fever protease domain does not process dengue cleavage sites whereas the dengue protease domain appears to process yellow fever cleavage sites more efficiently than its own cleavage sites. This suggests that dengue cleavage sites have not evolved an optimal sequence or structure, and that the suboptimal efficiency may be used to regulate the production of processing intermediates (which can have different functions than end products) (6, 41). The dengue protease

has a broader substrate specificity than the yellow fever protease, and would be potentially useful in assaying for a broad spectrum flavivirus proteinase inhibitor.

The small, hydrophobic nonstructural proteins NS2A and NS2B, which are only 19% and 33% conserved between yellow fever and dengue 2, are active substrates for proteolysis, which suggests that they have similar structures. Since the P2B3' intermediate is seen in both pYD1- and pDY1.1-programmed translations, cleavage occurs with the correct order and presumably the same specificity as *in vivo* (Chapter 1). The protease domain of NS3 is 55% conserved between dengue and yellow fever, and our data suggests that these proteases must have similar structures. The pDY1 and pDY3 chimeric proteases are not active, presumably because the substrate binding pocket of yellow fever does not efficiently recognize dengue cleavage sites, but it is possible that long- or short-range folding effects have perturbed the protease structure. It is important to note that any negative effects caused by misfolding can be suppressed by changing only three amino acids in the substrate binding pocket (see pDY1.2 translation Fig. 4B.).

Fine-scale mapping of the protease domain: Site-directed mutagenesis and crystallography studies of the bacterial subtilisins have produced the most detailed picture of serine protease enzyme-substrate interactions. Four short stretches of amino acids (which are distributed over a 126 amino acid segment of the protease) are brought into close proximity with substrate by tertiary folding of the polypeptide chain (37). By analogy, it seems likely that multiple determinants of substrate

specificity are distributed throughout the flaviviral nonstructural proteinase. The molecular chimeras allowed us to localize one region of the protease domain that was important for substrate recognition and we found that residues which differed between yellow fever and dengue could account for some of the differences in cleavage efficiency.

Only multiple, simultaneous changes within the protease domain noticeably affected cleavage efficiency or specificity. Individual mutations and many of the multiple mutations had no effect. The phenotype of the mutant also depended on the genetic background in which the mutations were made. Protease activity was not observed in pDY3-programmed translations, and therefore it was surprising that pDD1.1-programmed translations produced an active protease. We would have predicted that changing four residues in the substrate binding pocket to their yellow fever counterparts would inactivate the dengue protease, but the pDD1.1 results suggest that there are other residues in the protease that are involved in enzyme-substrate interaction. These residues could act directly by interacting with bound substrate, or indirectly by altering the conformation of the protease domain. The *in vitro* expression assay used in our experiments is not sensitive enough to detect small changes in efficiency and will only readily detect a gain or loss of function. The subtle effects of multiple or single mutations could be assayed biochemically using purified proteases, or biologically using infectious clones (30) (Chapter 3).

We have identified one intragenic locus that is involved in enzyme-substrate interaction, but there are probably other loci which were not mapped by our experiments. The yellow fever protease domain will

process dengue cleavage sites when a minimum of three amino acids are changed, and the gain of function argues strongly that the introduced mutations are important for substrate recognition. There are a total of 54 amino acid differences between the protease domains of pDY1 and pYD1. By taking advantage of endogenous or engineered restriction sites it is possible to systematically exchange all of the divergent residues and to map other dominant loci that determine substrate specificity.

Designer proteases: The remarkable cleavage specificity of viral proteinases make them ideal biochemical targets for antiviral drugs, and recently developed proteinase inhibitors offer hope for the treatment of incurable disease (23). Knowledge of the sequence and structure of cleavage sites, and of the nature of the substrate binding pocket of the proteinase are necessary in order to design effective biochemical inhibitors. Designing a group-specific proteinase inhibitor for a viral family would require knowledge of the substrate preferences of individual family members. We have described a system that allows identification of residues which are directly or indirectly involved with enzyme-substrate interaction, that does not require a crystal structure, and can define the substrate preferences of individual members of a viral family. Our studies can be extended to other members of the flavivirus family to expand our understanding of substrate recognition. Our approach has limitations but can be useful for identifying residues which are important for enzyme-substrate interactions in other viral families.

Acknowledgements

References Cited

1. **Bazan, J. F., and R. J. Fletterick.** 1989. Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. Virology **171**:637-639.

2. **Castle, E., U. Leidner, T. Nowak, G. Wengler, and G. Wengler.** 1986. Primary structure of the West Nile flavivirus genome region coding for all the nonstructural proteins. Virology **149**:10-26.

3. **Coia, G., M. D. Parker, G. Speight, M. E. Byrne, and E. G. Westaway.** 1988. Nucleotide and complete amino acid sequences of Kunjin virus: Definitive gene order and characteristics of the virus-specified proteins. J. Gen. Virol. **69**:1-21.

4. **Dalgarno, L., D. W. Trent, J. H. Strauss, and C. M. Rice.** 1986. Partial nucleotide sequence of Murray Valley encephalitis virus: Comparison of the encoded polypeptides with yellow fever virus structural and nonstructural proteins. J. Mol. Biol. **187**:309-323.

5. **Dasso, M. C., and R. J. Jackson.** 1989. On the fidelity of mRNA translation in the nuclease treated rabbit reticulocyte lysate system. Nucl. Acids. Res. **17**:3129-3144.

6. **de Groot, R. J., W. R. Hardy, Y. Shirako, and J. H. Strauss.** 1990. Cleavage site preferences of Sindbis virus polyproteins containing the nonstructural proteinase: Evidence for temporal regulation of polyprotein processing *in vivo*. EMBO, J. in press.

7. **Dewalt, P. G., W. S. Blair, and B. L. Semler.** 1990. A genetic locus in mutant poliovirus genomes involved in overproduction of RNA polymerase and 3C proteinase. Virology **174**:504-514.

8. **Dewalt, P. G., M. A. Lawson, R. J. Colonno, and B. L. Semler.** 1989. Chimeric picornavirus polyproteins demonstrate a common 3C proteinase substrate specificity. J. Virol. **63:**3444-3452.

9. **Doolittle, R. F., D.-F. Feng, M. S. Johnson, and M. A. McClure** 1989. Origins and evolutionary relationships of retroviruses. Quarterly Review of Biology **64:**1-30.

10. **Falgout, B., R. Chanock, and C. J. Lai.** 1989. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2a. J. Virol. **63:**1852-1860.

11. **Hahn, C. S., J. M. Dalrymple, J. H. Strauss, and C. M. Rice.** 1987. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. Proc. Natl. Acad. Sci. **84:**2019-2023.

12. **Hahn, Y. S., R. Galler, T. Hunkapiller, J. Dalrymple, J. H. Strauss, and E. G. Strauss.** 1988. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. Virology **162:**167-180.

13. **Hashimoto, H., A. Nomoto, K. Watanabe, T. Mori, T. Takezawa, C. Aizawa, T. Takegami, and K. Hiramatsu.** 1988. Molecular cloning and complete nucleotide sequence of the genome of Japanese encephalitis virus Beijing-1 strain. Virus Genes **1:**305-317.

14. **Holland, J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol, and S. VandePol.** 1982. Rapid evolution of RNA genomes. Science **215:**1577-1585.

15. **Jones, P. T., P. H. Dear, J. Foote, M. S. Neuberger, and G. Winter.** 1986. Replacing the complementarity determining regions in a human antibody with those from a mouse. Nature **321:**522-525.

16. **Kotler, M., R. A. Katz, W. Danho, J. Leis, and A. M. Skalka.** 1988. Synthetic peptides as substrates and inhibitors of a retroviral protease. Proc. Natl. Acad. Sci. USA. **85:**4185-4189.

17. **Kraeusslich, H. G., R. H. Ingraham, M. T. Skoog, E. Wimmer, P. V. Pallai, and C. A. Carter.** 1989. Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural and synthetic peptides. Proc. Natl. Acad. Sci. USA **86:**807-811.

18. **Mackow, E., Y. Makino, B. Zhao, Y.-M. Zhang, L. Markoff, A. Buckler-White, M. Guiler, R. M. Chanock, and C.-J. Lai.** 1987. The nucleotide sequence of dengue type 4 virus: Analysis of genes coding for nonstructural proteins. Virology **159:**217-228.

19. **Mandl, C. W., F. X. Heinz, E. Stoeckl, and C. Kunz.** 1989. Genome sequence of tick-borne encephalitis virus (western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. Virology **173:**291-301.

20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

21. **Markoff, L.** 1989. *In vitro* processing of dengue virus structural proteins: Cleavage of the pre-membrane protein. J. Virol. **63:**3345-3352.

22. **Mas, M. T., C. Y. Chen, R. A. Hiteman, and A. D. Riggs.** 1986. Active human-yeast chimeric phosphoglycerate kinases engineered by domain interchange. Science **233:**788-790.

23. **McQuade, T. J., A. G. Tomaselli, L. Liu, V. Karacostas, and B. Moss.** 1990. A synthetic HIV protease inhibitor with antiviral activity arrests HIV-like particle maturation. Science **247:**454-456.

112

24. **Miller, M., M. Jaskolski, J. K. M. Rao, J. Leis, and A. Wlodawer.** 1989. Crystal structure of a retroviral protease proves relationship to aspartic protease family. Nature **337:**576-579.

25. **Navia, M. A., P. M. D. Fitzgerald, B. M. McKeever, C. T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer.** 1989. Three dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature **337:**615-620.

26. **Nicklin, M. J. H., H. Toyoda, M. G. Murray, and E. Wimmer.** 1988. Poliovirus 3C: Large-scale expression, purification and specific cleavage activity on natural and synthetic substrates *in vitro*. J. Virol. **62:**4586-4594

27. **Nowak, T., P. M. Farber, G. Wengler, and G. Wengler.** 1989. Analyses of the terminal sequences of West Nile virus structural proteins and of the translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. Virology **169:**365-376.

28. **Osatomi, K., and H. Sumiyoshi.** 1990. Complete nucleotide sequence of dengue type 3 virus genome RNA. Virology. in press.

29. **Pallai, P. V., F. Burkhardt, M. Skoog, K. Schreiner, P. Bax, K. A. Cohen, G. Hansen, D. E. H. Palladino, K. S. Harris, M. J. Nicklin, and E. Wimmer.** 1989. Cleavage of synthetic peptides by purified poliovirus 3C proteinase. J. Biol. Chem. **264:**9738-9741.

30. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers.** 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by *in vitro* ligation. New Biol. **1:**285-296.

31. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: Mapping of lethal mutations, rescue of a temperature-sensitive marker,

and *in vitro* mutagenesis to generate defined mutants. J. Virol. **61**:3809-3819.

32. **Rice, C. R., E. G. Strauss, and J. H. Strauss.** Structure of the flavivirus genome, pp. 279-327. In The Togaviridae and Flaviviridae. S. Schlesinger and M. J. Schlesinger (ed.), Plenum Publishing Corp., New York, 1986.

33. **Ruiz-Linares, A., A. Cahour, P. Despres, M. Girard, and M. Bouloy.** 1989. Processing of the yellow fever virus polyprotein: Role of cellular proteases in the maturation of the structural proteins. J. Virol. **63**:4199-4209.

34. **Schneider, J., and S. B. H. Kent.** 1988. Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. Cell **54**:363-368.

35. **Steinhauer, D. A., and J. J. Holland.** 1987. Rapid evolution of RNA viruses. Ann. Revs. Microbiol. **41**:409-433.

36. **Trent, D. W., R. M. Kinney, B. J. Johnson, A. V. Vorndam, J. A. Grant, V. Deubel, C. M. Rice, and C. Hahn.** 1987. Partial nucleotide sequence of St. Louis encephalitis virus RNA: Structural proteins, NS1, NS2A, and NS2B. Virology **156**:293-304.

37. **Wells, J. A., B. C. Cunningham, T. P. Graycar, and D. A. Estell.** 1987. Recruitment of substrate specificity properties from one enzyme into a related one by protein engineering. Proc. Natl. Acad. Sci. USA **84**:5167-5171.

38. **Wharton, R. P., E. L. Brown, and M. Ptashne.** 1984. Substituting an alpha helix switches the sequence-specific DNA interactions of a repressor. Cell **38**:361-369.

39. **Wharton, R. P., and M. Ptashne.** 1985. Changing the binding specificity of a repressor by redesigning an alpha helix. Nature **316:**601-605.

40. **Yoshinaka, Y., I. Katoh, T. D. Copeland, G. W. Smythers, and S. Oroszlan.** 1986. Bovine leukemia virus protease: Purification, chemical analysis and *in vitro* processing of the gag precursor. J. Virol. **57:**826-832.

41. **Ypma-Wong, M. F., P. G. Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166:**265-270.

Chapter 3

In Vitro Transcription of Genomic RNA from Dengue 2 cDNA Templates.

In Vitro Transcription of Genomic RNA from Dengue 2 cDNA Templates.

Frank Preugschat and James H. Strauss*.

Division of Biology 156-29

California Institute of Technology,

Pasadena, CA. 91125.

*Author to whom correspondence should be addressed.

Abstract

Dengue viruses are members of the flavivirus family, a diverse group of viruses responsible for a number of human and veterinary diseases. The dengue 2 PR159 S1 strain is a first-generation candidate vaccine strain that has been used successfully in human subjects. We have developed techniques to produce full-length genomic RNA from cloned dengue 2 PR159 S1 cDNA templates. Methods are described for the transcription, transfection and detection of infectious dengue RNA. These experiments lay the foundation for a second-generation approach for the development of dengue virus vaccines.

Introduction

The dengue viruses are members of the flavivirus family, and are responsible for rapidly expanding disease outbreaks in the tropical and subtropical worlds. The four serotypes of dengue viruses are transmitted to man during a blood meal taken by infected female *Aedes* mosquitoes. Dengue infection causes fever, rash, and joint pain, and the most severe forms invoke a general fatigue that can last for months after the initial attack. Dengue fever is rarely fatal, but roughly 5% of infected individuals develop a secondary shock syndrome (DSS) or hemorrhagic fever (DHF) which can be fatal (5, 6, 8). The factors contributing to the development of DSS and DHF are not understood, but it has been suggested that cross-reacting antibodies aid the establishment of a secondary infection by a different serotype (19, 20).

All four serotypes of dengue viruses are present in the Americas and conditions for a DHF epidemic resemble those present in southeast Asia, where over 600,000 cases were reported in 1987 alone (17, 19). The worldwide incidence of DSS and DHF is increasing, and with over 1.5 billion people at risk for infection the need for an effective vaccine has increased proportionally. There is no animal model for dengue fever, and an additional complication for vaccine development is that each of the four serotypes is comprised of a number of groups or topotypes which have diverged genetically by approximately 10% (7, 35). Recently a number of dengue viral genomes have been partially or completely sequenced, and we now have a better understanding of the genetic relationships between sero- and topotypes, which will be useful for the design of second-generation vaccines (11, 12, 16, 18, 21, 24, 26, 27, 36, 37).

In 1969 a temperature-sensitive dengue 2 virus was isolated from the blood of an infected Puerto Rican, and after 19 passages in cell culture a candidate vaccine strain was developed (13, 14). The dengue 2 PR159 S1 strain exhibits a small plaque phenotype and has reduced virulence for monkeys and suckling mice. 70% of vaccinated human volunteers developed antibodies against dengue 2 and the seroconversion percentage was even higher for individuals that had been previously immunized against yellow fever. The vaccine strain has not been put into general use because the virus has not been completely attenuated (2, 3, 4, 33).

The sequence of all but the 5' terminal 15 nucleotides of the dengue 2 PR159 S1 genome was derived from virus that had been passaged 21 times in cell culture (18). The randomly primed cDNA library contains approximately 50-75 genome equivalents and should be representative of the small plaque vaccine strain. The cDNA library was used to construct several clones that can be transcribed to produce full-length genomic RNA, and this chapter describes the development of a number of methods and reagents that lay a foundation for the construction of infectious cDNA clones.

Infectious cDNA clones and cDNA clones from which infectious RNA can be transcribed, have been constructed for a number of positive stranded RNA viruses (1, 10, 22, 28, 30, 34). The flaviviruses differ from other positive stranded RNA viruses in that full-length cDNA copies of the genome can not be propagated using plasmid vectors. Full-length yellow fever genomes have been constructed in lambda vectors, but are genetically unstable (29). Insertions, deletions and nucleotide substitutions have been observed with large dengue cDNA clones, but smaller subgenomic cDNA

clones appear to be stable. The dengue genome can be propagated as two plasmids, a structural region clone containing nucleotides 1 to 2505 and a nonstructural region clone containing nucleotides 2040 to 10712. Site-directed mutagenesis was used to insert a T7 RNA polymerase promoter at the 5' end of the structural region clone, and to engineer unique restriction sites in the nonstructural region clone. *In vitro* ligation and transcription of overlapping clones was used to produce full-length genomic RNA, which was transfected into several different cell lines. Conditions were established for the transfection of viral RNA that allowed recovery of dengue virus and alternative methods for isolating biologically active cDNA clones are discussed.

Materials and methods

Structural and nonstructural region clones: All recombinant plasmids were constructed with cDNA from the PR159 S1 vaccine strain, using standard recombinant DNA techniques (18, 25). Our initial approach for constructing an infectious clone used cDNA clones that were partially or completely sequenced. Several polymorphisms were noted during the original sequencing of the dengue genome and additional polymorphisms were detected during the construction of dengue transcription vectors. Nucleotides 65 and 103 are A residues in pDNC3 and are A residues in all other sequenced dengue viruses. The published PR159 sequence lists positions 65 and 103 as G residues. The polymorphism at position 103 results in an aspartate to asparagine change in the capsid protein, whereas the polymorphism at position 65 occurs in the 5' noncoding region. Asparagine is found at this position in all other sequenced dengue viruses. Nucleotide 4208 is a T residue in pDN2 which changes a threonine to isoleucine in NS2B, destroying a potential glycosylation site. The published PR159 sequence lists position 4208 as a C residue. An additional change in the 5' noncoding region was detected during the engineering of the 5' end, and is discussed below.

pBSK 5427 was constructed using fragments from three different dengue cDNA clones. Clone pDNP7 was digested with *Xmn*I and *Eco*RI to generate a fragment containing nucleotides 10232 to 10712 of the dengue genome. This fragment also contains a 21-nucleotide-long poly (A) tract and an *Eco*RI linker at the 3' end. pDN8 was digested with *Ava*I and *Xmn*I to generate a fragment containing nucleotides 7058 to 10231 of the dengue genome, and pDN5 was digested with *Xho*I and *Ava*I to generate a

fragment containing nucleotides 5427 to 7057 of the dengue genome. These three fragments were ligated together and were then ligated to *Xho*I and *Eco*RI digested Lambda Zap (Stratagene). After packaging, recombinant phage were isolated by plaque hybridization and were converted into plasmids by superinfection with filamentous phage R408. pBSK 5427 contains nucleotides 5427 to 10712 of the dengue genome in vector Bluescript SK (-). A sequence polymorphism at nucleotide 6845 in pDN8 (which causes a frame shift) was avoided by using the *Ava*I site. The *Xmn*I site was chosen to avoid an asparagine to histidine mutation (at nucleotide 10030) present in clone pDNP7. Clone pDN5 contains a valine to leucine polymorphism at position 5815; valine and threonine residues are observed at this position in other flaviviruses.

pBSK 5427 was digested with *Xho*I and treated with bacterial alkaline phosphatase. Clone pDN2 was digested with *Xho*I and a fragment containing nucleotides 3552 to 5426 was ligated to the vector to generate pBSK 3552. pBSK 3552 was digested with *Spe*I and *Pst*I to generate a fragment containing nucleotides 3580 to 10712 of the dengue genome (*Pst*I cuts within the vector polylinker). The insert was ligated to *Spe*I and *Pst*I digested pBSK (-) to generate pBSK 3580. This cloning maneuver flips the transcriptional orientation of the insert, removes several redundant restriction sites, and was necessary to facilitate construction of a complete nonstructural region clone. pBSK 3580 was digested with *Spe*I and *Xba*I and was ligated to a *Spe*I to *Xba*I fragment (from clone pDN2) containing nucleotides 2040 to 3579 of the dengue genome to generate pBSK 2040. pBSK 2040 contains the entire nonstructural region of the dengue genome and is a composite of four

dengue cDNA clones, pDN2, pDN5, pDN8 and pDNP7. A silent sequence polymorphism at nucleotide 4965 (in pDN2) eliminates a *Stu*I site, making the *Stu*I site at nucleotide 4077 unique.

pDNC3 was digested with *Hind*III to generate a fragment containing nucleotides 16 to 1547 of the dengue genome, and pDNC8 was digested with *Hind*III and *Eco*RI to generate a fragment containing nucleotides 1548 to 2505 of the dengue genome. Vector pGEM1 (Promega) was digested with *Hind*III and *Eco*RI and was ligated to the fragments described above to generate pDNC38. This clone contains the entire structural region of the dengue genome with the exception of the 5' terminal 15 nucleotides. pDNC38 is stable in MC1061.1 but is unstable in XL1 Blue (Stratagene) *Escherichia coli* cells. An asparagine to serine polymorphism at nucleotide 959 occurs in pDNC8, but it is not known whether this mutation is present in pDNC3 or pDNC38. The polymorphism is found within the envelope protein gene, and asparagine is found at this position in other dengue viruses.

We tried several different methods to generate a plasmid containing a full-length copy of the dengue genome but all of these attempts failed because the plasmids were unstable in *Escherichia coli*. Since it was clear that the dengue genome was stable in two pieces, we concentrated on developing methods that would allow production of full-length RNA from *in vitro* ligated templates. In order to do so, a *Bst*XI restriction site in the nonstructural clone was eliminated by site-directed mutagenesis and an *Xba*I site was created at the 3' end of the genome. A T7 RNA polymerase promoter and 15 nucleotides were added to the 5' end of pDNC38 to allow

transcription of dengue RNA. These methods and the cloning steps are detailed below.

5' end engineering: pDNC3 was digested with *Eco*RI and *Pst*I and a fragment containing nucleotides 16 to 213 of the dengue genome was cloned into *Eco*RI and *Pst*I digested pGEM1 to yield pGEM16/213. Oligonucleotide FP72 (containing the sequence 5' dTAATACGACTCACTATAGGAGTTGTTAGTCTACGTGGCCCGACAA AGACAGATTCTTTGAGGGAGCTGAGCT) was kinased and gel purified, prior to annealing with a kinased, complementary oligonucleotide, FP5A (containing the sequence 5' dCAGCTCCCTCAAAG). The annealed product was made double stranded using Escherichia coli DNA polymerase Klenow fragment and was digested with *Sac*I. The double stranded DNA oligonucleotide contains a T7 RNA polymerase promoter immediately adjacent to the first 55 nucleotides of the dengue genome (the sequence of the first 15 nucleotides was assumed to be the same as for the dengue 2 Jamaica strain (12), and was inserted into *Sac*I/*Pvu*II digested pGEM16/213, yielding vector p5'213. p5'213 and pGEM4 were digested with *Pst*I and *Nhe*I and polylinker sequences of pGem4 (a *Pst*I to *Nhe*I fragment) were inserted to yield vector p5'L213. Vector p5'L213 contains several unique restriction sites that allow cloning of sequences adjacent to the 5' end of the dengue genome (Chapter 1).

p5'L213 was digested with *Bgl*II and *Eco*RI, and a *Bgl*II- and *Eco*RI-digested fragment from pDNC38 (containing nucleotides 89 to 2505 of the dengue genome), was inserted to yield p5'TOTOA. p5'TOTOA was

digested with *Bgl*II and *Bst*EII, and a *Bgl*II- and *Bst*EII-digested fragment from pDNC2 (containing nucleotides 89 to 1648 of the dengue genome) was inserted to yield p5'TOTOB. p5'TOTOB was constructed to generate an independent clone containing the envelope gene (to minimize any effects of sequence polymorphism at nucleotide 959).

When the 5' ends of p5'TOTOA and pDNC38 were sequenced it was noted that the published sequence for nucleotide 19 was incorrect (18). This nucleotide is an A residue in clones pDNC38 and pDNC2 and is absolutely conserved among the dengue viruses. The incorrect sequence had been incorporated in the design of the 5' end and was corrected by site-directed mutagenesis. The original T7 RNA polymerase promoter was designed for maximum transcription efficiency and included two additional G residues which are not found in dengue RNA. In order to minimize the affect of additional residues on infectivity, one of the additional G residues was removed by site-directed mutagenesis. Two oligonucleotide primers and the polymerase chain reaction (PCR) were used to amplify nucleotides 1 to 2422 of the dengue genome. A minus strand primer (FP20) containing the sequence 5' dCGGCCTGCACCATAACTCCC, and a plus strand primer (FP42GA) containing the sequence 5' dTAATACGACTCACTATAGAGTTGTTAGTCTACGTGGACCG were used to amplify fragments from pools I and III of the original dengue cDNA library (32). The plus strand primer removes an additional G at the transcription start site and changes position 19 to an A. PCR products were kinased, digested with *Bst*XI, and were gel purified using low melting temperature agarose (Seakem). pDNC38 was digested with *Bst*XI

and *Eco*RI and a fragment containing nucleotides 2255 to 2505 was gel purifed. pGEM1 was digested with *Pvu*II and *Eco*RI, and was ligated with the two gel isolated fragments to create 5'TOTO42GA. Several independent clones were isolated and were used as a pool in transcription reactions.

3' end engineering: pBSK 5427 was digested with *Sac*I and *Pst*I and a fragment containing nucleotides 9601 to 10712 was ligated to *Sac*I and *Pst*I digested M13mp18 (replicative form) to yield pM18 9601. Kinased oligonucleotide primers FP15.1 (containing the sequence 5' dCACCACCCTGCTCCCTG) and FP18 (containing the sequence 5' dTTTTTTTTTCTAGAACCTG) were used to mutagenize single strand M18 9601 essentially as described (23). Primer FP15.1 destroys a *Bst*XI site at position 9800 and primer FP18 creates an *Xba*I site at the 3' terminus of the dengue genome. The replicative form of a double mutant was digested with *Sac*I and *Xba*I, and was ligated to *Sac*I and *Xba*I digested pGEM1 to generate pGEM 9603. pBSK 2040 was digested with *Bst*XI and *Sac*I to generate a fragment containing nucleotides 2255 to 9600, and pDN2 was digested with *Eco*RI and *Bst*XI to generate a fragment containing nucleotides 2040 to 2254 of the dengue genome. These fragments were ligated to *Sac*I and *Eco*RI digested pGEM 9603 to yield p3'TOTO. p3'TOTO contains the entire nonstructural region of the dengue genome, and can be ligated (using a unique *Bst*XI site) with a 5'TOTO series clone to generate a full-length genomic template for RNA transcription.

Transcription of RNA from *Xba*I linearized templates produces RNA that contains two additional nucleotides at the 3' terminus. *Xba*I digested 3'TOTO was partially filled in with Klenow, dATP, dCTP, dGTP

and ddTTP, then phenol extracted and ethanol precipitated. DNA was resuspended in S1 nuclease digestion buffer (50 mM sodium acetate pH 4.5, 250 mM sodium chloride, 5 mM zinc chloride) and was treated with 100 units per ml of S1 nuclease (Boehringer Mannheim) for 5 minutes at room temperature. S1 nuclease digestions were phenol extracted, ethanol precipitated and resuspended in restriction endonuclease digestion buffer. After digestion with *Bst*XI, linearized fragments containing nucleotides 2256 to 10,712 of the dengue genome were purified using low melting temperature agarose gels.

Structural and nonstructural region shuttle vectors: Vector p5'213 was digested with *Bgl*II, the 5' overhang was made blunt by Klenow treatment and was then digested with *Hind*III. pBSK 2040 was digested with *Sma*I and *Bam*HI to generate a fragment containing nucleotides 2040 to 10,019, and pGEM 9603 was digested with *Bam*HI and *Hind*III to generate a fragment containing nucleotides 10,020 to 10712 of the dengue genome. Ligation of these two fragments with vector produces p5'3' 2040. *Bst*XI digested PCR or cDNA fragments can be cloned unidirectionally into *Bst*XI digested p5'3' 2040 to generate a library of nonstructural region clones. A three-piece ligation is required to produce a full-length genomic template using this vector. *Bgl*II and *Bst*XI digested PCR or cDNA fragments can be cloned into *Bgl*II and *Bst*XI digested p5'TOTO clones to generate a library of structural region clones.

PCR amplification of dengue cDNA: Independent structural and nonstructural region clones can be assembled using existing cDNA, newly

synthesized cDNA, or by amplifying portions of an existing cDNA library with PCR technology. Several primers were designed to amplify 2 to 3.7 kilobase fragments of dengue cDNA that contained unique restriction sites for directional ligation and cloning. cDNA from pools I and III of the dengue library were mixed and amplified using standard PCR reaction conditions, as recommended by the manufacturer (Cetus). Primers were used at a concentration of 1 nanomole per ml and templates were amplified using a starting mass of 200 nanograms per ml. A standard PCR cycle consisted of a 1 minute, 94° C denaturation step, followed by a 1 minute 50° C hybridization step, and a 2 minute 72° C polymerization step. After 25 cycles, PCR reaction mixes were phenol-extracted, ethanol-precipitated and resuspended in restriction endonuclease digestion buffer. Fragments were gel isolated and ligated to appropriately digested p5'TOTO42GA and p5'3'TOTO vectors to create libraries of structural and nonstructural region clones.

In vitro ligations and transcriptions: 25μg of p5'TOTOA was digested with *Bst*XI and *Eco*RI to generate a fragment containing nucleotides 1 to 2254 of the dengue genome. 50 μg of p3'TOTO was treated as described earlier to generate a fragment containing nucleotides 2255 to 10712 of the dengue genome. The gel isolated fragments were resuspended in a total volume of 25 μl and were ligated using 10 U/ml of T4 DNA ligase (New England Biolabs) for 2 hours at 15°C. Ligated templates were digested with 50 μg per ml of proteinase K (Boehringer Mannheim), and were then phenol-extracted and ethanol-precipitated. Templates were resuspended to a mass of 0.5 μg/ml in diethylpyrocarbonate treated,

ribonuclease free water. Transcription reactions containing 0.5 mM rNTPs (Pharmacia), 0.5 mM cap analog $m^7G(5')ppp(5')G$ or $m^7G(5')ppp(5')A$ (New England Biolabs), 50 ng/µl of template, and 2 units/µl of T7 RNA polymerase (Pharmacia) in 1X transcription buffer were incubated at 37°C for 45 minutes. Quality of transcribed RNA was checked by nondenaturing agarose electrophoresis. Transcription reactions were treated with RNase free DNase I prior to electrophoresis.

Cells and virus stocks: Stocks of the PR159 dengue 2 S1 isolate (3, 4, 33) were prepared on *Aedes albopictus* C6/36 cells essentially as described (9). C6/36 cells were propagated at 30°C in Dulbecco's modified Eagle medium containing 10% fetal calf serum and supplemented with nonessential amino acids. BHK-21 clone 13 cells were obtained from the American Type Culture Collection. BHK-21 clone 15 cells were obtained from Dr. Joel Dalrymple, USAMRIID, Bethesda MD, and were used for plaque purifying virus and RNA transfections. BHK cells were propagated at 37°C in minimum essential medium supplemented with nonessential amino acids and containing 5% fetal calf serum. Plaque assays were performed at 34°C on BHK-21 clone 15 cells using a 1% low melting temperature agarose (Seakem) overlay containing minimum essential medium and 5% fetal calf serum. Plaques were readily visible after 4 days by direct visualization or after staining the monolayer with neutral red.

Preparation of viral RNA and transfection assays: C6/36 cells were infected using a multiplicity of infection of 0.1, and the culture medium was harvested 3 and 6 days after infection. 3×10^9 plaque forming units

were concentrated by precipitation with 8% polyethylene glycol and were resuspended in TNE buffer (50 mM sodium chloride, 100 mM Tris pH 8.6, 1 mM EDTA, and 0.02% w/v bovine serum albumin) Virus was isolated using linear velocity potassium tartrate gradients (31). Gradients were centrifuged for 3 hours at 24,000 rpm using an SW27 rotor. Bands containing virus were diluted tenfold in a proteinase K buffer (20 mM Tris pH 7.5, 1 mM EDTA, 0.2% sodium dodecyl sulfate and 50 μg per ml proteinase K) and were digested for 30 minutes at 50°C. The viral RNA was phenol-extracted, ethanol-precipitated, and stored under 70% ethanol at -20°C. The yield of genomic RNA was estimated by comparison with ethidium bromide stained markers and was 1.5 μg per 1 x10^9 PFU.

5 to 10 ng of dengue viral RNA, 200 ng of *in vitro* transcribed dengue RNA, or 50 ng of *in vitro* transcribed Sindbis virus RNA was diluted with distilled water to a volume 25 μl and was mixed with 25 μl of lipofectin (BRL) (15). Sindbis virus RNA was transcribed from an infectious cDNA clone (TOTO51) essentially as described (30). Transfection mixes were incubated for 10 minutes at room temperature before applying to nearly confluent 35 or 60 mm tissue culture plates of BHK-21 clone 13, or clone 15 cells. Monolayers were washed twice with Opti-MEM (GIBCO-BRL), and were transfected with a 50 μl mixture of RNA and lipofectin in 1 ml of Opti-MEM for one hour at 37°C. Cells were washed once with minimum essential medium containing 5% fetal calf serum, and were either processed for a plaque assay (see above) or were incubated for four days at 34°C. Plaque assays were performed on tissue culture supernatants from solution infections, as described above.

Results and Discussion

Transcription of genome length RNA: We and others have empirically determined that full-length genomes of flaviviruses can't be propagated as plasmids in *Escherichia coli* (29). The cause of plasmid instability is unknown, but presumably the dengue genome contains protein or DNA sequences that are deleterious to the host or plasmid. Structural region clones were stable in MC1061.1 (this was the parental strain for the dengue cDNA library) but underwent deletions and rearrangements when transfected into XL1 Blue. These strains have compatible restriction-modification and recombination systems and the reason for the strain-dependent instability is not understood. Strain-independent rearrangements and deletions were observed when full-length genomes were assembled using plasmid or lambda vectors.

An *in vitro* ligation approach was developed that enabled transcription of full-length genomic RNA from two stable plasmid templates. Our initial approach used cDNA clones (from a randomly primed cDNA library) that had been partially or completely sequenced. Several potentially lethal sequence polymorphisms were avoided by choosing appropriate restriction sites and by constructing independent clones (Fig. 1A). A bacteriophage T7 RNA polymerase promoter was placed at the 5' end of structural region clones to allow transcription of dengue genomic RNA containing one or two additional G residues at the 5' end (Fig. 1B). The first 15 nucleotides of the dengue 2 S1 genome have not been sequenced and was assumed to be the same as the dengue 2 Jamaica strain. The original T7 RNA polymerase promoter was designed for maximal transcription efficiency and transcribed RNA contains 2

Fig.1. Schematic of dengue cDNA clones and modifications used in constructing transcription templates.

Panel A: Individual cDNA clones used for constructing 5' and 3' structural and nonstructural region clones. Individual cDNA clones and restriction sites used for generating structural and nonstructural region clones are aligned with a diagram detailing the genomic organization of the dengue 2 S1 strain. Restriction sites are designated by name and a vertical bar between individual cDNA clones. The single large open reading frame is depicted as an open box and is functionally divided into structural and nonstructural regions. The 5' and 3' untranslated regions are depicted as thin lines protruding from either end of the open box. The relative positions of sequence polymorphisms are displayed by asterisks within individual cDNA clones.

Panel B: 5' and 3' end modifications necessary for production of full-length genomic RNA. The upper half of Fig. 1B depicts the promoter and partial 5' untranslated leader sequences of clones p5'TOTOA and B, and p5'TOTO42GA. The transcription start site of p5'TOTOA and B is indicated by a rightward facing arrow. The T7 RNA polymerase promoter of p5'TOTO42GA and the first nucleotide of the dengue genome are indicated by underlining and a rightward facing arrow, respectively. The lower half of Fig. 1B details the cloning steps necessary to produce run off transcripts identical to dengue genomic RNA. After *Xba*I digestion of p3'TOTO the 5' overhang is partially filled in, then blunted with S1 nuclease. The terminal ddTTP residue is designated in bold type. Full details of the conditions used to engineer the 5' and 3' ends are listed in Materials and Methods.

A.

B.
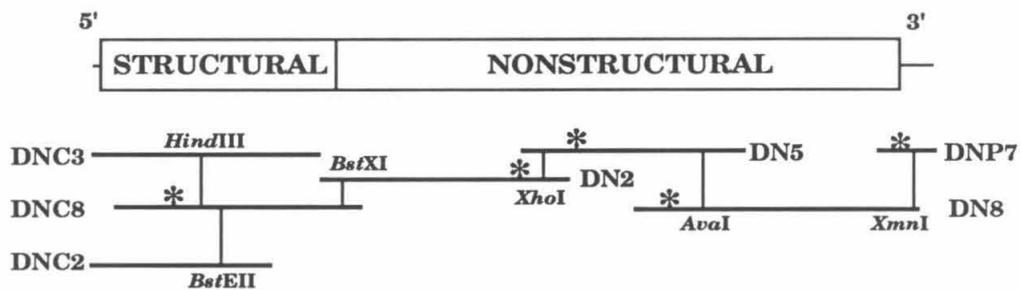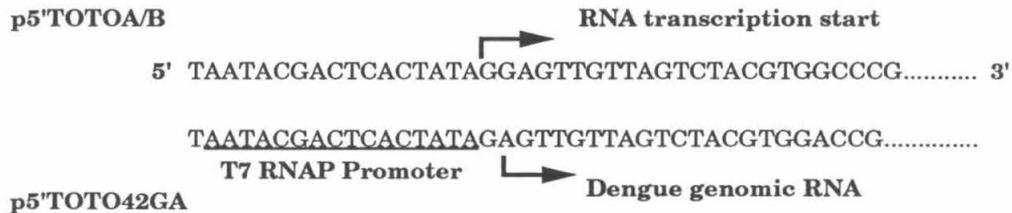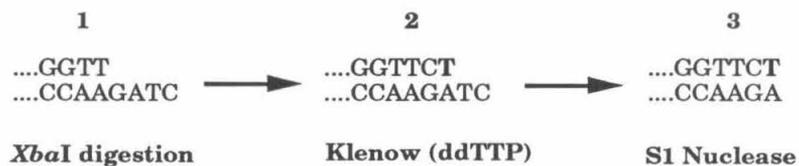
additional G residues at the 5' terminus. A second promoter was designed that removed one of the G residues and corrected the sequence at position 19. Transcription efficiency is approximately one-third that of the 2 G promoter and is decreased 2- to 3-fold when transcripts are initiated with the cap analogue $m^7G(5')ppp(5')A$ (data not shown).

An *Xba*I restriction site was added to the 3' end of a nonstructural clone, and run off transcripts contain 2 additional residues at the 3' terminus. Partial fill in of the 5' overhang with Klenow and ddTTP, followed by S1 nuclease digestion removes the additional residues and run off transcripts terminate with the correct residue. It is not known whether this step is necessary for the production of infectious RNA, as yellow fever RNA transcripts containing extra 3' residues are fully infectious (29).

Transfection of viral and transcribed RNA: A number of p5'TOTO clones were ligated with p3'TOTO and transcribed to produce full-length genomic RNA. A full-length run off transcript of 10.7 kilobases and a run off transcript (from the 5' structural region clone) of 2.25 kilobases were observed, when gel purified templates were used for transcription.(Fig. 2). Virus could not be recovered when transcribed RNAs were transfected into BHK-21 clone 13 or 15 cells. SP6 transcribed Sindbis virus genomic RNA and dengue viral RNA were used as positive controls for transfection (Table 1). Transfection efficiency varied from experiment to experiment and was particularly low for the dengue viral RNA positive control. This may reflect the quality of the RNA preparation or the methods used to quantitate viral and transcribed RNAs. Alternatively the dengue RNA may have a low endogenous transfection efficiency. Apparently our

Fig. 2. *In vitro* ligation and transcription strategy used to produce genome length RNA.

The upper half of Fig. 2 describes the structure of p5' and p3'TOTO clones. The T7 RNA polymerase promoter of a p5'TOTO clone is designated by a rightward facing arrowhead. Nucleotides 1 to 2505 of the dengue genome are designated by a closed box, and the positions of *Bst*XI and *Eco*RI sites used to prepare the 5' structural clone for ligation are indicated. Nucleotides 2506 to 10712 of the dengue genome are designated by an open box and the positions of *Bst*XI and *Xba*I sites used to prepare the 3' nonstructural clone for ligation are indicated. The individual steps necessary to prepare the 5' and 3' clones for ligation are summarized on the left and right sides of the downward facing arrow. Full details are found in the Materials and Methods. The lower half of Fig. 2. describes the *in vitro* transcription methodology. The lower left quadrant of Fig. 2. schematically diagrams the linear structure of a full-length transcription template. The rightward facing arrowhead delineates the 5' end, and the *Xba*I restriction site delineates the 3' end of the dengue genome. The lower right quadrant of Fig. 2. displays a nondenaturing agarose gel containing *in vitro* transcribed RNA and BHK-21 cell RNA. One tenth of a p5'TOTOA/3'TOTO programmed transcription reaction (lane 1) and one µg of BHK-21 total cellular RNA (lane 2) are displayed. *In vitro* transcriptions produce full-length genomic RNA from ligated, and a 2.25 kilobase RNA from unligated p5'TOTOA templates. 28 and 18S ribosomal RNA markers are labeled. Lane λ contains a *Hind*III digest of λ DNA.
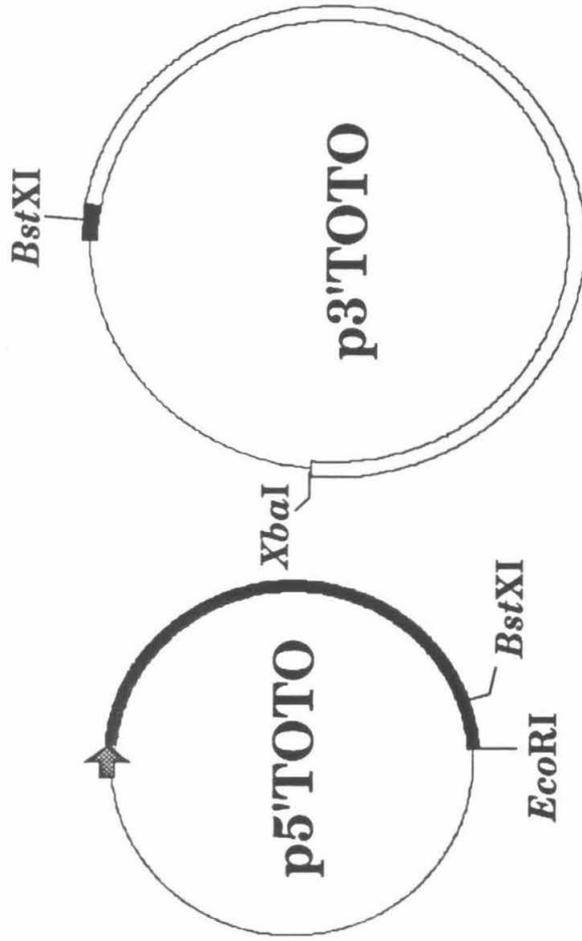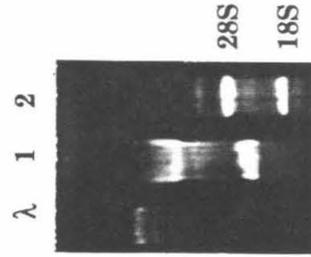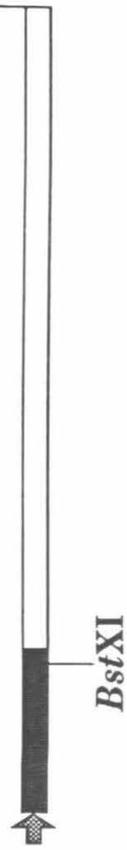
Table 1: Transfection of viral and transcribed RNAs.

RNA source	PFU/µg[a]	PFU/ml[b]
5'TOTOA and 3'TOTO	0	0
5'TOTOB and 3'TOTO	0	0
5'TOTO42GA and 3'TOTO	0	0
TOTO 51	2-5000	N.T.
Dengue 2 RNA	1-5000	20,000-100,000

Footnotes:

a

Transfections and plaque assays were performed as described in Materials and Methods. Viral and transcribed RNAs were quantitated by comparison with ethidium bromide stained standards. BHK-21 clone 13 and clone 15 cell lines were used for transfections.

b

Four days after transfection, tissue culture supernatants were assayed for the presence of dengue virions. BHK-21 clone 13 cells consistently gave lower titers (about fivefold) than clone 15 cells. N.T. indicates not tested.

lipofectin transfection assays can be optimized further to increase the sensitivity of detection. Solution infections of BHK clone 15 cells produced viral titers three orders of magnitude lower than those obtained using mosquito cells. Mosquito cells are a superior cell line for the detection of virus produced by solution transfection and methods need to be adapted to optimize the transfection efficiency with this cell line.

Alternative strategies for isolating infectious cDNA: Our initial attempts at constructing full-length genomic templates used clones that had been completely or partially sequenced. Once it became evident that our cDNA clones did not produce infectious RNA, it became necessary to construct independent p5' and p3'TOTO clones. Newly synthesized cDNA or PCR amplified cDNA from an existing library can be cloned into p5'TOTO or p5'3' 2040 (Fig. 3). All but the 5' terminal 89 nucleotides of p5'TOTO and the 3' terminal 912 nucleotides of p5'3' 2040 are unique to each independent structural and nonstructural clone. Nucleotides 1 to 89 and 9800 to 10,712 of the dengue genome were resequenced (to ensure that no mutations had occurred during cloning), and no nucleotide changes were detected. Several primers were designed to amplify a large portion of the dengue genome using PCR technology. The primers were designed in such a way that the PCR segments could be digested with restriction endonucleases to generate unique 5' or 3' overhangs (to facilitate directional cloning) (Table 2). This increases the probability of a productive ligation event and minimizes the effort needed in screening for independent 3' nonstructural region clones. Libraries of 5' and 3' clones have been constructed, but have not been fully characterized. It is unclear

at this time whether this approach will produce cDNA clones that can be transcribed to produce infectious RNA.

The nature of lethal mutations: A large portion of the cDNA used for the construction of structural and nonstructural region clones was sequenced during the original characterization of the dengue genome and later during construction of transcription vectors. The integrity of the cDNA was tested independently by constructing *trp*E fusions that expressed over 60% of the dengue polyprotein, and by *in vitro* translations expressing all of the dengue polyprotein. All of the *trp*E fusions and *in vitro* translations produced polypeptides of the expected molecular weight, indicating that no in-frame deletions or frameshifts had occurred. Point mutations that alter a critical amino acid are the most likely cause of lethality. These can be relatively conservative changes (a valine for alanine) yet can have severe effects (30). The high error frequency of the viral RNA polymerase and reverse transcriptase used to synthesize cDNA, suggest that screening libraries of ligated 5' and 3' clones for their ability to transcribe infectious RNA would be an efficient method to isolate biologically active cDNA clones.

Fig. 3. An alternative strategy for the construction of independent structural and nonstructural region clones.

Panel A: The upper half of panel A displays a schematic diagram of the dengue genome along with the relative positions of primers and restriction sites used to generate independent 5' and 3'clones. Four plus strand and four minus strand primers are used to amplify four segments of dengue cDNA. Appropriately digested segment 1 and p5'TOTO can be ligated together to generate independent structural region clones. Appropriately digested segments 2, 3, 4 and p5'3'2040 can be ligated together to generate independent nonstructural region clones. Full details of the primers and methods used are found in Table 2 and in the Materials and Methods.

Panel B: Agarose gel displaying dengue cDNA that was amplified using the polymerase chain reaction. Pools I and III of the dengue cDNA library were pooled and amplified with the primers listed in Table 2 and the Materials and Methods. Lanes 1 and 2 contain segment 1 that was amplified using FP72 and FP42GA plus strand primers. Lanes 3, 4 and 5 display segments 2, 4 and 3 respectively. The plus and minus strand primers used to amplify all four segments are listed in Table 2. Lane λ contains a *Hind*III digested λ marker.

A.

p5'TOTO series   p5'3' 2040

B.

PCR Products
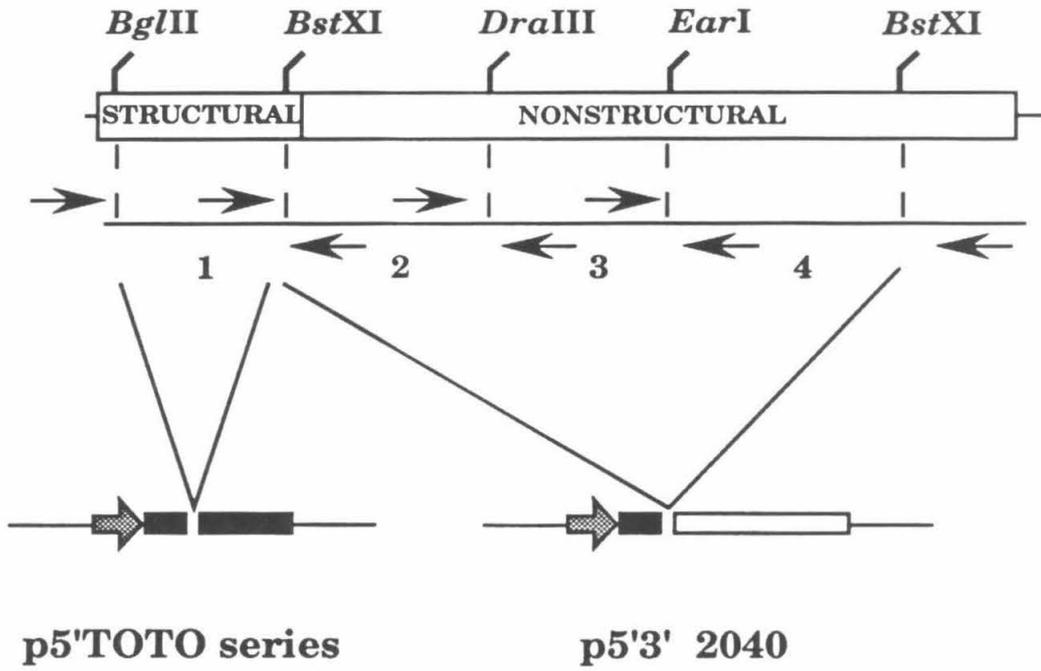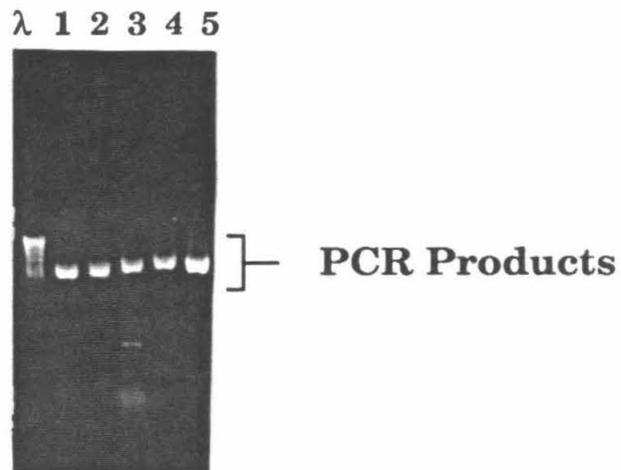
Table 2: PCR amplification of dengue cDNA.

Segment	Primers	Primer Sequence[a]	Size[b]	R.E. Digestion[c]
1	FP42GA or FP72 and FP20	FP20-dcggcctgcaccataactccc	2.3 kB	*Bgl*II and *Bst*XI
2	FP19 and FP5087A	FP19-dtgggcgacacagcctggga	3.0 kB	*Bst*XI and *Dra*III
3	FP4477 and FP4BS	FP4477-dccggtatcaataccaa FP4BS-dttatgcggccaccactgtga	3.0 kB	*Dra*III and *Ear*I
4	FP6366 and FP22	FP6366-datgctaaccctgaacctaatc FP22-dgagccgcaccatggtcttctc	3.7 kB	*Ear*I and *Bst*XI

Footnotes:

[a] The sequences of primers FP72, 42GA and 5087A (Chapter 2) are described in Materials and Methods.

[b] Size of the cDNA segment amplified.

[c] Restriction enzymes used for digestion of PCR fragments.

Acknowledgements

References Cited

1. **Ahlquist, P., R. French, M. Janda, and L. S. Loesch-Fries.** 1984. Multicomponent RNA plant virus infection derived from cloned viral cDNA. Proc. Natl. Acad. Sci. USA **81:**7066-7070.

2. **Bancroft, W. H.** 1987. Current status of dengue vaccines and prospects for the future. Puerto Rico Health Sci. J. **6:**23-6.

3. **Bancroft, W. H., R. M. Scott, K. H. Eckels, C. H. Hoke, T. E. Simms, K. D. T. Jesrani, P. L. Summers, D. R. Dubois, D. Tsoulos, and P. K. Russell.** 1984. Dengue virus type 2 vaccine: Reactogenicity and immunogenicity in soldiers. J. Inf. Dis. **149:**1005-1010.

4. **Bancroft, W. H., F. H. Top, K. H. Eckels, J. H. Anderson, J. M. McCown, and P. K. Russell.** 1981. Dengue-2 vaccine: Virological, immunological, and clinical responses of six yellow fever-immune recipients. Infect. Immun. **31:**698-703.

5. **Barnes, W. J. S., and L. Rosen.** 1974. Fatal hemorrhagic disease and shock associated with primary dengue infection on a Pacific island. Am. J. Trop. Med. Hyg. **23:**495-506.

6. **Bhamarapravati, N.** 1981. Pathology and pathogenesis of DHF, pp. 207-214. In Dengue Hemorrhagic Fever. S. Hotta (ed.), International Center for Medical Research, Kobe University School of Medicine, Kobe, Japan.

7. **Blok, J., S. Samuel, A. j. Gibbs, and U. T. Vitarana.** 1989. Variation of the nucleotide and encoded amino acid sequences of the envelope gene from eight dengue-2 viruses. Arch Virol **105:**39-53.

8. **Cohen, S. N., and S. B. Halstead.** 1966. Shock associated with dengue infection. I. Clinical and physiologic manifestations of dengue hemorrhagic fever in Thailand, 1964. J. Pediatr. **68:**448-456.

9. **Condreay, L. D., R. H. Adams, J. Edwards, and D. T. Brown.** 1988. Effect of actinomycin D and cycloheximide on replication of Sindbis virus in *Aedes albopictus* (mosquito) cells. J. Virol. **62:**2629-2635.

10. **Dasmahapatra, B., R. Dasgupta, K. Saunders, B. Selling, T. Gallagher, and P. Kaesberg.** 1986. Infectious RNA derived by transcription from cloned cDNA copies of the genomic RNA of an insect virus. Proc. Natl. Acad. Sci. USA. **83:**63-66.

11. **Deubel, V., R. M. Kinney, and D. W. Trent.** 1986. Nucleotide sequence and deduced amino acid sequence of dengue type 2 virus, Jamaica genotype. Viorology **155:**365-377.

12. **Deubel, V., R. M. Kinney, and D. W. Trent.** 1988. Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins of dengue type 2 virus, Jamaica genotype: Comparative analysis of the full-length genome. Virology **165:**234-244.

13. **Eckels, K. H., W. E. Brandt, V. R. Harrusin, J. M. McCown, and P. K. Russell.** 1976. Isolation of a temperature-sensitive dengue-2 virus under conditions suitable for vaccine development. Infect. Immun. **14:**1221-1227.

14. **Eckels, K. H., V. R. Harrison, P. L. Summers, and P. K. Russell.** 1980. Dengue-2 vaccine: Preparation from a small-plaque virus clone. Infect. Immun. **27:**175-180.

15. **Felgner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielson.** 1987. Lipofectin: A highly efficient, lipid-mediated DNA transfection procedure. Proc. Natl. Acad. Sci. USA **84:**7413-7417.

16. **Gruenberg, A., W. S. Woo, A. Biedrzycka, and P. J. Wright.** 1988. Partial nucleotide sequence and deduced amino acid sequence of the

structural proteins of dengue virus type 2, New Guinea C and PUO-218 strains. J. Gen. Virol. **69:**1391-1398.

17. **Gubler, D. J.** 1987. Dengue and dengue hemorrhagic fever in the Americas. Puerto Rico Health Sci. J. **6:**107-11.

18. **Hahn, Y. S., R. Galler, T. Hunkapiller, J. Dalrymple, J. H. Strauss, and E. G. Strauss.** 1988. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. Virology **162:**167-180.

19. **Halstead, S. B.** 1988. Pathogenesis of dengue: Challenges to molecular biology. Science **239:**476-81.

20. **Halstead, S. B., J. S. Chow, and N. J. Marchette.** 1973. Immunological enhancement of dengue virus replication. Nature (London) New Biol. **243:**24-26.

21. **Irie, K., P. M. Mohan, R. Putnak, and R. Padmanabhan.** 1989. Sequence analysis of cloned dengue virus type 2 genome (New Guinea C strain). Gene **75:**197-211.

22. **Kaplan, G., J. Lubinski, A. Dasgupta, and V. R. Racaniello.** 1985. *In vitro* synthesis of infectious poliovirus RNA. Proc. Natl. Acad. Sci. USA **82:**8424-8428.

23. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Nat. Acad. Sci. USA **82:**488-492.

24. **Mackow, E., Y. Makino, B. Zhao, Y.-M. Zhang, L. Markoff, A. Buckler-White, M. Guiler, R. M. Chanock, and C.-J. Lai.** 1987. The nucleotide sequence of dengue type 4 virus: Analysis of genes coding for nonstructural proteins. Virology **159:**217-228.

25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

26. **Mason, P. W., P. C. McAda, T. Mason, and M. J. Fournier.** 1987. Sequence of the dengue-1 genome in the region encoding the three structural proteins and the major nonstructural protein NS1. Virology **161:**262-267.

27. **Osatomi, K., and H. Sumiyoshi.** 1990. Complete nucleotide sequence of dengue type 3 virus genome RNA. Virology. in press.

28. **Racaniello, V. R., and D. Baltimore.** 1981. Cloned poliovirus complementary DNA is infectious in mammalian cells. Science **214:**916-919.

29. **Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers.** 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by *in vitro* ligation. New Biol. **1:**285-296.

30. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: Mapping of lethal mutations, rescue of a temperature-sensitive marker, and *in vitro* mutagenesis to generate defined mutants. J. Virol. **61:**3809-3819.

31. **Roehrig, J. T., J. W. Day, and R. M. Kinney.** 1982. Antigenic analysis of the surface glycoproteins of a Venezuelan equine encephalomyelitis virus (TC-83) using monoclonal antibodies. Virology **118:**269-278.

32. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic

amplification of DNA with a thermostable DNA polymerase. Science **239**:487-491.

33. **Scott, R. M., K. H. Eckels, W. H. Bancroft, P. L. Summers, J. M. McCown, J. H. Anderson, and P. K. Russell.** 1983. Dengue 2 vaccine: Dose response in volunteers in relation to yellow fever status. J. Inf. Dis. **148**:1055-1060.

34. **Taniguchi, T., M. Palmieri, and C. Weissmann.** 1978. Qß DNA containing hybrid plasmids giving rise to Qß phage formation in the bacterial host. Nature **274**:223-228.

35. **Trent, D. W., J. A. Grant, L. Rosen, and T. P. Monath.** 1983. Genetic variation among dengue 2 viruses of different geographic origin. Virology **128**:271-285.

36. **Yaegashi, T., V. N. Vakharia, K. Page, Y. Sasaguri, R. Feighny, and R. Padmanabhan.** 1986. Partial sequence analysis of cloned dengue virus type 2 genome. Gene **46**:257-267.

37. **Zhao, R., E. Mackow, A. Buckler-White, L. Markoff, R. M. Chanock, C.-J. Lai, and Y. Makino.** 1986. Cloning full-length Dengue type 4 viral DNA sequences: Analysis of genes coding for structural proteins. Virology **155**:77-88.