

RICE UNIVERSITY

**Specificity of Membrane Helix-helix Interactions  
by Mutagenesis and Structural Analysis**


by


**Endah Susilowati Sulistijo**

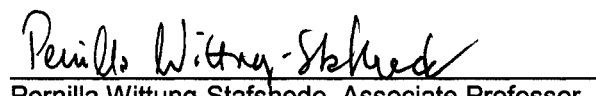
A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

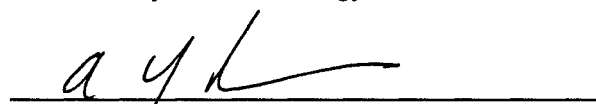
**Doctor of Philosophy**

APPROVED, THESIS COMMITTEE:

  
Kevin R. MacKenzie, Assistant Professor, Chair,  
Biochemistry and Cell Biology

  
Edward P. Nikonowicz, Associate Professor,  
Biochemistry and Cell Biology

  
Pernilla Wittung-Stafshede, Associate Professor,  
Biochemistry and Cell Biology

  
Yousif Shamoo, Associate Professor,  
Biochemistry and Cell Biology

  
Lydia Kavrakli, Noah Harding Professor, Computer  
Science

Houston  
September, 2007

UMI Number: 3309967

Copyright 2007 by  
Sulistijo, Endah Susilowati

All rights reserved.

## INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

**UMI**<sup>®</sup>

---

UMI Microform 3309967

Copyright 2008 by ProQuest LLC.

All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

ProQuest LLC  
789 E. Eisenhower Parkway  
PO Box 1346  
Ann Arbor, MI 48106-1346

Copyright  
Endah Susilowati Sulistijo  
2007

## ABSTRACT

### Specificity of Membrane Helix-Helix Interactions by Mutagenesis and Structural Analysis

by

Endah Susilowati Sulistijo

The activity of apoptosis protein BNIP3 has been associated with its ability to form homodimeric and heteromeric associations through its carboxy-terminal transmembrane domain (TMD), but little is known about the chemical or physical basis of these interactions. In this thesis, I describe two approaches to examine the sequence requirements for BNIP3 TMD dimerization and the properties that drive and stabilize this association. The first approach employs saturation mutagenesis to generate a library of single mutants in the context of a fusion protein construct and SDS-PAGE combined with Western blotting to characterize the mutant dimerization phenotypes. The mutagenesis data maps the BNIP3 TMD dimerization region and identifies five interacting residues for BNIP3 TMD dimerization: Ala176, Gly180, and Gly184 form tandem GxxxG motifs that allow close approach of the helix backbones, and His173 and Ser172 form inter-monomer hydrogen bonds. The mutagenesis data also show that the sequence context in which these five critical residues are embedded affects the strength of TMD helix-helix interactions because several mutations at or near the dimer interface that leave the small residues and the inter-monomer hydrogen bonding intact can abolish or profoundly lower dimerization.



The second approach to the study of BNIP3 TMD dimerization involves determining the structure of the TMD dimer using solution NMR spectroscopy. Reconstitution of the BNIP3 TMD peptide in the non-ionic detergent dodecylphosphocholine (DPC) and addition of dipalmitoylphosphatidylcholine (DPPC) yields peptide spectra with excellent peak dispersion and resolution. This quality enables collection of chemical shift, *J* coupling, and NOE distance restraints, thus allowing determination of the BNIP3 TMD dimer structure. The NMR structure of the BNIP3 TMD dimer reveals the details of how the elements of the BNIP3 TMD sequence cooperate to support dimerization and provides a context to interpret the effects of the saturation mutagenesis results. Results from the saturation mutagenesis and structural analyses establish an understanding of the BNIP3 TMD dimerization and provide a framework for further studies of BNIP3, which include but are not limited to thermodynamic studies, functional analyses, and molecular dynamics modeling of the BNIP3 TMD associations.

# Acknowledgements

The work that leads to the completion of this dissertation is only possible with the assistance, suggestions, and constructive criticisms from talented and intelligent scientists at Rice: faculty, graduate students, postdoctoral fellows, technicians, and undergraduate students. I am very grateful for the opportunity to interact and share experience with them during my career at Rice. At the risk of offending many people, I would like to single out certain individuals for special thanks.

First and foremost, I would like to thank my advisor Kevin MacKenzie who has been the source of inspiration, advice, and motivation through out my career in his laboratory. His knowledge has been a never ending source of information, and it is not an overstatement to call him a walking encyclopedia. I thank him for being patience with me during the ups and downs, for helping me to solve any problems, either scientific or non-scientific, and for believing that it is possible to solve an NMR structure of BNIP3 transmembrane domain dimer.

I thank Dr. JS Olson for being a father figure, lending an ear when needed, giving advice on how to deal when problems come, and making sure that I know how to handle situations properly. I appreciate that he keeps nagging me with job seeking and thesis writing, so that everything goes in place when finishing time comes.

My thesis committee, Drs. Ed Nikonowicz, Yousif Shamoo, Pernilla Wittung-Stafshede, and Jordan Konisky, have been sources of reasons. They make sure that

my project does not deviate from the main goal, so that I am able to finish in a timely manner, and for that I am very grateful.

Dr. Sean Moran has been a great source of advice and assistance for work with the NMR spectrometers. I appreciate his help and suggestions during my time at Rice.

I thank my co-workers in the MacKenzie lab: Ian Dews, Todd Jaszewski, Adina Maximciuc, Mehul Joshi, Chris Rodriguez, and past and present undergraduate students: Erin Wittman, Yasuko Takeda, Mylinh Duong, and Charles Lawrie, for their assistance, suggestions, and constructive criticism. The spectrum of personalities that they have shown has made the lab quite amusing.

I also thank past and present members of the Shamoo lab: John Bruning, Jeff Myers, Siyang Sun, Rafael Couñago, Gang Wu, Megan Guelker, Matt Peña, and Kasia Walkiewicz for help, suggestions and discussion, and for keeping our lab meeting enjoyable.

Essentially everyone at Keck Hall has been very supportive and have made working environment very pleasant. Members of the Olson lab, especially Angela Hvitved, George Blouin, Philip Graves, Todd Molan, Jayashree Soman, and Eileen Singleton, members of the Tao lab, members of the Matthews lab, members of the Nikonowicz lab, members of the Pernilla lab, Izabela Tworowska, Faiza Hussain, and the Keck staffs, Lisa Blinn, Karen Ethun, Deb Miller, and Melissa Glueck have been such good friends and I enjoy the opportunities to interact and share experiences with them both scientifically and non-scientifically.

A special thank for Corey Wilson for being the sound of reasons and Catherine McCollum my comrade in this battle. (Catherine, the light is finally here!!)

Last but not least, I thank my family, my late father, Mohamad Sulistijo, my mother, Sarojini, my brother, Susilo Utomo, and my sister, Retno Puspitosari, for giving me love, support, guidance, and the opportunity to pursue my goals. None of my success can be achieved without their confidence and belief in me. I dedicate this volume to you.

# Table of Contents

	Page
Abstract	iii
Acknowledgements	v
Table of Contents	viii
List of Figures	xiii
List of Tables	xvii
Terms, Symbols, and Abbreviations	xviii
1 Introduction	1
1.1 Apoptosis plays important roles in development and disease	1
1.2 Bcl-2 superfamily of apoptosis proteins regulates mitochondria-dependent apoptosis pathway	2
1.3 The multidomain Bcl-2 and Bax subfamily are directly involved in the initiation of mitochondria-dependent apoptosis pathway	4
1.4 BH3-only proteins may inhibit anti-apoptosis proteins or activate pro-apoptosis proteins	7
1.5 BNIP3 is a member of the Bcl-2 superfamily of apoptosis proteins	8
1.6 BNIP3 may activate apoptosis through multiple mechanisms	11
1.7 BNIP3 interactions through the transmembrane domain are implicated in apoptosis	13
1.8 Rationale	14
1.9 Outline of the thesis contents	15
2 Sample preparation for analysis of BNIP3 transmembrane domain dimerization in detergent micelles	18
2.1 Introduction	18
2.2 Optimization of the SN-BNIP3TMD fusion protein overexpression	19
2.3 Optimization of detection method for SDS-PAGE analysis	21
2.4 Experimental methods for SDS-PAGE analysis	22
2.4.1 Construction of plasmid bearing gene encoding SN-BNIP3TMD fusion protein	22
2.4.2 Generation of single-point mutation library of SN-BNIP3TMD fusion protein	22

2.4.3	Small scale expression and preparation of SN-BNIP3TMD fusion proteins	23
2.4.4	SDS-PAGE analysis and Western blotting	24
2.5	Experimental methods for circular dichroism and Fourier transform infrared spectroscopy	25
2.5.1	Large scale expression and purification of the SN-BNIP3TMD fusion protein	25
2.5.2	Sample preparation for circular dichroism and Fourier transform infrared spectroscopy	27
2.5.3	IR spectra collection and analysis	27
2.5.4	Circular dichroism spectroscopy	28
3	Analysis of BNIP3 transmembrane domain dimerization in detergent micelles	29
3.1	The BNIP3 transmembrane domain is sufficient for dimerization	29
3.2	BNIP3 transmembrane domain dimerization in SDS micelles and in <i>E. coli</i> membranes are stabilized by the same interactions	31
3.3	BNIP3 transmembrane domain dimerization is reversible and not kinetically trapped	34
3.4	Dissociation constant for BNIP3 transmembrane domain dimerization is influenced by the chosen analytical conditions	36
3.5	Substitutions in the BNIP3 transmembrane domain cause a range of dimerization phenotypes	39
3.6	Different substituting residues have different effects on BNIP3 transmembrane domain dimerization	42
3.7	Positional effects of mutations implicate six residues in BNIP3 transmembrane domain dimerization	48
3.8	Mutations to proline map the transmembrane domain region important for dimerization	52
3.9	The carboxy-terminal tail of BNIP3 is not critical for dimerization	55
3.10	Hydrophobic substitutions identify the BNIP3 transmembrane domain dimer interface	57
3.11	Small residues Ala176, Gly180, and Gly184 promote close helix-helix packing through a tandem GxxxG motif	59
3.12	Most large aliphatic residues flanking the tandem GxxxG motif are not important for dimerization	60

3.13	His173 and Ser172 participate in the BNIP3 transmembrane domain dimer stability by making intermonomer hydrogen bonds	60
3.14	Interactions that stabilize BNIP3 transmembrane domain dimerization are inter-dependent	65
3.15	Disruption at non-interfacial residues correlates with changes in hydrophobicity	66
3.16	Fourier transform infra-red spectroscopy data suggest that mutations to polar residue cause dimer disruption due to differences in solvent exposure	68
4	Preparation of the BNIP3 transmembrane domain peptide for solution NMR structure determination	74
4.1	Introduction	74
4.2	Optimization of BNIP3 transmembrane domain peptide length	75
4.3	Optimization of protein overexpression	75
4.4	Optimization of proteolysis conditions	76
4.5	Optimization of NMR sample preparation	79
4.6	Preparation of BNIP3 transmembrane domain peptide for NMR spectroscopy	81
4.6.1	Overexpression of isotopically labeled SN-BNIP3TMD fusion protein	81
4.6.2	Extraction and purification of the SN-BNIP3TMD fusion protein	82
4.6.3	Isolation and purification of the BNIP3 transmembrane domain peptide	83
4.6.4	NMR sample preparation	84
4.6.5	Preparation of phospholipid stocks	85
5	Determination of the BNIP3 transmembrane domain dimer structure by solution NMR spectroscopy	86
5.1	Introduction	86
5.2	Initial assessment of sample and experimental conditions for structure determination of the BNIP3 transmembrane domain	87
5.3	Added phospholipid affects BNIP3 transmembrane peptide chemical shifts in detergent micelles	90
5.4	NMR data collection	95
5.4.1	Sequential assignment of the BNIP3 transmembrane domain peptide backbone atoms	96
5.4.2	Assignment of the side chain chemical shifts	101

5.4.3	Determination of backbone dihedral angles	107
5.4.4	Determination of side chain torsion angles	109
5.4.5	Assignment of intra- and inter-monomer distance restraints	117
5.5	Calculation of the BNIP3 transmembrane domain dimer structure	122
5.5.1	Preliminary structure calculation	123
5.5.2	Initial structure refinement	125
5.5.3	Final structure refinement	128
5.6	Different phospholipids can also modulate structure or stability of the BNIP3 transmembrane dimer	131
6	Interpretation of the BNIP3 transmembrane domain dimer structure	136
6.1	The NMR structure identifies the same interfacial residues as those predicted from saturation mutagenesis	136
6.2	Small residues provide a context for close packing between BNIP3 transmembrane domain helices	138
6.3	BNIP3 TMD dimer exhibits a ridges-to-grooves helix-helix packing	141
6.4	The AxxxGxxxG motif promotes a network of inter-monomer backbone-to-backbone $C^{\alpha}-H^{\alpha}\cdots O=C$ hydrogen bonds	144
6.5	The structure of BNIP3 transmembrane domain dimer confirms inter-monomer interaction between Ser172 and His173 side chains	148
6.6	Leu169 is at the dimer interface but contributes minimally to the stability of the BNIP3 transmembrane domain dimer	150
6.7	The structure is not sufficient explain the importance of Ile183 for BNIP3 transmembrane domain dimerization	151
6.8	Comparison of BNIP3 and GpA transmembrane domain dimer structures	152
6.9	Comparison with a published BNIP3 transmembrane domain dimer structure	158
7	Summary and future work	163
7.1	Combining mutagenesis with structure determination reveals the basis for the specificity of membrane helix-helix associations	163
7.2	Findings from saturation mutagenesis provide a framework for further studies of BNIP3 transmembrane domain dimerization	164
7.2.1	Analysis of BNIP3 transmembrane domain dimerization in real membranes	164



7.2.2	Thermodynamic studies of BNIP3 transmembrane domain dimerization	165
7.2.3	Functional studies of BNIP3 transmembrane domain dimerization	166
7.3	Results from solution NMR spectroscopy call for additional analyses	167
7.3.1	Determination of BNIP3 transmembrane domain peptide-phospholipid contacts	167
7.3.2	Study of BNIP3 transmembrane domain dimer dynamic	168
7.4	Conclusion	169
Bibliography		170
Appendix A	Oligonucleotide primers	184
Appendix B	Glycerol- and urea-SDS-PAGE	189
Appendix C	List of assigned chemical shifts	193
Appendix D	Lists of experimentally derived restraints	203
Appendix E	Modified ARIA 1.2 code	219

# List of Figures

		Page
Figure 1.1	Two major apoptosis pathways in mammalian cells	3
Figure 1.2	The Bcl-2 superfamily of apoptosis proteins	5
Figure 1.3	A model for apoptosis regulation by members of the Bcl-2 superfamily of proteins	6
Figure 1.4	Alignment of the transmembrane domains of BNIP3, Nix, ceBNIP3, and deBNIP3	10
Figure 1.5	Possible pathways for BNIP3 apoptosis activity	12
Figure 3.1	The SN-BNIP3TMD fusion protein dimerizes on SDS-PAGE	30
Figure 3.2	Quantification of CAT expression levels reports the strength of transmembrane domain interactions in <i>E. coli</i> membranes	32
Figure 3.3	Limited saturation mutagenesis data indicates that the same interactions stabilize BNIP3 transmembrane domain dimerization on SDS-PAGE and in <i>E. coli</i> membranes	33
Figure 3.4	Mixing the SN-BNIP3TMD fusion protein with the BNIP3 transmembrane domain peptide generates a protein-peptide heterodimer	35
Figure 3.5	Serial dilutions of purified wild type and His173Trp mutant of the SN-BNIP3TMD fusion proteins	37
Figure 3.6	Serial two-fold dilution of Leu169Phe and Ile183Leu mutants of the SN-BNIP3TMD fusion proteins in glycerol- and urea-SDS-PAGE	38
Figure 3.7	Representative of dimerization phenotypes caused by substitutions in the BNIP3 transmembrane domain	41
Figure 3.8	The library of BNIP3 transmembrane domain single substitution mutants and their corresponding phenotypes	43

Figure 3.9	Positional dependence of disruption caused by substitution mutations	49
Figure 3.10	Projection of the average disruptions on BNIP3 transmembrane domain helix	51
Figure 3.11	Phenotypes for Pro scanning substitutions in the BNIP3 transmembrane domain sequence	54
Figure 3.12	Single substitution mutations at residues 185-188 and their corresponding phenotypes	56
Figure 3.13	Substitution mutations at residues Ser172 and His173 to polar amino acids and their corresponding phenotypes	62
Figure 3.14	Ser172 mutants and His173 mutants do not form heterodimer in SDS micelles	64
Figure 3.15	Correlations between disruptive effects of non-interfacial mutations and loss of hydrophobicity	69
Figure 3.16	Circular dichroism (CD) spectra of wild type and selected mutants of the SN-BNIP3TMD fusion protein	71
Figure 3.17	Original and difference FTIR spectra for wild type and selected mutants of SN-BNIP3TMD	72
Figure 4.1	Sequences of various Trypsinolysis products of SN-BNIP3TMD	78
Figure 5.1	$^1\text{H}/^{15}\text{N}$ -HSQC temperature series of the BNIP3 transmembrane domain in 5% dodecylphosphocholine	89
Figure 5.2	Effects of added dipalmitoylphosphatidylcholine (DPPC) on peak dispersion and resolution on the $^1\text{H}/^{13}\text{C}$ -HSQC spectrum at the methyl region	91
Figure 5.3	Effects of dipalmitoylphosphatidylcholine (DPPC) addition on peak dispersion and resolution on the $^1\text{H}/^{15}\text{N}$ -HSQC spectrum	94
Figure 5.4	The $^1\text{H}/^{15}\text{N}$ -HSQC spectrum of BNIP3 transmembrane domain peptide prepared in 128 mM DPC, 10 mM DPPC, 10 mM phosphate, pH 5.10	97
Figure 5.5	Slices from 3D HNCA spectrum of the BNIP3 transmembrane domain	98

Figure 5.6	Assigned resonances of the BNIP3 transmembrane domain $^1\text{H}/^{15}\text{N}$ -HSQC	100
Figure 5.7	$\text{H}^\alpha/\text{C}^\alpha$ region of the $^1\text{H}/^{13}\text{C}$ -HSQC of the BNIP3 transmembrane domain	102
Figure 5.8	2D spectra for the aromatic residues of the BNIP3 transmembrane domain	105
Figure 5.9	CSI analysis of the chemical shifts of backbone resonances identifies $\alpha$ -helical region of the BNIP3 transmembrane domain	108
Figure 5.10	Reference and spin-echo difference spectra used to determine coupling between $\gamma$ methyl C and backbone N or C	111
Figure 5.11	Reference and spin-echo difference spectra used to determine the couplings of isoleucine and leucine $\text{C}^\delta$ methyls to their $\text{C}^\alpha$	115
Figure 5.12	Slices from 3D NOESY-HSQC and half-filter NOESY-HSQC	118
Figure 5.13	The number of NOE constraints per residue	120
Figure 5.14	Summary of NOE correlation data of the BNIP3 transmembrane domain peptide	121
Figure 5.15	Initial structure ensemble of the BNIP3 transmembrane domain dimer	124
Figure 5.16	Structure ensemble of the BNIP3 transmembrane domain dimer after initial refinement	126
Figure 5.17	Final structure ensemble of the BNIP3 transmembrane domain dimer	129
Figure 5.18	The effects of adding DPPC or DPPA to a sample of BNIP3 transmembrane domain peptide	132
Figure 5.19	$^1\text{H}/^{13}\text{C}$ -HSQC spectra acquired with samples prepared in 10 mM DPPC and 10 mM DPPA	133
Figure 5.20	$^1\text{H}/^{15}\text{N}$ -HSQC spectra of BNIP3 transmembrane domain peptide prepared in the presence of POPA or <i>E. coli</i> lipids	135
Figure 6.1	Structure of the BNIP3 transmembrane domain dimer	137

Figure 6.2	Close up view of the BNIP3 transmembrane domain dimer interface at the region for the tandem GxxxG motif	139
Figure 6.3	Ridges-to-grooves helix-packing between the monomers of the BNIP3 transmembrane domain dimer	143
Figure 6.4	The network of inter-monomer $C^{\alpha}-H^{\alpha}\cdots O=C$ hydrogen bonds in BNIP3 transmembrane domain dimer	145
Figure 6.5	His173 and Ser172 form inter-monomer hydrogen bonds	149
Figure 6.6	Structure of the GpA transmembrane domain dimer	153
Figure 6.7	Alignment of the BNIP3 and GpA transmembrane domains interacting residues	154
Figure 6.8	Superposition of BNIP3 and GpA transmembrane domain structures when the BNIP3 A176xxxG180 motif is aligned with the GpA GxxxG motif	155
Figure 6.9	Superposition of BNIP3 and GpA transmembrane domain structures when the BNIP3 G180xxxG184 motif is aligned with the GpA GxxxG motif	156
Figure 6.10	The structure of the BNIP3 transmembrane domain dimer solved by the Arseniev's group	159
Figure 6.11	Superposition of BNIP3 transmembrane domain dimer structure with the structure solved by the Arseniev's group	160

## List of Tables

		Page
Table 3.1	Phenotype distributions for each substituting residue and for pools	45
Table 3.2	Phenotype distributions for each substituting residue compared with pools	46
Table 3.3	Correlation coefficients for regression analyses between phenotype for a given substituting residue and the pool phenotype means at each position along the sequence	58
Table 3.4	Correlation between hydrophobicity change and phenotype for non-interfacial mutations within the sensitive region	67
Table 5.1	List of backbone atom chemical shifts for the BNIP3 transmembrane domain peptide obtained from HNCA, CBCA(CO)NNH, and HNCO spectra	103
Table 5.2	$^3J_{\text{HNH}\alpha}$ coupling for residues at the amino terminus of the BNIP3 transmembrane domain peptide	110
Table 5.3	Three bond heteronuclear $J$ couplings between $\gamma$ carbons and backbone C and N nuclei	112
Table 5.4	$^3J_{\text{CC}\alpha}$ couplings between Leucine and Isoleucine $\delta$ methyls and their $\text{C}^\alpha$ carbons	116
Table 5.5	NMR structural statistics and atomic RMS differences	130
Table 6.1	Geometry of inter-monomer $\text{C}^\alpha\text{--H}^\alpha\cdots\text{O}=\text{C}$ contacts in the BNIP3 transmembrane domain dimer	146

## Terms, Symbols, and Abbreviations

2D	two dimensional
3D	three dimensional
$^3J_{\text{HNH}\alpha}$	$\text{H}^{\text{N}}\text{-H}^{\alpha}$ through bond coupling
$^3J_{\text{CC}\alpha}$	aliphatic- $\alpha$ $^{13}\text{C}\text{-}^{13}\text{C}$ through bond coupling
$^3J_{\text{CC}}$	aliphatic-carbonyl $^{13}\text{C}\text{-}^{13}\text{C}$ through bond coupling
$^3J_{\text{CN}}$	$^{13}\text{C}\text{-}^{15}\text{N}$ through bond coupling
Å	Angstrom, $10^{-10}$ m
ARIA	Ambiguous Restraints for Iterative Assignment
BH	Bcl-2 homology
BNIP3	Bcl2/nineteen-kDa protein interacting protein 3
C	backbone carbonyl carbon
$\text{C}^{\alpha}$	$\alpha$ carbon
CAT	chloramphenicol acetyltransferase
CBCA(CO)NNH	3D NMR experiment that correlates $\text{H}^{\text{N}}$ , N, $\text{C}^{\alpha}$ , $\text{C}^{\beta}$ resonances
CD	circular dichroism
COSY	correlation spectroscopy
CSI	chemical shift index
CT-HSQC	constant time heteronuclear single quantum correlation
cyt c	cytochrome c
DMPC	dimyristoylphosphatidylcholine
DPC	dodecylphosphocholine
DPPA	dipalmytoylphosphatidic acid
DPPC	dipalmytoylphosphatidylcholine
EDTA	ethylenediaminetetraacetic acid
<i>E. coli</i>	<i>Escherichia coli</i>
FRET	fluorescence resonance energy transfer
FTIR	Fourier transform infrared
GpA	glycophorin A

HFIP	hexafluoro isopropanol
H <sup>α</sup>	α proton
H <sup>N</sup>	amide proton
HNCA	3D NMR experiment that correlates H <sup>N</sup> , N, and C <sup>α</sup> resonances
HNCO	3D NMR experiment that correlates H <sup>N</sup> , N, and C resonances
HPLC	high pressure liquid chromatography
HSQC	heteronuclear single-quantum coherence
Hz	hertz; unit of frequency in cycles per second
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kiloDaltons; unit of mass equivalent to 1000 atomic mass unit
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy
MHz	megahertz; 10 <sup>6</sup> cycles per second
N	amide nitrogen
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
NOESY	nuclear overhauser enhancement spectroscopy
OD <sub>600</sub>	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PC	phosphatidylcholine
PMSF	phenylmethylsulfonylfluoride
ppm	part per million
PTP	permeability transition pore
RMS	root mean square
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SN	<i>Staphylococcal</i> nuclease
SN-BNIP3TMD	<i>Staphylococcal</i> nuclease-BNIP3 transmembrane domain fusion protein
TFAA	trifluoroacetic acid
TFE	trifluoro ethanol
TMD	transmembrane domain
TNFR1	tumor necrosis factor receptor 1



TSP	3-(trimethylsilyl)- propionic acid-d <sub>4</sub> , sodium
v/v	volume to volume ratio
w/v	weight to volume ratio

# **1 Introduction**

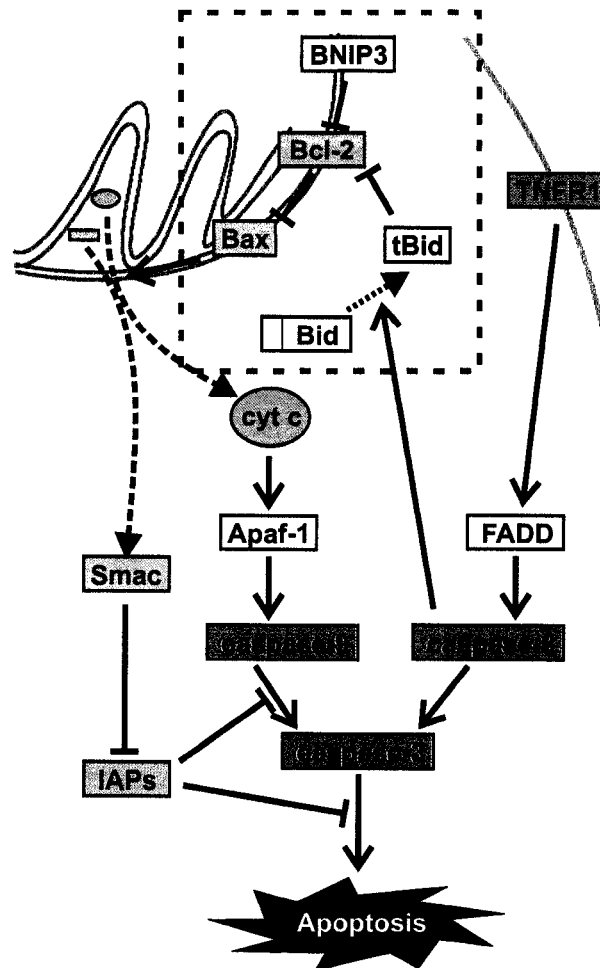
## **1.1 Apoptosis plays important roles in development and disease**

Apoptosis, or programmed cell death, is an essential mechanism for normal development and tissue homeostasis, and it is involved in sculpturing organs, deleting unwanted structures, adjusting cell numbers, and eliminating damaged, abnormal, or misplaced cells (Vaux and Korsmeyer 1999; Meier et al. 2000; Cory and Adams 2002; Opferman and Korsmeyer 2003). Apoptosis is also implicated in defense mechanism. By destroying infected cells and depriving viruses or pathogens of host cells, apoptosis ensures that the spread of infection is limited. Failure in apoptosis regulation has been associated to multiple diseases, ranging from autoimmune and degenerative diseases to cancer (Cory et al. 1994; Thompson 1995; Strasser et al. 1997; Yuan and Yankner 2000; Green and Evan 2002; Kirkin et al. 2004). Because elucidating the regulation and the mechanism of apoptosis can provide insight into development of strategies to treat these illnesses, the study of apoptosis has become one of the fastest growing area in biomedical research (Nicholson 2000; Reed 2001; Reed 2002).

## **1.2 Bcl-2 superfamily of apoptosis proteins regulates mitochondria-dependent apoptosis pathway**

The mechanism that leads to apoptosis can be considered as a complex and tightly controlled network because it implicates many regulatory proteins at both the initiation and the execution stages. Apoptosis can be activated through various pathways and signaling cascades, and there are two major apoptosis pathways in mammalian cells: 1) the death-receptor pathway, which involves members of the death-receptor superfamily, including tumor necrosis factor receptor 1 (TNFR1); 2) the mitochondria-dependent pathway, which involves members of the Bcl-2 superfamily (Figure 1.1) (Hengartner 2000). Despite implicating different regulators, both pathways lead to an activation of a series of caspases, the machineries that are responsible for the breaking down of organelles, and thus cell death.

The mitochondria-dependent pathway is initiated by the release of apoptosis factors cyt c (cytochrome c) and Smac (second mitochondria-derived activator of caspases) from the mitochondria. These factors then activate downstream regulatory proteins, which regulate the action of caspases and other effectors of apoptosis (Figure 1.1) (Adams and Cory 1998; Eskes et al. 1998; Pastorino et al. 1998; Finucane et al. 1999; Gross et al. 1999; Antonsson and Martinou 2000; Hengartner 2000). In this pathway, the release of cyt c and Smac is regulated by proteins that belong to the Bcl-2 superfamily of apoptosis proteins, which include pro-apoptotic (Bax, BNIP3, Bid) and anti-apoptotic members (Bcl-2, Bcl-x<sub>L</sub>). Members of this superfamily of proteins are characterized by four short conserved sequence regions called the Bcl-2 homology (BH) domains, and many possess a carboxy-terminal transmembrane domain (TMD),



**Figure 1.1 Two major apoptosis pathways in mammalian cells**

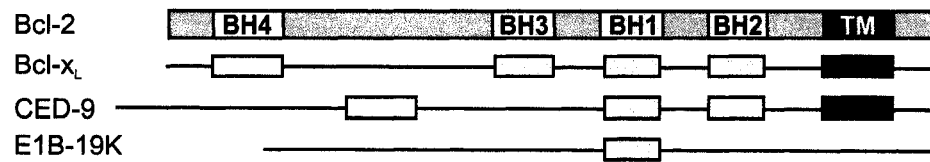
Apoptosis in mammalian cells can be initiated through the death-receptor pathway, which involves members of the death-receptor superfamily, such as TNFR1, or the mitochondria-dependent pathway, which involves members of the Bcl-2 superfamily (boxed). Both pathways activate a series of proteolytic caspases, which are the effectors of cell death.

which allow these proteins to be targeted to the mitochondrial membrane or other organelles (Reed 1997; Adams and Cory 1998; Kelekar and Thompson 1998; Gross et al. 1999; Antonsson and Martinou 2000). Based on structural similarities and functional criteria, these proteins can be classified further into three sub-categories: the anti-apoptosis, multi-domain Bcl-2 subfamily; the pro-apoptosis, multi-domain Bax subfamily; and the pro-apoptosis BH3-only subfamily (Figure 1.2) (Adams and Cory 1998).

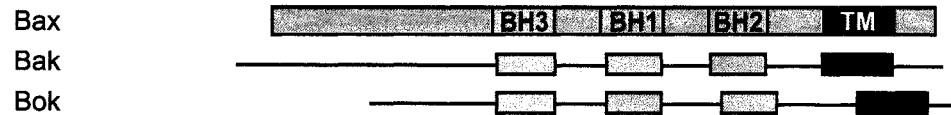
### **1.3 The multidomain Bcl-2 and Bax subfamily are directly involved in the initiation of mitochondria-dependent apoptosis pathway**

The multi-domain Bcl-2 and Bax subfamily proteins share similar structural fold with the pore-forming subunit of diphtheria toxin and bacterial colicins, prompting the speculation that these proteins may directly regulate the permeability of mitochondrial membranes (Muchmore et al. 1996; Schendel et al. 1998; Chou et al. 1999; McDonnell et al. 1999; Fesik 2000; Suzuki et al. 2000; Hinds et al. 2003). However, only members of the pro-apoptosis Bax subfamily, such as Bax and Bak, are associated with the release of cyt c and other apoptogenic proteins. Therefore, it is postulated that Bax and Bak during apoptosis initiation, these proteins form channels that provide a passage for the release of apoptosis factors (Figure 1.3). Other studies suggest that these proteins cause the release of apoptosis factors by perturbing the integrity of the mitochondrial permeability transition pore (PTP) (Korsmeyer et al. 2000). Either action can be antagonized by the anti-apoptosis members of the Bcl-2 superfamily, such as Bcl-2 and Bcl-x<sub>L</sub>, which form interactions with the pro-apoptosis protein through the BH3 domain, thus preventing these proteins from forming channels (Sattler et al. 1997).

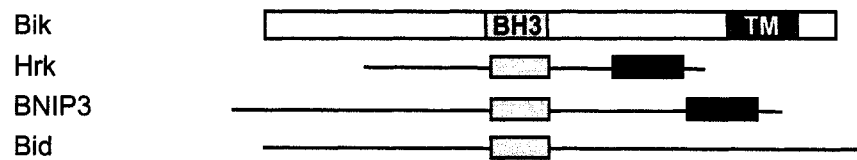
### Anti-apoptosis: Bcl-2 subfamily



### Pro-apoptosis: Bax subfamily

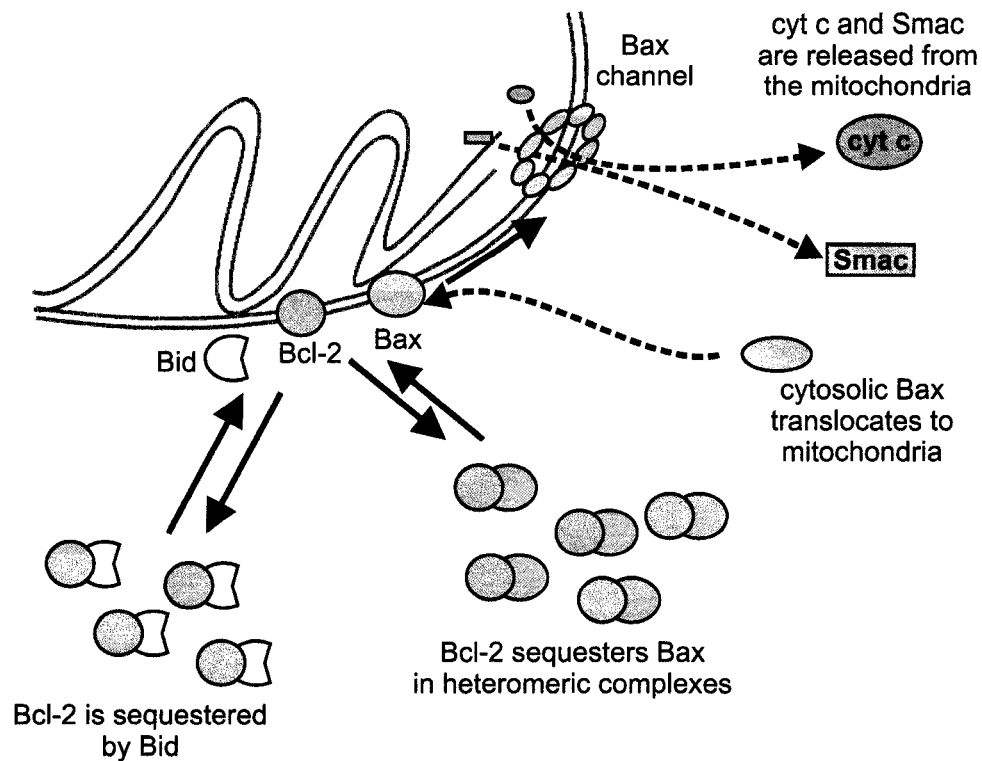


### Pro-apoptosis: BH3-only subfamily



**Figure 1.2 The Bcl-2 superfamily of apoptosis proteins**

Members of the Bcl-2 superfamily of proteins share sequence homology through the BH1-BH4 domains, and many have a carboxy-terminal transmembrane domain, which allow them to transport and anchor to the mitochondria or other cellular organelles. Based on structural similarities and functional criteria, these proteins can be classified into three subfamilies.



**Figure 1.3 A model for apoptosis regulation by members of the Bcl-2 superfamily of proteins**

Upon signal for apoptosis, cytosolic Bax translocates to the mitochondria where it form channels or perturbs the mitochondria membrane permeability transition pore and causes the release of apoptosis factors cyt c and Smac, which then activate downstream effectors of apoptosis. The release of apoptosis factors can be inhibited by Bcl-2 binding to Bax, but this inhibition can be prevented if Bcl-2 interacts with Bid or other BH3-only proteins. Thus, cellular levels of members of the Bcl-2 superfamily can determine cell fate.

Given the extensive structural similarities between the pro- and anti-apoptosis multi-domain members of the Bcl-2 superfamily, it remains unclear why some members can promote apoptosis while others block apoptosis. Nevertheless, the relative cellular levels of these proteins seem to play a significant role in determining whether cells are committed to apoptosis.

## **1.4 BH3-only proteins may inhibit anti-apoptosis proteins or activate pro-apoptosis proteins**

Despite lacking extensive sequence similarities to the multi-domain proteins, members of the BH3-only subfamily, such as Bid and Bik, are classified as members of the Bcl-2 superfamily of apoptosis proteins based on limited sequence homology to the BH3 domain. The lack of the pore-forming fold suggests that these proteins may not directly cause the release of apoptosis factors from the mitochondria. Yet, members of this subfamily can form interactions with the anti-apoptotic Bcl-2 or Bcl-x<sub>L</sub> through the BH3 domain. Because of this property, it is proposed that the BH3-only proteins participate in the induction of apoptosis by sequestering the anti-apoptotic members of the Bcl-2 superfamily, which in turn prevents these proteins from antagonizing the activity of the pro-apoptotic members (Figure 1.3) (Kelekar and Thompson 1998). Other studies suggest that some members of the BH3-only proteins may actually promote the apoptosis function of Bax or Bak by facilitating the insertion and oligomerization of these proteins at the mitochondrial membrane (Korsmeyer et al. 2000; Kuwana et al. 2002; Kuwana et al. 2005). Therefore, even though the BH3-only proteins may not directly induce apoptosis, their interactions with the anti- or pro-apoptotic members of the Bcl-2 superfamily can lead to the activation of apoptosis.



## 1.5 BNIP3 is a member of the Bcl-2 superfamily of apoptosis proteins

BNIP3 (Bcl-2/Adenovirus E1B 19 kDa-protein interacting protein 3) is a mammalian, mitochondrial membrane-associated protein that was first identified for its ability to interact with Bcl-2 and the adenovirus E1B 19-kDa protein (Boyd et al. 1994; Yasuda et al. 1998; Ray et al. 2000). BNIP3 expression can be detected in various tissues, and its cellular levels are kept low in normal cells. BNIP3 expression can be induced under hypoxic conditions following stroke, heart failure, and ischemia through the action of transcription factor HIF-1 $\alpha$ , and combination of high levels of BNIP3 and acidosis is associated to apoptosis (Bruick 2000; Guo et al. 2001; Sowter et al. 2001; Kubasiak et al. 2002; Frazier et al. 2006). Studies have shown that the levels of BNIP3 are also increased in the hypoxic region of tumor or cancer cells, that stronger BNIP3 expression is observed in more advanced tumor stages, and that the loss of BNIP3 is associated with evasion of hypoxia-induced cell death in malignant and metastatic cancer cells (Erkan et al. 2005; Manka et al. 2005; Lee and Paik 2006; Leo et al. 2006). Because of its clinical significance, the studies of BNIP3 can provide insight into prevention of hypoxic cell death following stroke, heart failure, and ischemia, and they may contribute to improvements in cancer treatments.

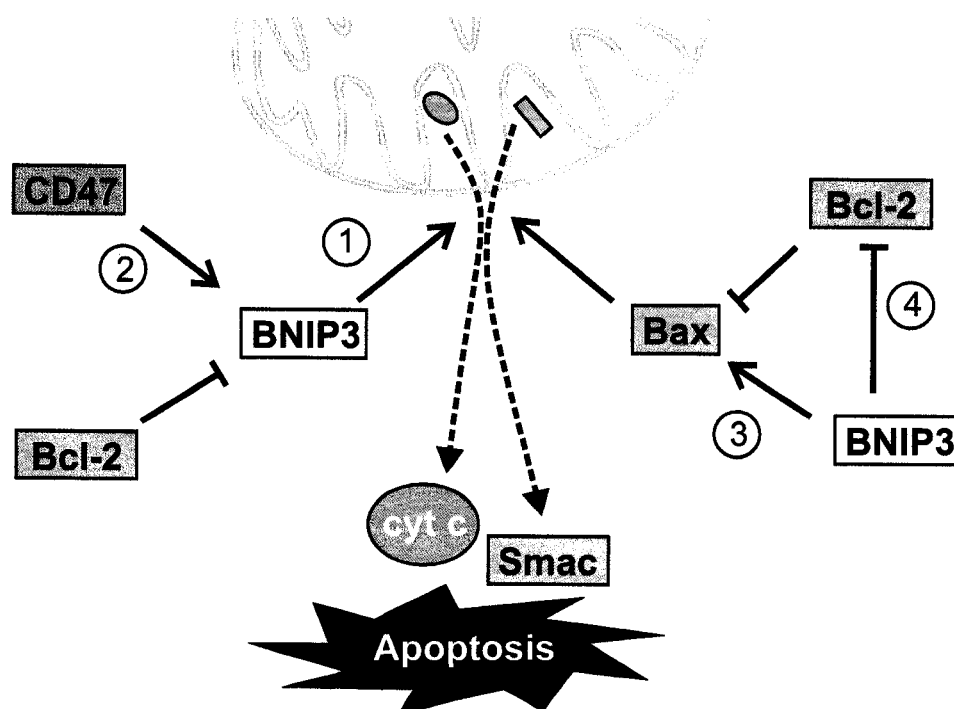
BNIP3 is classified as a member of the BH3-only subfamily based on limited sequence homology to the BH3 domain and the presence a of carboxy-terminal TMD (Figure 1.2) (Yasuda et al. 1998; Ray et al. 2000). However, unlike other BH3-only subfamily members, whose associations with Bcl-2 or Bcl-x<sub>L</sub> are mediated by the BH3 domain, interactions between BNIP3 and Bcl-2 or Bcl-x<sub>L</sub> are mediated, at least partially,

by the transmembrane domain (TMD) (Ray et al. 2000). The same interacting behavior is also observed for Nix, a mammalian homolog of BNIP3, and the *Caenorhabditis elegans* and *Drosophila melanogaster* orthologs of BNIP3 (ceBNIP3 and dmBNIP3, respectively) (Chen et al. 1999; Cizeau et al. 2000). Although BNIP3 and its homologs share only limited sequence homology, the TMD sequences of these homologs are somewhat similar. All four sequences contain the GxxxG motif, which is a common helix-helix interacting motif (Russ and Engelman 2000; Senes et al. 2000), and they possess polar residues in the middle of the TMDs. These similarities suggest that interactions of BNIP3 and its homologs through their TMD may be modulated and stabilized by the same chemical properties (Figure 1.4). Based on these distinct characteristics, their mode for association with Bcl-2, and data from phylogenetic analysis, it is proposed that the BNIP proteins and their homologs can be grouped as a separate class in the Bcl-2 superfamily of apoptosis proteins (Reed et al. 2004; Aouacheria et al. 2005).



## **1.6 BNIP3 may activate apoptosis through multiple mechanisms**

High level expression of BNIP3 has been associated with the induction of apoptosis. However, it is not yet clear how BNIP3 can exert this function. Some studies suggest that BNIP3 can directly promote apoptosis by affecting the mitochondrial permeability transition pore (PTP), and that interaction with the anti-apoptotic Bcl-2 can inhibit this activity (Figure 1.5, pathway 1) (Vande Velde et al. 2000; Kim et al. 2002; Regula et al. 2002; Wan et al. 2003; Webster et al. 2005). Others propose that BNIP3 apoptosis activity may require associations with other apoptosis-associated proteins, such as CD47 (Figure 1.5, pathway 2) (Lamy et al. 2003). Recent studies indicate that BNIP3 may not directly cause apoptosis because expression of BNIP3 at high levels in budding yeast does not result in apoptosis (Guscetti et al. 2005). In addition, it has been reported that induction of apoptosis in mouse embryonic fibroblasts requires co-expression of BNIP3 and Bax or Bak (Kubli et al. 2007), suggesting that BNIP3 may actually mediate apoptosis by promoting the activity of Bax or Bak (Figure 1.5, pathway 3), or by sequestering inhibitors of these proteins, such as Bcl-2 (Figure 1.5, pathway 4).



**Figure 1.5 Possible pathways for BNIP3 apoptosis activity**

Even though overexpression of BNIP3 is associated with increased apoptosis levels, there is no consensus as to how BNIP3 can induce apoptosis. Some believe that BNIP3 can directly cause apoptosis by perturbing mitochondria membrane permeability transition pore (pathway 1), while others suggest that this function requires interactions with other proteins (pathway 2). It is also proposed that BNIP3 can activate apoptosis by enhancing the function of pro-apoptosis members of the Bcl-2 superfamily (pathway 3) or by sequestering the action of the anti-apoptosis members (pathway 4).

## **1.7 BNIP3 interactions through the transmembrane domain are implicated in apoptosis**

Even though the mechanism for BNIP3 apoptosis activity remains unclear, the studies of BNIP3 apoptosis function agree on one condition: the carboxy-terminal TMD of this protein is required for apoptosis function because TMD deletion results in reduced levels of apoptosis in cultured cells (Chen et al. 1997; Ray et al. 2000; Vande Velde et al. 2000; Kim et al. 2002). However, the significance of these reports may have been obscured because the BNIP3 TMD is also necessary for subcellular localization to the mitochondria, where the protein is proposed to function. In addition to cellular localization, the TMD of BNIP3 is required for homodimerization and interactions with other members of the Bcl-2 superfamily of apoptosis proteins (Boyd et al. 1994; Ohi et al. 1999; Ray et al. 2000). These interactions play a role in modulating BNIP3 apoptosis activity because association with Bcl-2 is correlated with the reduction of apoptosis levels, and association with Nix can either promote or inhibit apoptosis, depending on which variant of Nix interacts with BNIP3. Hence, it is reasonable to propose that BNIP3 ability to form homodimer through the TMD is involved in the regulation of the protein's apoptosis activity.

## 1.8 Rationale

In spite of the potential importance for regulation of apoptosis, little is known about the properties that drive and stabilize BNIP3 TMD homodimer. My thesis is aimed at characterizing the chemical and physical basis for BNIP3 TMD self-association. The approach that I used for this study is similar to those that have been used in the pioneering studies of glycophorin A (GpA) TMD dimerization and phospholamban pentamerization (Lemmon et al. 1992; Lemmon et al. 1992; Arkin et al. 1994; MacKenzie et al. 1997). I identified the sequence requirements for BNIP3 TMD dimerization and characterized the chemical and physical basis for this association using saturation mutagenesis, which was coupled with *in vitro* analysis using SDS-PAGE and Western blotting. In addition, I employed solution NMR spectroscopy to determine the structure of the BNIP3 TMD dimer and to establish knowledge of the BNIP3 TMD dimer interface at the atomic level. The BNIP3 TMD dimer structure also allows me to interpret the saturation mutagenesis data in terms of a three-dimensional structure. Though the mutagenesis and structural data do not explicitly explain how BNIP3 can participate in the induction of apoptosis, they do explain the sequence specificity of dimerization, and they provide insight for further studies of BNIP3, such as thermodynamic analysis of BNIP3 TMD dimer, functional characterization of the BNIP3 mutants, and computational modeling of BNIP3 TMD associations.

## 1.9 Outline of the thesis contents

The introduction to the thesis, Chapter 1, contains a brief literature review of apoptosis and the proteins that participate in the regulation of mitochondria-dependent apoptosis in mammalian cells, including BNIP3. It also presents the current knowledge of BNIP3 involvement in apoptosis regulation and indicates the importance of studies of BNIP3 TMD dimerization to further understand the role of BNIP3 and BNIP3 associations in apoptosis regulation.

Chapter 2 describes in detail sample preparation and experimental procedures for *in vitro* analysis of BNIP3 TMD dimerization in detergent micelles. This chapter also explains optimization strategies for overexpression of the *Staphylococcal* nuclease-BNIP3 TMD fusion protein, providing starting conditions for optimization of large scale protein expression needed for biophysical and structural studies of BNIP3 TMD dimerization.

Results obtained from the *in vitro* analysis of BNIP3 TMD dimerization in detergent micelles and interpretation of the data are described in Chapter 3. Many of the materials described in this chapter have been previously published (Sulistijo et al. 2003; Sulistijo and MacKenzie 2006), and I include unpublished data in the discussion as appropriate. In brief, data from saturation mutagenesis enable mapping of residues that form an interaction face for BNIP3 TMD dimerization, which allowed identification of interactions that stabilize the TMD association. Though the mutagenesis data provide hypotheses about the chemical and physical requirements for BNIP3 TMD dimerization,



a high resolution structure of the TMD dimer is needed to confirm structural inferences from these data.

Chapter 4 describes sample preparation methods for the solution NMR study of BNIP3 TMD dimer. In this chapter, I discuss optimization strategies for protein overexpression, generation and isolation of BNIP3 TMD peptide, and conditions for sample preparation. I also describe in detailed the methods for preparing NMR samples of BNIP3 TMD peptide which I used to collect distance and experimentally-derived restraints for determining the structure of BNIP3 TMD dimer.

In Chapter 5, I discuss how I used the data from NMR experiments to determine the structure of the BNIP3 TMD dimer. First, I describe the method used to screen optimal sample and experimental conditions, and the approach employed to collect data for sequential assignments of the backbone atoms, identification of the side chain resonances, and determination of the NOE correlations. Then, I explain the process of calculating the structure of the BNIP3 TMD dimer and show how supplementing distance restraints with additional experimentally-derived restraints can improve the resolution of the BNIP3 TMD dimer structure.

Chapter 6 presents interpretation of the BNIP3 TMD dimer structure in the context of saturation mutagenesis data that are discussed in Chapter 3. In general, the information obtained from the BNIP3 TMD dimer structure is consistent with the conclusions drawn from saturation mutagenesis as both methods identify the same interfacial residues and the types of interactions that stabilize the TMD association. The BNIP3 TMD structure also provides a visual to explain the dimerization phenotypes of

BNIP3 TMD mutants, and enables identification of other stabilizing interactions that are overlooked or cannot be identified from the mutagenesis data.

The final chapter summarizes the results from saturation mutagenesis and solution NMR spectroscopy, and provides a perspective for further studies, which include but are not limited to analysis of BNIP3 TMD dimerization in real membranes, investigation of the biological importance of BNIP3 TMD self-association, and examination of biophysical characteristics of the TMD dimer.

Four appendices are included in this thesis: Appendix A contains the oligonucleotide primer designs used to generate the library of BNIP3 TMD mutants; Appendix B contains the recipes for preparing SDS-PAGE containing glycerol or urea; Appendix C contains experimental restraint files and input files used to generate the family of BNIP3 TMD dimer structures in ARIA version 1.2; and Appendix D contains ARIA code that I modified for calculating structures of the BNIP3 TMD dimer.

## **2 Sample preparation for analysis of BNIP3 transmembrane domain dimerization in detergent micelles**

### **2.1 Introduction**

In this chapter, I discuss the methods that I used to perform *in vitro* analysis of BNIP3 TMD dimerization in detergent micelles. I determined the sequence requirements for BNIP3 transmembrane domain (TMD) self-association in detergent micelles using saturation mutagenesis and SDS-PAGE, similar to the approach that has been described by Lemmon and colleagues (Lemmon et al. 1992; Lemmon et al. 1992). I generated single-point substitution mutations in the BNIP3 TMD sequence of a *Staphylococcal* nuclease (SN)-BNIP3TMD fusion protein construct, and analyzed the effects of mutations on BNIP3 TMD dimerization using SDS-PAGE. The use of the SN-BNIP3TMD fusion protein instead of the full-length BNIP3 protein in this analysis allows high level expression of the protein, which is advantageous for biophysical studies of BNIP3 TMD association, and the purification technique for the SN-TMD fusion protein that I developed is applicable for subsequent biophysical and structural studies. In the first half of the chapter, I explain the approaches to find optimal sample preparation procedure and experimental conditions. Then, I describe the methods for sample preparation and protocols for data collection, which allowed me to elucidate the sequence specificity of BNIP3 TMD dimerization in SDS micelles.

## 2.2 Optimization of the SN-BNIP3TMD fusion protein overexpression

For overexpression of the SN-BNIP3TMD fusion protein in *Escherichia coli* (*E. coli*) cells, the genes for the fusion protein were inserted in a plasmid derived from the pET11c vector. Because different *E. coli* strains may have different efficiency for SN-BNIP3TMD protein overexpression, I examined the expression levels of the SN-BNIP3TMD fusion protein in the BL21(DE3), BL21(DE3)pLysS, HMS174(DE3), and HMS(DE3)pLysS strains. The first strains that I tested were BL21(DE3)pLysS and HMS174(DE3)pLysS because these strains contain the pLysS plasmid that confers stringent control for protein expression, and thus eliminate concerns for possible toxicity caused by leaky expression of the SN-BNIP3TMD fusion protein. However, protein expression using the BL21(DE3)pLysS and HMS174(DE3)pLysS cells resulted in extremely low yields of the SN-BNIP3TM fusion protein as I could not detect the fusion protein from crude cell extract using SDS-PAGE stained with Coomassie blue. Interestingly, when I expressed the protein from these same cells in cultures that had been grown without chloramphenicol, I could detect the presence of the fusion protein when I resolved the crude cell extracts using SDS-PAGE, indicating that the initial failure was not due to issues with the expression plasmid or *E. coli* expression system in general. Because addition of chloramphenicol to the medium is necessary to maintain the pLysS plasmid, this result also suggests that the stringency of the pLysS plasmid inhibits high levels expression of the SN-BNIP3TMD fusion protein, and that the use of *E. coli* strains lacking pLysS plasmid, the BL21(DE3) and HMS174(DE3) strains, might be preferable. Examination of protein expression levels obtained with these two cell strains demonstrated that the BL21(DE3) strains yields higher levels of the SN-

BNIP3TMD fusion protein, and based on this result, I used the BL21(DE3) cell strain for subsequent protein overexpression.

In addition to cell strains, I screened other conditions for optimal expression of the SN-BNIP3TMD fusion protein, including the optical density at which protein expression was initiated ( $OD_{600}$  of 0.6-1.0), the amount of IPTG added for induction (0.4-0.8 mM), and the incubation period (1-4 hours) for expression prior to harvesting. I found that varying IPTG concentrations for induction did not significantly affect the amount of protein expressed, and that I can use as low as 0.4 mM IPTG to induce high level protein expression. In contrast, induction at higher  $OD_{600}$  and longer incubation period resulted in higher protein yield. A concern with induction at high  $OD_{600}$  is an increased number of cell deaths, which can lower the yield. An excessively long incubation period may not be beneficial either because the protein of interest may start degrading. Based on these concerns and the analytical results, I decided to induce protein expression is at  $OD_{600}$  of ~1.0, when *E.coli* cells are still in exponential growth, and to harvest the cells three to four hours post-induction.

## 2.3 Optimization of detection method for SDS-PAGE analysis

Because BNIP3 forms SDS resistant dimers through the TMD (Chen et al. 1997), SDS-PAGE is an excellent tool to simply and quickly analyze the effects of single-substitution mutations on the TMD self-association, and a variety of detection methods are available to determine the effects of substitutions in the TMD sequence on the ability of the SN-BNIP3TMD fusion protein to form dimers. Coomassie staining is the most common technique for visualizing proteins resolved with SDS-PAGE. However, this detection technique is not sensitive and can introduce a problem when expression levels are low. Another issue related to Coomassie staining for detection is that the stain does not discriminate between proteins. This means that other proteins that remain in the sample will also be detected and can cause overestimation of the SN-BNIP3TMD band intensities. This concern also eliminates the use of silver staining as a detection method even though this method is more sensitive than Coomassie staining. Thus, Western blotting is the preferred detection method: it is more efficient than either Coomassie or silver staining because it is selective and detects only the target protein, and it is sensitive and allows detection of small amounts of protein. In this *in vitro* analysis, I performed Western blotting using a primary antibody raised against SN. This antibody enables detection of mutants of the SN-BNIP3TMD, and it can be used to probe other SN-TMD constructs, making this approach cost-effective.

## 2.4 Experimental methods for SDS-PAGE analysis

### 2.4.1 Construction of plasmid bearing gene encoding SN-BNIP3TMD fusion protein

The expression plasmid bearing the gene that codes for the *Staphylococcal* nuclease-BNIP3 transmembrane (TM) domain fusion protein (SN-BNIP3TMD) was generated by my advisor Kevin MacKenzie from a previous SN-GpATMD fusion protein vector (pT7SN/GpATM) in a manner described by Lemmon *et al* (Lemmon et al. 1992). For this construction, the gene for the GpA TMD was cleaved with Apa I/BamH I digestion, and the DNA encoding residues 146-194 of mammalian BNIP3 was inserted in-frame at the Apa I site, which encodes a GlyPro linker. The coding sequence of the pT7SN/BNIP3TM plasmid was verified using automated dideoxynucleotide sequencing. The amino acid sequence of BNIP3 inserted in the fusion protein construct, with the predicted TMD underlined, is the following:

**GPRNTSVMKKGIFSAEFLKVFLPSLLLSHLLAIGLGIYIGRRLTTSTSTF\***

### 2.4.2 Generation of single-point mutation library of SN-BNIP3TMD fusion protein

I generated single substitution mutations in the SN-BNIP3TMD fusion construct by oligonucleotide-directed mutagenesis using the QuikChange kit (Stratagene, La Jolla CA). The sets of primers used in the reactions consist of a long primer and a short primer. The long primer contains a redundant codon, which encodes two to eleven different amino acids at the target position, and a silent mutation that eliminates or

introduces a convenient restriction enzyme site (see Appendix A). The short primers are non-mutagenic and anneal to the 3' end of the long primer without overlapping the mutation sites. In the first round of DNA amplification, short primers that anneal to the long primers are extended to generate long oligonucleotides that complement the mutagenic long primers. The use of non-mutagenic short primers lowers the efficiency of the mutagenesis reactions to some degree. However, the short primers are inexpensive, and they can complement a number of different long primers, and thus these primers can be used in multiple mutagenesis reactions. The sequences of the mutagenesis products were confirmed by automated dideoxynucleotide sequencing.

#### **2.4.3 Small scale expression and preparation of SN-BNIP3TMD fusion proteins**

Glycerol stocks for overexpression of the SN-BNIP3TMD fusion protein were made by transforming the pT7SN/BNIP3TM vector and the mutant plasmids into BL21(DE3) cells and plated onto Luria Bertani (LB) plates containing  $50 \mu\text{g ml}^{-1}$  ampicillin. The growing colonies were inoculated in 5 ml LB medium plus  $50 \mu\text{g ml}^{-1}$  carbenicillin and cultures were grown at  $37^\circ\text{C}$  with shaking. When cultures reached  $\text{OD}_{600} \sim 0.2$ , the cultures were brought to 15% (v/v) glycerol and stored at  $-80^\circ\text{C}$ . For small scale protein expression, I inoculated the glycerol stocks into 5 ml LB medium supplemented with carbenicillin and let them grow overnight at  $37^\circ\text{C}$ . The overnight cultures were then diluted 1:200 in 5 mL fresh LB medium plus  $50 \mu\text{g ml}^{-1}$  carbenicillin, the cultures were grown with shaking at  $37^\circ\text{C}$ , and protein expression was induced at  $\text{OD}_{600} \sim 1.0$  by adding 0.5 mM IPTG. The cells were harvested 3 hours after induction by centrifugation at  $16,000 \times g$  for 5 min using a benchtop centrifuge. Next, the cell pellet



was resuspended in 1:20 culture volume of lysis buffer (20 mM Tris-HCl, 2 mM EDTA, pH 8.0). The cells were lysed by subjecting the cell resuspension to three rounds of freezing and thawing and probe sonication. Then, 10 mM  $\text{CaCl}_2$  and 0.1 mg  $\text{ml}^{-1}$  hen egg white lysozyme were added, and the samples were incubated on ice for 30 min to allow digestion of the chromosomal DNA by SNase. Afterward, the soluble and insoluble fractions were separated by centrifugation at 16,000 x  $g$  for 5 min. The supernatant was removed with decantation and the insoluble fraction, which contains the SN-BNIP3TMD fusion protein, was treated with lysis buffer plus 1 M ammonium acetate. The SN-BNIP3TMD protein remained in the insoluble fraction with this treatment and was recovered by centrifugation at 16,000 x  $g$  for 5 min. For storage and SDS-PAGE analysis, I resuspended the pellet in 1:20 culture volume of lysis buffer.

#### **2.4.4 SDS-PAGE analysis and Western blotting**

I prepared samples for SDS-PAGE by mixing sample stocks in SDS-PAGE loading buffer (2% (w/v) SDS, 100 mM 1,4-dithiothreitol (DTT), 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue, and 50 mM Tris-HCl pH 6.8). Typically the sample mix contains  $\sim 10 \mu\text{g ml}^{-1}$  of SN-BNIP3TMD fusion protein. Prior to loading, sample mixes were heated at 95 °C for 5 min and allowed to cool to room temperature before being loaded on 15% pre-cast polyacrylamide gels (Bio-Rad). The proteins were separated using SDS-PAGE with a running buffer containing 0.2% (w/v) SDS.

For Western blotting, I blotted the proteins, separated using SDS-PAGE, onto nitrocellulose membrane in a transfer buffer containing 48 mM Tris, 39 mM glycine, 20% (v/v) methanol, 0.1% (w/v) SDS, pH 9.2. The SN-BNIP3TMD proteins were probed on the blot with an anti-*Staphylococcal* nuclease antibody (prepared by Bethyl Industries) and visualized the probed proteins with an antirabbit horseradish peroxidase conjugate and ECL reagent (GE Healthcare). Most chemiluminescence images were captured with Kodak film, though at later stage, a Fluorchem 5500 documentation system was used to capture digital images of the chemiluminescence signal.

## **2.5 Experimental methods for circular dichroism and Fourier transform infrared spectroscopy**

### **2.5.1 Large scale expression and purification of the SN-BNIP3TMD fusion protein**

Large scale expression of the SN-BNIP3TMD fusion protein is similar to the small scale expression, except that the overnight cultures were diluted 1:100 in 100 mL fresh LB medium. Cells were harvested 3 hours after induction by centrifugation at 12,000 x *g* for 5 min and resuspended in 1:20 culture volumes of lysis buffer. Cells were lysed with three rounds of freezing and thawing and probe sonication, 10 mM CaCl<sub>2</sub> and 0.1 mg ml<sup>-1</sup> hen egg white lysozyme were added, and the sample was incubated on ice for 30 min for digestion of chromosomal DNA. The soluble and insoluble fractions were separated by centrifugation at 12,000 x *g* for 10 min, and the soluble fraction was removed with decantation. Then, the insoluble fraction was treated with lysis buffer containing 1 M ammonium acetate and with lysis buffer containing 2% (w/v) detergent

Thesit (Fluka). To extract the SN-BNIP3TMD fusion protein, the insoluble fraction was treated with lysis buffer containing both 1 M ammonium acetate and 2% (w/v) Thesit. The resuspension was probe sonicated, and the soluble and insoluble fractions were separated by centrifugation at 12,000 x *g* for 20 min. At this stage the SN-BNIP3TMD fusion protein is found in the supernatant. To reduce the concentrations of ammonium acetate and Thesit, the supernatant was dialyzed twice against 20 volumes of lysis buffer plus 100 mM ammonium acetate and 0.2% (w/v) Thesit. Aggregates resulting from the dialysis were removed by centrifugation at 12,000 x *g* for 20 min. The dialyzed extract supernatant was passed over a DE52 column (Whatman) to remove nucleic acids, and then bound to a CM52 column (Whatman) and eluted with a buffer containing 400 mM ammonium acetate and 0.2% (w/v) Thesit. Prior to final purification with reverse-phase HPLC, the ammonium acetate concentration in the sample was lowered by dialysis against water. Then, the sample was dried by lyophilization and treated with hexane to remove Thesit. Reverse-phase HPLC purification was carried out using a phenyl stationary phase with a gradient of acetonitrile:isopropyl alcohol (3:2), 0.1% (v/v) 2,2,2-trifluoroacetic acid (TFAA). The protein elutes from the column when the organic solvents composition is roughly 55%.

### **2.5.2 Sample preparation for circular dichroism and Fourier transform infrared spectroscopy**

Samples were prepared by dissolving purified wild type and mutant SN-BNIP3TM fusion proteins in a 1:1 mixture of 2,2,2-trifluoroethanol (TFE) and water. Then, buffer containing SDS and Tris-HCl pH 8.0 was added, and the samples were lyophilized until completely dry. Next, the samples were reconstituted in water and relyophilized. This step was repeated two or three times to ensure that residual TFAA from HPLC purification is removed as much as possible.

### **2.5.3 IR spectra collection and analysis**

Prior to FTIR spectroscopy, the dried samples were redissolved in D<sub>2</sub>O to concentrations of 0.5 mM fusion protein, 4% (w/v) SDS, and 50 mM Tris-HCl pH 8.0, and the samples were allowed to equilibrate at room temperature for 20 hours. For FTIR, 20  $\mu$ l of these samples was loaded between CaF<sub>2</sub> windows with a 50  $\mu$ m path length spacer. The IR spectra were collected with <sup>2</sup>H<sub>2</sub>O background at 25 °C on a Nicolet Nexus 470 FTIR spectrometer, which is equipped with a Mercury-Cadmium-Telluride (MCT) detector and was purged continuously with dry N<sub>2</sub>. A total of 64 scans were recorded per spectrum at a resolution of 1 cm<sup>-1</sup>. For correction, buffer signal (50 mM Tris-HCl pH 8.0, 4% SDS in D<sub>2</sub>O) was subtracted from the spectra, and atmospheric vapor signal was removed such that the featureless region between 1900 and 2100 cm<sup>-1</sup> had an absorbance of zero. To obtain clean spectra of SN/BNIP3TM fusion proteins, signals from residual TFAA were removed such that the band at 1145 cm<sup>-1</sup> had an absorbance of zero. The (mutant – wild type) difference spectra were generated by

subtracting the spectra of the wild type fusion protein from those of the mutant fusion proteins. Variations in fusion protein concentrations were corrected by adjusting the scale of the intensity of the prominent tyrosine band (at  $1515\text{ cm}^{-1}$ ) (Venyaminov and Kalnin 1990) that arises from the nine Tyr residues in the nuclease region and the single Tyr in the BNIP3 TMD region of the fusion protein.

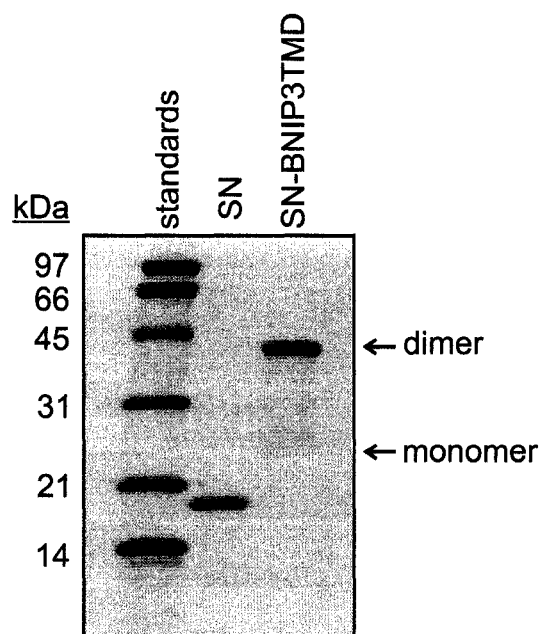
#### **2.5.4 Circular dichroism spectroscopy**

Prior to CD spectroscopy, the dried samples were redissolved in water and brought to  $20\text{ }\mu\text{M}$  protein,  $50\text{ mM}$  Tris-HCl pH 8.0, and 4% (w/v) SDS. CD spectra were acquired between 260-200 nm at  $25\text{ }^{\circ}\text{C}$ . Three scans were collected and averaged for each sample. Spectra of the buffer without the fusion protein were also collected for background correction.

### **3 Analysis of BNIP3 transmembrane domain dimerization in detergent micelles**

#### **3.1 The BNIP3 transmembrane domain is sufficient for dimerization**

Analyses using SDS-PAGE and yeast two-hybrid have indicated that the transmembrane domain (TMD) of BNIP3 is necessary for dimerization (Chen et al. 1997). To test whether the TMD is sufficient for dimerization, I analyzed the mobility of an in-frame fusion of *Staphylococcal* nuclease (SN)-BNIP3TMD using SDS-PAGE. This fusion protein migrates predominantly as a 42 kDa band (Figure 3.1), which is approximately twice the size of the SN-BNIP3TMD fusion protein (MW ~22 kDa), indicating that the BNIP3 TMD is sufficient to drive dimerization of a heterologous protein in SDS micelles. The fact that SN migrates as a 19 kDa band confirms that the association of the SN-BNIP3TMD fusion protein is mediated by the BNIP3 TMD and not by SN (Figure 3.1). The ability of the BNIP3 TMD to drive SDS-resistant dimerization places the BNIP3 protein among a few membrane spanning proteins that can form oligomers in the conditions for SDS-PAGE. Proteins in this category include single-span membrane proteins glycophorin A (GpA) (Lemmon et al. 1992) and phospholamban (Imagawa et al. 1986; Fujii et al. 1989; Arkin et al. 1994; Cornea et al. 1997), the porins and other  $\beta$ -barrel outer membrane proteins (Watanabe 2002; Voulhoux et al. 2003), and the *E. coli* ammonium transporter AmtB (Blakey et al. 2002).



**Figure 3.1 The SN-BNIP3TMD fusion protein dimerizes on SDS-PAGE**

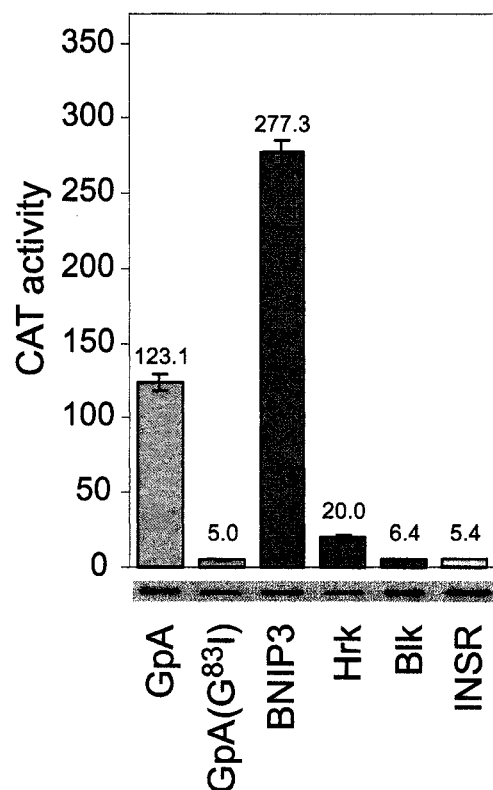
The SN-BNIP3TMD fusion protein migrates as a 42 kDa dimer band when separated using 15% SDS-PAGE (labeled as dimer). The fusion protein has a molecular weight of 22,249 dalton, but a band of this molecular weight can be barely observed under this analytical conditions (labeled as monomer). *Staphylococcal* nuclease (SN) is included in this analysis to confirm that the SN-BNIP3TMD fusion protein dimerization is mediated by the BNIP3 transmembrane domain, and not by SN. (Adapted from Sulistijo et al. 2003)

To determine whether the BNIP3 TMD can drive dimerization in real membranes, Todd Jaszewski performed an analysis using TOXCAT assay, a method that reports the strength of TMD helix-helix associations in *E. coli* membranes through measurement of the expression levels of a reporter gene chloramphenicol acetyltransferase (CAT) using a spectrophotometric enzymatic assay (Russ and Engelman 1999; Sulistijo et al. 2003). He found that the BNIP3 TMD can mediate dimerization in *E. coli* membranes, and that this association is stronger than GpA TMD dimerization (Figure 3.2). For comparison, Todd Jaszewski also measured the associations of the TMDs of other BH3-only proteins, Hrk and Blk, and of human insulin receptor (INSR), and he found that these TMDs cause low levels expression of CAT, suggesting that TMD sequences with similar length, composition, and membrane-targeting properties do not constitutively activate the expression of CAT.

### **3.2 BNIP3 transmembrane domain dimerization in SDS micelles and in *E. coli* membranes are stabilized by the same interactions**

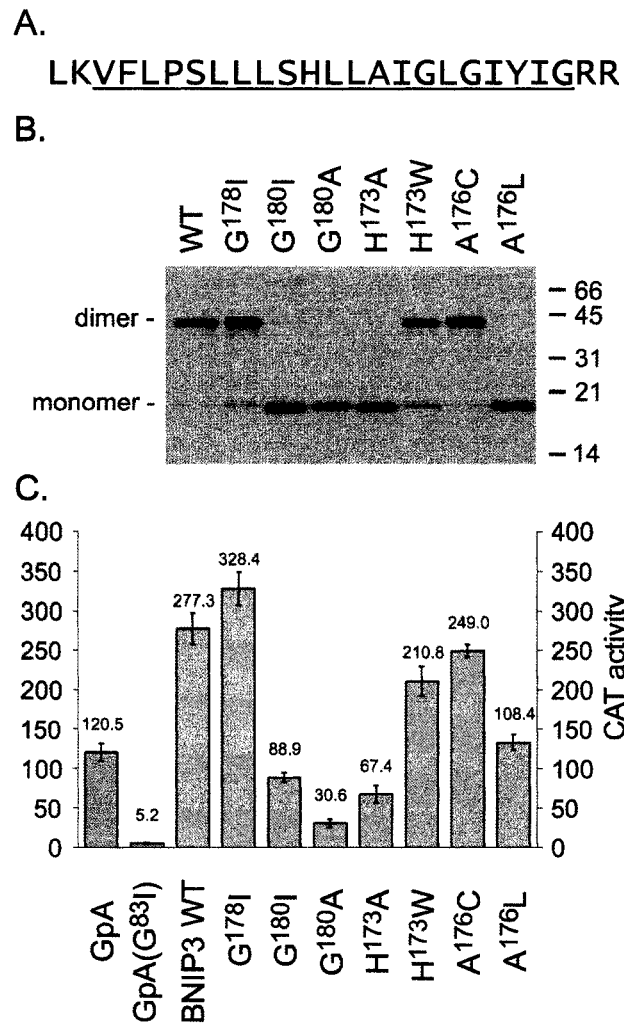
To verify that BNIP3 TMD dimerization on SDS-PAGE and in *E. coli* membrane are stabilized by the same interactions, Todd Jaszewski and I analyzed the effects of substitution mutations using TOXCAT and SDS-PAGE, respectively (Sulistijo et al. 2003). We found that mutants that show disruptive phenotypes on SDS-PAGE, such as the Gly180 to Ala mutant, have reduced levels of CAT expression, while those that show phenotypes similar to the wild type, such as the Ala176 to Cys mutant, have high levels of CAT expression (Figure 3.3). The limited mutagenesis data also show that there is sequence dependence for BNIP3 TMD dimerization.





**Figure 3.2 Quantification of CAT expression levels reports the strength of transmembrane domain interactions in *E. coli* membranes**

Dimerization of BNIP3 transmembrane domain (TMD) in *E. coli* membranes was measured by Todd Jaszewski using spectrometric TOXCAT assay, which reports the strength of the association as the expression level of a reporter gene chloramphenicol acetyltransferase (CAT). Here, the amount of CAT expressed is translated as the levels of CAT activity. The positive control for this analysis is glycophorin A (GpA) TMD, which can mediate strong association in *E. coli* membrane, and the negative control is a monomeric mutant of the GpA TMD, the GpA (G83I). Also included in the analysis are the TMDs of Hrk, BIK, and INSR to show that TMDs with similar length, composition, and membrane targeting properties do not always activate the expression of CAT. The Western blot under the histogram shows the amount of fusion protein containing these TMDs present in crude cell extract and indicates that the different levels of CAT expression is not the result of varying levels of fusion proteins present in the cells. (Adapted from Sulistijo et al. 2003)



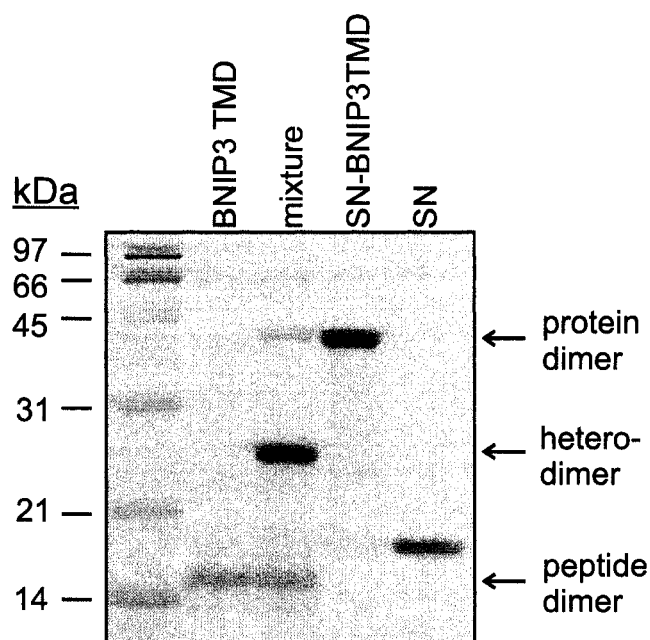
**Figure 3.3 Limited saturation mutagenesis data indicates that the same interactions stabilize BNIP3 transmembrane domain dimerization on SDS-PAGE and in *E. coli* membranes**

Presented are the sequence of the BNIP3 transmembrane domain (TMD) (panel A), the effects of substitution mutations on BNIP3 TMD dimerization on SDS-PAGE (panel B), the effects of the same mutations on BNIP3 TMD dimerization in *E. coli* membranes (panel C). The analysis on SDS-PAGE was performed by me, and the analysis in *E. coli* membranes was performed by Todd Jaszewski. The same effects observed in both systems suggest that BNIP3 TMD dimerization on SDS-PAGE and in *E. coli* membranes are mediated by the same interactions. (Adapted from Sulistijo et al. 2003)

For instance, Gly180 to Ile mutation abrogates dimerization abrogation in both SDS-PAGE and TOXCAT, whereas the Gly178 to Ile mutation has no effects on dimerization in either system. Moreover, specific chemical characteristics are involved in the stability of the TMD dimer because the His173Ala mutation can completely disrupt BNIP3 TMD dimerization, but the His173Trp mutation reduces dimerization only slightly. Together, these results confirm that BNIP3 TMD association in SDS micelles and in *E. coli* membranes involves the same stabilizing properties, but a more complete mutagenesis data is needed to elucidate sequence requirements for BNIP3 TMD dimerization in either system.

### **3.3 BNIP3 transmembrane domain dimerization is reversible and not kinetically trapped**

Analysis of the BNIP3 TMD dimerization on SDS-PAGE also showed that this association is reversible. Mixing the SN-BNIP3TMD fusion protein with purified BNIP3 TMD peptide, followed by heating at 95 °C and separation with SDS-PAGE, results in a loss of the 42-kDa protein homodimer band and an appearance of a new 26 kDa band (Figure 3.4). The migration position of the new band corresponds roughly to the size of the SN-BNIP3TMD fusion protein monomer plus the BNIP3 TMD peptide monomer, thus it is assigned as a protein-peptide heterodimer. The depletion of the protein homodimer band and the appearance of the protein-peptide heterodimer band demonstrate that the BNIP3 TMD peptide is able to compete with the SN-BNIP3TMD fusion protein to form TMD-TMD complexes, indicating that SN-BNIP3TMD dimers are in equilibrium with the monomers and are not kinetically trapped.



**Figure 3.4 Mixing the SN-BNIP3TMD fusion protein with the BNIP3 transmembrane domain peptide generates a protein-peptide heterodimer**

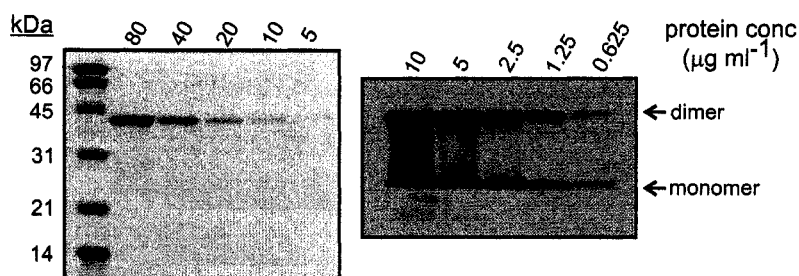
Purified BNIP3 transmembrane domain (TMD) peptide and SN-BNIP3TMD fusion protein migrate as dimers on 15% SDS-PAGE. Mixing the TMD peptide with the fusion protein cause a depletion of the peptide and protein homodimer and appearance of a new band that migrates at roughly 26 kDa. The size of the new band corresponds to a protein-peptide heterodimer (22 kDa + 4 kDa). (Adapted from Sulistijo et al. 2003)

### **3.4 Dissociation constant for BNIP3 transmembrane domain dimerization is influenced by the chosen analytical conditions**

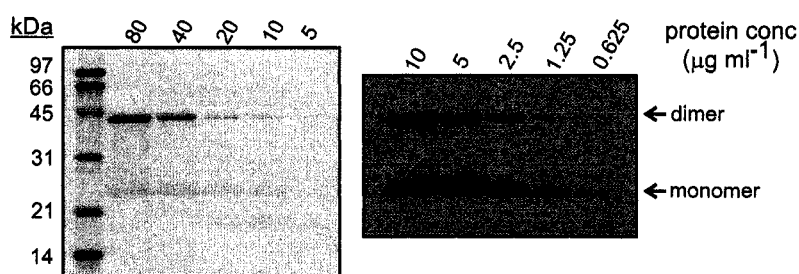
A series of two-fold dilutions of the SN-BNIP3TMD demonstrate that the apparent dimer-to-monomer ratio on SDS-PAGE is concentration dependent (Figure 3.5), further supporting the idea that the BNIP3 TMD dimer and monomer are in equilibrium. At higher concentrations, the wild type SN-BNIP3TMD appears entirely dimeric as shown with Coomassie blue staining, and as the protein is diluted, the monomer band can be detected. At a concentration of  $1.25 \mu\text{g ml}^{-1}$  (55 nM) the dimer and monomer band intensities are approximately the same, and because of this concentration dependence, further dilution of the wild type SN-BNIP3TMD would have shown the monomer as the major species. A less tightly associating mutant of the SN-BNIP3TMD, His173Trp, shows detectable monomer with Coomassie staining, and the dimer and monomer band have approximately the same intensities at a higher concentration of  $10 \mu\text{g ml}^{-1}$  (450 nM). The apparent dissociation constants of the wild type and the mutant SN-BNIP3TMD can be calculated from the relative dimer-to-monomer ratio, and they are approximately 55 and 450 nM, respectively.

BNIP3 TMD dimerization is also dependent on the conditions of the analysis. Separating somewhat disruptive Leu169Phe and Ile183Leu mutants with electrophoresis using SDS-PAGE containing 13% (w/v) glycerol results in more intense dimer bands (Figure 3.6 B), while the same mutants resolved on SDS-PAGE in the presence of 8 M urea show a depletion of the dimer bands (Figure 3.6 C). The conditions for the glycerol- and urea-SDS-PAGE are listed in Appendix B.

### A. Wild type SN-BNIP3TMD

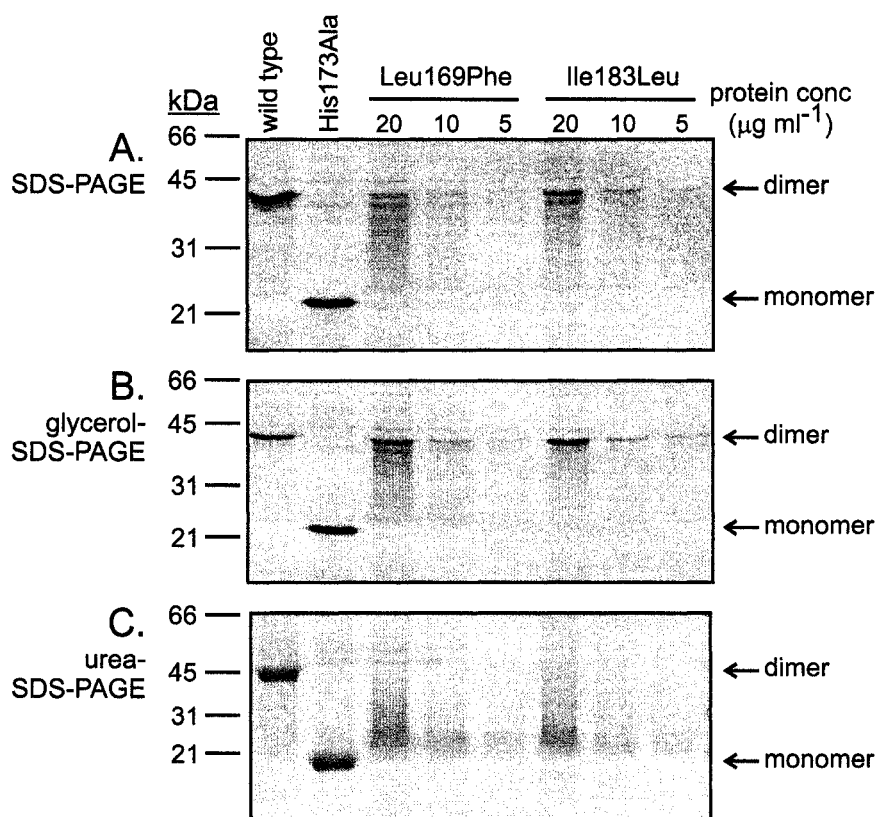


### B. His173Trp SN-BNIP3TMD



**Figure 3.5 Serial dilutions of purified wild type and His173Trp mutant of the SN-BNIP3TMD fusion proteins**

Purified wild type SN-BNIP3TMD fusion protein (A) and His173Trp mutant (B) were subjected to two-fold serial dilution in loading buffer containing 4% (w/v) SDS, separated by SDS-PAGE, and detected with Coomassie staining (left) and Western blotting (right). The concentration of the protein is indicated in each lane, and the positions of the dimer and monomer band are shown on the right. (Adapted from Sulistijo et al. 2003)



**Figure 3.6 Serial two-fold dilution of Leu169Phe and Ile183Leu mutants of the SN-BNIP3TMD fusion proteins in glycerol- and urea-SDS-PAGE**

Leu169Phe and Ile183Leu mutants of SN-BNIP3TMD fusion protein were subjected to two-fold serial dilution in loading buffer containing 2% (w/v) SDS, separated by electrophoresis on regular (A), 13% (w/v) glycerol (B), or 8 M urea (C) SDS-PAGE. The proteins were detected with Coomassie staining. Wild type and monomeric mutant (His173Ala) of the SN-BNIP3TMD were included to indicate the positions of the dimer and monomer bands. The concentration of the protein is indicated in each lane.

Together, these data indicate that protein concentrations and the conditions for SDS-PAGE can influence the apparent dimer-to-monomer ratio. Therefore, when analyzing effects of mutations on BNIP3 TMD dimerization, it is important to keep experimental conditions identical, so that influences that can alter the apparent dimer-to-monomer ratio, and thus the apparent dissociation constant, should cancel out. This precaution ensures that the observed variation between the wild type and mutants SN-BNIP3TMD fusion proteins accurately represent the relative effects of the mutations on self-association.

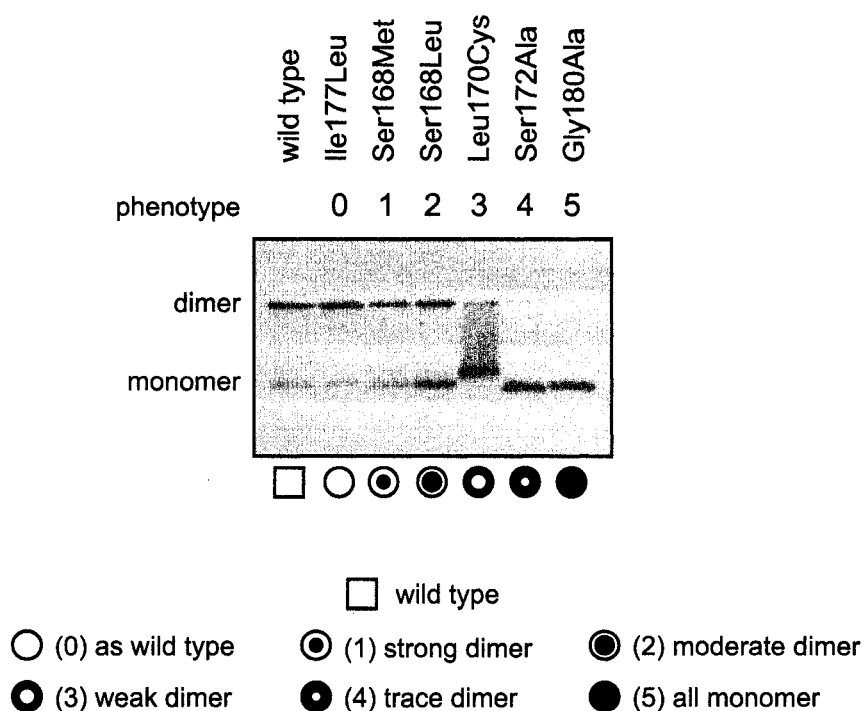
### **3.5 Substitutions in the BNIP3 transmembrane domain cause a range of dimerization phenotypes**

As mentioned in section 3.2, sequence dependence of the BNIP3 TMD dimerization can be elucidated from analysis of saturation mutagenesis data. For this purpose, I generated a library of single point substitution mutants in the context of the SN-BNIP3TMD fusion protein and characterized the effects of these mutations on dimerization using SDS-PAGE and Western blotting. Each residue in the predicted BNIP3 TMD sequence (residues 164-184) was subjected to site-directed mutagenesis using redundant oligonucleotides designed to give hydrophobic and slightly polar amino acids. In generating this library, I avoided introducing mutations to polar or charged amino acids because the study of GpA TMD dimerization indicated that substitution to polar and charged amino acids can disrupt TMD dimerization whether the substitutions are located at or away from the dimer interface (Lemmon et al. 1992). In contrast, substitutions to hydrophobic and slightly polar amino acids perturb GpA TMD dimerization only at a subset of positions that were first identified as the dimer interface



by mutagenesis analysis (Lemmon et al. 1992) and then confirmed with the NMR structure (MacKenzie et al. 1997). With the approach described in section 2.4.2, I have generated a total of 195 single substitution mutations to hydrophobic and slightly polar residues over the 21 amino acid-long TMD sequence.

Figure 3.7 presents the extent to which the SN-BNIP3TMD dimerizes on SDS-PAGE under the chosen experimental conditions and the effects of six single substitution mutations on dimerization. The wild type SN-BNIP3TMD migrates predominantly as a dimer, with a trace of monomer. Mutations in the BNIP3 TMD sequence result in a variety of SN-BNIP3TMD dimerization phenotypes, ranging from a phenotype similar to the wild type to complete disruption of the dimer. Based on the relative dimer-to-monomer ratio, these mutants are grouped into six classes: those that dimerize “as wild type”, those that exhibit “all monomer”, and four intermediate classes. These classes represent roughly three-fold decreases in the apparent dimer-to-monomer ratio. To simplify data recording, a symbol and a score are assigned to represent each class: white circle (score=0) for the “as wild type” mutants, black circle (score =5) for “all monomer” mutants, and circles with different black and white arrangements (score=1-4) for the intermediate classes, where higher proportion of black and a higher score indicate more a disruptive phenotype. Roughly 30% of the SN-BNIP3TMD mutants show smearing phenotypes that can be observed either upward from the monomer position or downward from the dimer position (Leu170Cys, lane 3). This is probably caused by the mutant proteins associating and dissociating as they are being resolved using SDS-PAGE. The different association constants and the different rates for association and dissociation determine whether the smearing is upward or downward.



**Figure 3.7 Representative of dimerization phenotypes caused by substitutions in the BNIP3 transmembrane domain**

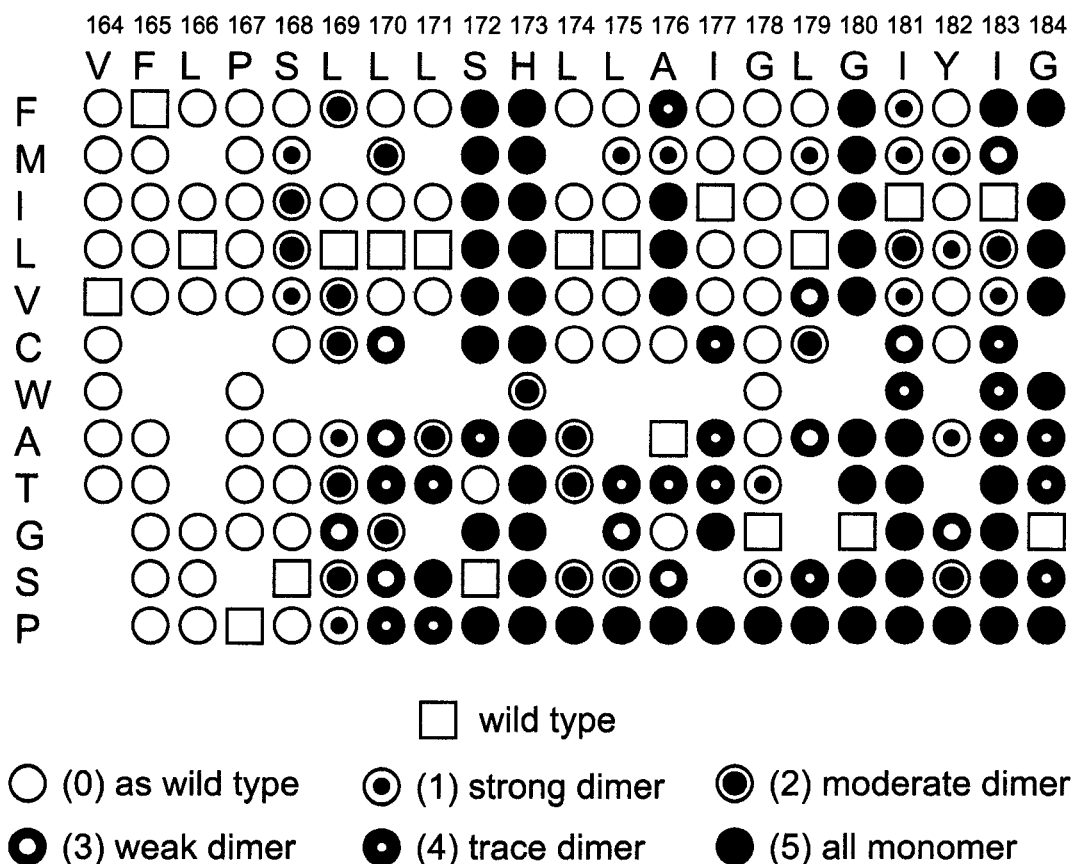
The Western blot presents the dimerization phenotypes of wild type SN-BNIP3TMD and six different mutants. The identity and phenotype score for each lane are given above, and the phenotype symbol at the bottom. The wild type and mutants with wild type-like phenotype, such as Ile177Leu are predominantly dimeric with a trace of monomer. These wild type-like mutants are given a score of 0 and symbol of white circle. Mutants with increased monomer band intensities are given phenotype scores of 1-5, with phenotype 1 (white circle) shows a detectable increased in monomer band intensity, phenotype 2 has approximately 1:1 dimer-to-monomer band intensities, phenotype 3 shows more monomer than dimer, phenotype 4 is mostly monomeric, and phenotype 5 (black circle) shows only the monomer band with no trace of the dimer. (Adapted from Sulistijo and MacKenzie 2006).

A limitation of SDS-PAGE in this analysis is that this method is unable to distinguish mutants that may exhibit stronger association than wild type under the chosen experimental conditions. A different experimental approach, such as sedimentation equilibrium analytical ultracentrifugation or fluorescence resonance energy transfer (FRET), will be needed to identify these mutants.

### **3.6 Different substituting residues have different effects on BNIP3 TMD dimerization**

Figure 3.8 shows the library of 195 BNIP3 TMD single substitution mutants and their corresponding phenotypes. Mutations to polar amino acids and at positions carboxy-terminal of the predicted TMD region were made, and they will be presented in other sections in this chapter. Because the substitution mutations were targeted for hydrophobic and less polar amino acids, only a limited number of substitutions to Met and Trp were generated, and substitutions to Tyr are absent.

Inspection of the phenotype chart in Figure 3.8 reveals that different substituting residues have different effects on BNIP3 TMD dimerization. For example, substitutions to Phe or Val show a periodic disruption of dimerization, substitutions to Ala show more disruptive phenotypes at most positions, and substitutions to Pro in the carboxy-terminal two-thirds of the TMD cause complete disruption of the BNIP3 TMD dimer. In the studies of GpA TMD dimerization, a periodic disruption pattern was interpreted to map the residues that are located at the dimer interface (Lemmon et al. 1992). The same approach can be applied in this study, and by analyzing the average disruption pattern at each position, we expect



**Figure 3.8 The library of BNIP3 transmembrane domain single substitution mutants and their corresponding phenotypes**

The library of single substitution mutants are presented with the BNIP3 transmembrane domain sequence listed on top and the substituting residues listed on the left. The symbols and scores on the legend represent various effects of substitutions on SN-BNIP3TMD dimerization on SDS-PAGE as discussed in section 3.5 and shown in Figure 3.7. (Adapted from Sulistijo and MacKenzie 2006).

to be able to make stronger conclusions about the interacting residues for the BNIP3 TMD dimerization. However, selecting which substitutions are to be averaged is not that simple because different substituting residues exhibit different disruption patterns. Furthermore, classifying the substituting residues based on their disruption pattern will introduce bias in the final results. Therefore, statistical measures that ignore positional effects of substitutions were applied to generate unbiased pools of substituting residues.

Table 3.1 shows the phenotype distributions for the substituting residues. Row “all” pools all mutagenesis data and reveals that of 195 mutants, 74 (38%) associate as wild type (score=0), 56 (29%) are all monomer (score=5), and the remaining 65 (33%) are distributed among the four intermediate phenotypes. The table also shows that the phenotype scores distribution depends strongly on the identity of the substituting residues. Mutations that introduce Ile have the lowest average score of 1.50, whereas mutations that introduce Pro have the highest average score of 3.89. A chi-squared statistic was used to evaluate the significance of these phenotype distributions, to identify the similarities among the different substituting residues, and to pool similar residues (Table 3.2). As indicated in Table 3.1, data from substitutions to Phe, Met, Ile, Leu, and Val are grouped in pool 1, and data from substitutions to Cys, Trp, Ala, Thr, Gly, and Ser are grouped in pool 2. The chi-squared analysis indicates that the distributions for Phe, Met, Ile, Leu, or Val substitutions correspond closely to pool 1 ( $P > 0.28$ ), whereas the distributions for Cys ( $P = 0.032$ ) and Gly ( $P = 0.065$ ) match poorly to this pool. Similarly, the distributions for Cys, Trp, Ala, Thr, Gly, and Ser substitutions correspond well to pool 2 ( $P > 0.45$ ), whereas the distributions for Phe ( $P = 0.19$ ) and Val ( $P = 0.14$ ) match poorly.

**Table 3.1** Phenotype distributions for each substituting residue and for pools

substituting residue	mutagenesis phenotype						$N_{res}$	phenotype mean
	0	1	2	3	4	5		
Phe	12	1	1	0	1	5	20	1.60
Met	5	6	1	1	0	3	16	1.63
Ile	12	0	1	0	0	5	18	1.50
Leu	5	1	3	0	0	5	14	2.29
Val	10	3	1	1	0	5	20	1.65
Cys	7	0	2	1	3	2	15	1.93
Trp	3	0	1	0	2	1	7	2.14
Ala	5	2	2	2	4	3	18	2.39
Thr	5	1	2	0	6	4	18	2.72
Gly	5	0	1	3	0	5	14	2.57
Ser	2	1	4	2	2	5	16	3.00
Pro	3	1	0	0	2	13	19	3.89
pool							$N_{pool}$	
All	74	16	19	10	20	56	195	2.27
Pool 1 <sup>a</sup>	44	11	7	2	1	23	88	1.70
Pool 2 <sup>b</sup>	27	4	12	8	17	20	88	2.50

<sup>a</sup> Pool 1: sum of distributions for Phe, Met, Ile, Leu, Val

<sup>b</sup> Pool 2: sum of distributions for Cys, Trp, Ala, Thr, Gly, Ser

(Adapted from Sulistijo and MacKenzie 2006)

**Table 3.2 Phenotype distributions for each substituting residue compared with pools**

residue	Pool 1: Phe Met Ile Leu Val				Pool 2: Cys Trp Ala Thr Gly Ser			
	entire pool		excluding self <sup>a</sup>		entire pool		excluding self <sup>a</sup>	
	$\chi^2$	P	$\chi^2$	P	$\chi^2$	P	$\chi^2$	P
Phe	3.11	0.80	5.97	0.43 <sup>b</sup>	8.78	0.19		
Met	7.46	0.28	13.2	0.040	19.1	0.0040		
Ile	3.78	0.71	5.04	0.54	12.2	0.059		
Leu	3.93	0.69 <sup>c</sup>	6.01	0.42 <sup>d</sup>	5.57	0.47 <sup>e</sup>		
Val	0.94	0.98	1.57	0.95	9.69	0.14		
Cys	13.8	0.032			2.49	0.87	3.44	0.75
Trp	17.2	0.0085			1.92	0.93	2.18	0.90
Ala	19.5	0.0034			1.65	0.95	2.90	0.82
Thr	26.4	1.9×10 <sup>-4</sup>			3.64	0.73	5.39	0.49
Gly	11.9	0.065			5.80	0.45	7.71	0.26
Ser	18.4	0.0054			3.55	0.74	5.15	0.52
Pro	20.1	0.0027			16.9	0.0096		

Phenotype distributions  $res_i$  for each residue on the left are compared with phenotype distributions  $pool_i$  for the two indicated pools of mutagenesis data using a chi-squared test:

$$\chi^2 = \sum_i \frac{\left( res_i \sqrt{N_{pool}/N_{res}} - pool_i \sqrt{N_{res}/N_{pool}} \right)^2}{N_{res} + N_{pool}} \quad \text{where} \quad N_{res} \equiv \sum_i res_i \quad N_{pool} \equiv \sum_i pool_i$$

The distributions  $res_i$  and  $pool_i$  are given in the top and bottom parts of Table 1, respectively. P values are calculated from  $\chi^2$  using the incomplete gamma function and 6 degrees of freedom (the number of bins in the phenotype distributions).

<sup>a</sup> In this case, the pool distribution excludes the substituting residue being tested against the pool; e.g., to calculate  $\chi^2$  for the Phe distribution against Pool 1 (excluding self), the pool consists of the sum of the distributions for Met, Ile, Leu, and Val (but not Phe).

<sup>b</sup> P values calculated using 5 degrees of freedom because the distributions  $res_i$  and  $pool_i$  both contain an empty bin ( $i = 4$ ), lowering the effective number of bins.

<sup>c</sup> Including the seven wild type Leu residues as phenotype 0 raises this P value to 0.88

<sup>d</sup> Including the seven wild type Leu residues as phenotype 0 raises this P value to 0.71

<sup>e</sup> Including the seven wild type Leu residues as phenotype 0 lowers this P value to 0.16

(Adapted from Sulistijo and MacKenzie 2006)

Reapplying the chi-squared test with the substituting residues being tested excluded confirms the similarities between the phenotype distributions of the substituting residues with their corresponding pools (Table 3.2, “excluding self”). The data still indicate that the distributions for Phe, Ile, Leu, and Val match the remaining four residues in pool 1 ( $P > 0.42$ ), though the correspondence for the Met distribution is weak ( $P = 0.04$ ). The distributions for Cys, Trp, Ala, Thr, Gly, and Ser match the pool of the remaining five residues in pool 2 ( $P > 0.26$ ). The  $P$  values for the Leu distribution suggest that it can be grouped to either pool 1 or pool 2. However, the Leu distribution is biased against low scores because seven out of 21 residues in the wild type sequence of the BNIP3 TMD are Leu, and hydrophobic substitutions at these positions have a mean score of 0.42. If the wild type Leu are considered as having a phenotype score of 0, the chi-squared test indicates that the Leu distribution corresponds more strongly to pool 1 and more weakly to pool 2 (see footnote to Table 3.2).

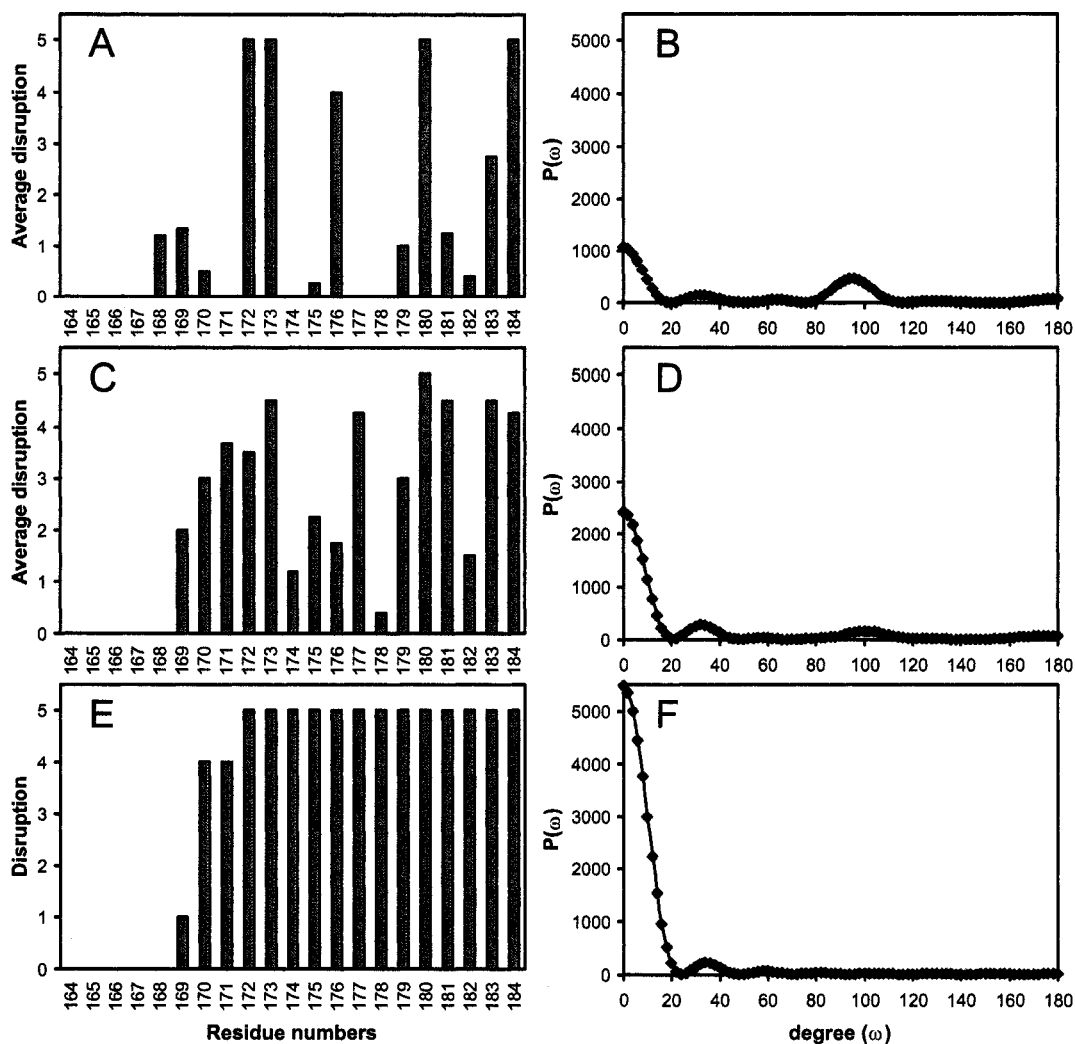
The chi-squared analysis clearly indicates that mutations that introduce pool 1 residues have quantitatively different effects from those that introduce pool 2 residues. The same analysis demonstrates that the distribution for mutations to Pro cannot be grouped into either pool because the Pro distribution is significantly different from pool 1 ( $P = 0.0027$ ) and pool 2 ( $P = 0.0096$ ). It is also interesting to note that even though no chemical information was used in this analysis, pool 1 comprises the five most hydrophobic amino acids in the GES hydrophobicity scale (Engelman et al. 1986), and five of the six most hydrophobic residues in a biological hydrophobicity scale (Hessa et al. 2005) and in the Kyte-Doolittle hydrophobicity scale (Kyte and Doolittle 1982).



### **3.7 Positional effects of mutations implicate six residues in BNIP3 transmembrane domain dimerization**

The library of BNIP3 TMD mutants (Figure 3.8) indicates that there is positional variation in the observed phenotypes over the TMD sequence. Of the 42 mutants identified over residues 164 to 168, 38 mutants have phenotype score of 0, two have a score of 1, and two have a score of 2. However, the phenotype scores are more distributed for the 153 mutants that are spread over residues 169-184 as 86 mutants (56%) have phenotype scores of 3 or higher, and 56 (37%) have “all monomer” phenotype. These positional distributions suggest that residues 164-168 are not involved in the stability or specificity of the BNIP3 TMD dimer, but the region that covers residues 169-184 is important for this self-association. Thus, the residues that participate in BNIP3 TMD dimerization can be identified by considering the average effects of mutations at residues 169-184.

The average disruption caused by mutations to residues in pool 1 varies strongly across the BNIP3 TMD sequence: seven positions have an average disruption score of 1 or less, and six positions (residues 172, 173, 176, 180, 183, and 184) have an average disruption score of 2.5 or more (Figure 3.9 A). The sensitivity of dimerization to replacements at these six positions indicates their importance to BNIP3 TMD dimerization. In addition, the spacing between these sites, every three or four residues, suggests that they are mapped on the same face of the BNIP3 TMD helix. For comparison, the variation of average disruption at positions 169-184 introduced by mutations to residues in pool 2 is less striking as only one position (residue 178) has an



**Figure 3.9 Positional dependence of disruption caused by substitution mutations**

The average phenotype at each position for substitutions to pool 1 (A), pool 2 (C) residues and by substitution to Pro (E) are presented. Fourier transform power spectra are generated from the average positional data using the formula:

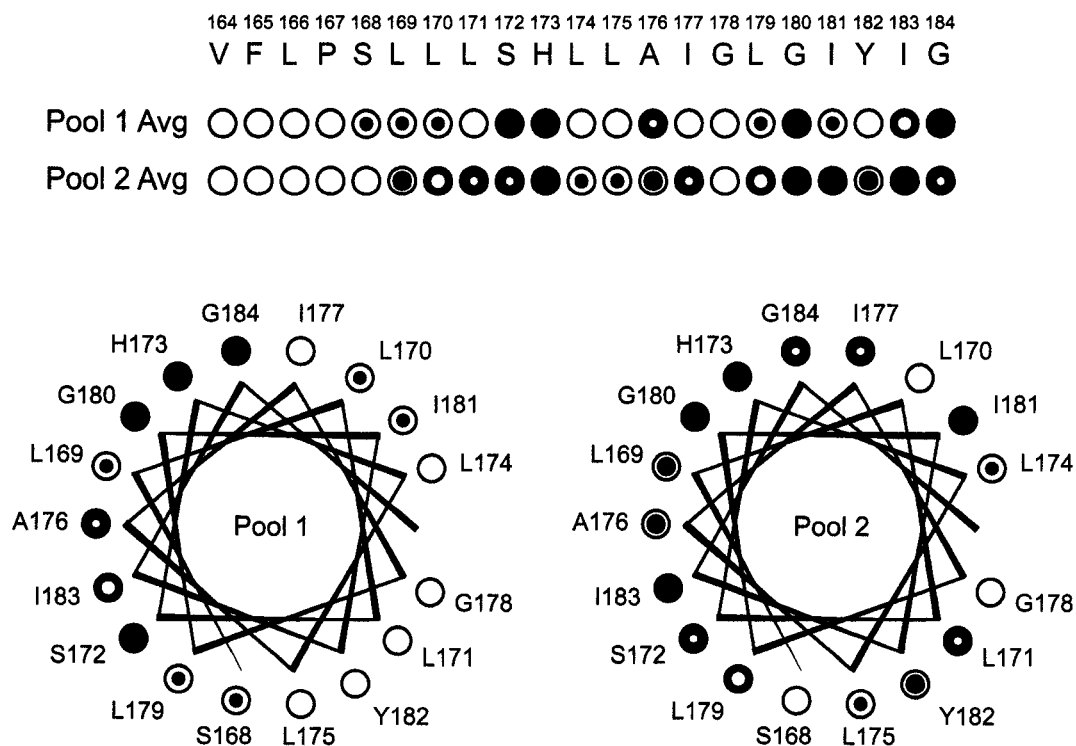
$$P(\omega) = [A(\omega)]^2 + [B(\omega)]^2$$

where  $A(\omega) = \sum V \cos(k\omega)$ ,  $B(\omega) = \sum V \sin(k\omega)$ , and  $V$  = average disruption at position  $k$ .

The pool 1 data exhibit a strong  $\alpha$ -helical repeat and a significant peak at  $95^\circ$ , suggesting that the BNIP3 transmembrane domain dimerizes with a right-handed crossing. The pool 2 data show weak periodicity, and the Pro substitutions are disruptive at interfacial and non-interfacial positions within the interacting region (res 172-184). (Adapted from Sulistijo and MacKenzie 2006).

average disruption score of less than 1, and ten positions have an average disruption score of larger than 2.5 (Figure 3.9 C). We use the average disruption of pool 1 residues to distinguish between the residues that are located at the dimer interface (Ser172, His173, Gly180, Gly184 have average scores=5; Ala176 has average score=4; Ile183 has average score=2.75) from those that are located away from the interface (residues with average scores <1.5).

We can make inferences about the relative orientation of the two BNIP3 helices from the Fourier transform power spectrum of the average disruption scores. The spectrum for the pool 1 average disruption shows a significant peak of 95° (Figure 3.9 B), while that for pool 2 average disruption shows a weak peak at 31° and a weaker peak at 101° (Figure 3.9 D). A Fourier transform power spectrum of interfacial positions on perfectly parallel  $\alpha$ -helices would show a peak at 100° (3.6 residues per turn periodicity). Helix-helix interactions with a left-handed crossing, such as leucine zippers, would show a peak at >100°, whereas those with a right-handed crossing, such as the GpA TMD dimer (Lemmon et al. 1992), would show a peak at <100°. The peak at 95°, or having interacting residues separated by 3.79 residues, suggests that the BNIP3 TMD helices have a right-handed crossing. Figure 3.10 shows the distribution of average disruption of pool 1 and pool 2 on a canonical BNIP3 TMD helix. The superiority of pool 1 average to distinguish interfacial from non-interfacial residues strongly indicates that the BNIP3 TMD dimer has a right-handed crossing.



**Figure 3.10 Projection of the average disruptions on BNIP3 transmembrane domain helix**

Average disruptions for substitutions to pool 1 and pool 2 residues are presented on a canonical helical wheel diagram of the BNIP3 transmembrane domain sequence after residue 167. Projection of pool 1 average maps sensitive residues (scores > 3) on one face of the helix, creating an interacting face for dimerization. Projection of pool 2 average, on the other hand, maps sensitive residues around the helix. (Adapted from Sulistijo and MacKenzie 2006).

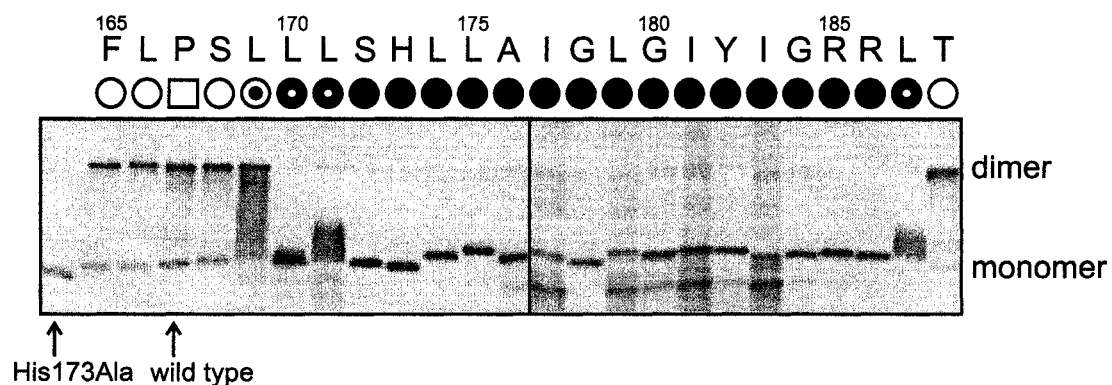
The residues that are important for dimerization map to the same face of a helical wheel representation of the BNIP3 TMD sequence, but so does Leu169, which is not critical for the TMD self-association (average disruption score=1.33). However, consistent with the idea that BNIP3 TMD helices cross at an angle, it is likely that Leu169 is away from the point of closest approach between the two helices, and hence is less important for BNIP3 TMD dimerization.

### **3.8 Mutations to proline map the transmembrane domain region important for dimerization**

Although substitutions to Pro are disruptive at most positions and therefore do not help map the interface, the Pro scanning data can be used to identify the BNIP3 TMD region that is important for dimerization. Figures 3.8 and 3.9 E indicate that substitutions to Pro at positions 165-168 are non-disruptive, substitutions at positions 169-171 have intermediate disruption phenotypes, and substitutions at residues 172-184 are completely disruptive. These data correspond well with the earlier discussion in section 3.7, which states that residues 164 to 168 are not involved in the BNIP3 TMD dimerization, and that residues 172 and 184 are the first and the last interfacial residues. These disruptive phenotypes may be caused by Pro allowing the BNIP3 TMD helix to kink or bend in the SDS micelles. A kink at the region that is not involved in dimerization does not alter the conformation of the BNIP3 TMD region that makes helix-helix contacts, thus allowing the TMD helices to associate. However, a backbone kink within the region that is involved in dimerization could helix-helix contacts whether the substitution were at or away from the dimer interface. Such alteration will prevent proper helix-helix packing, and thus only monomeric SN-BNIP3TMD can be observed with

SDS-PAGE. Therefore, the region in which substitutions to Pro are disruptive is the important region for TMD helix-helix association, and I suggest that Pro scanning mutagenesis should be used by others to map helix-helix interaction regions.

Data from substitutions to Pro allowed me to identify the BNIP3 TMD region that is important for dimerization, and the same approach can also be used to determine if residues located at the carboxy terminus of the TMD are also part of this interacting region. Figure 3.11 shows that substituting Arg185 and Arg186 with Pro cause complete disruption of the BNIP3 TMD dimer, substituting Leu187 significantly reduces dimerization, but substituting Thr188 has no effects on dimerization. These results suggest that the important region for BNIP3 TMD dimerization extends to residue 187 and imply that residues after position 188 are not involved in the TMD self-association. Figure 3.11 also shows that substitutions to Pro at positions 176, 177, 179, 180, 181, 183, and 184 display two monomer bands on Western blot. The second monomer band is probably a second conformation of the SN-BNIP3TMD mutants because an  $\alpha$ -helix containing Pro near a small residue can adopt either a more extended conformation or a bent conformation (Jacob et al. 1999). Taken as a whole, the Pro scanning data indicate that residues 169-187 comprise the interacting region for BNIP3 TMD dimerization.



**Figure 3.11 Phenotypes for Pro scanning substitutions in the BNIP3 transmembrane domain sequence**

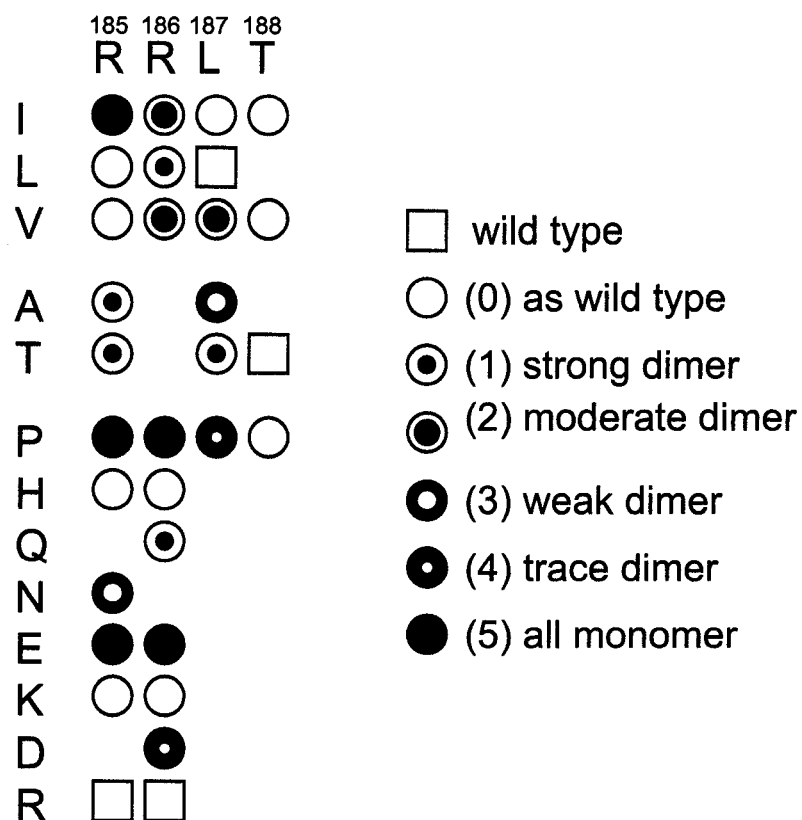
The effects of substitutions to Pro at residues 165-188 are presented. His173Ala mutant is included to indicate the position of the monomer band. Substitutions at residues 169-187 are significantly disruptive, whereas substitutions at residues 165-168 and 188 have wild type-like phenotype, suggesting that the interacting region for BNIP3 transmembrane domain dimerization covers residues 169-187. Substitutions at some positions display a second, faster-migrating monomer band, which may be a second conformation of the SN-BNIP3TMD fusion protein in SDS micelles.

### **3.9 The carboxy-terminal tail of BNIP3 is not critical for dimerization**

The involvement of residues located at the carboxy-terminus of the BNIP3 TMD in self-association can also be assessed by generating a series of truncation mutations. Analysis of these truncated mutants indicates that removing residues after Thr188 and Ser190 do not affect dimerization because these mutants display wild type-like phenotype on SDS-PAGE. Hence, it can be concluded that residues 189-194 are not important for BNIP3 TMD dimerization. However, truncation after Arg185 causes a complete disruption of the BNIP3 TMD dimer, indicating that residues 186-188 are needed for stability of the TMD dimer. This result also supports the conclusion from the Pro scanning mutagenesis, which suggests that residues 185-187 participate in the BNIP3 TMD self-association.

A limited set of mutations at residues Arg185, Arg186, Leu187, and Thr188 show that mutations to the pool 1 and pool 2 residues at these positions do not cause a major disruption of the BNIP3 TMD dimer, with the exception of Arg185Ile, thus indicating that none of these positions is directly involved in the association (Figure 3.12). The data also suggest that Thr188 is not part of the interacting region because all three substitutions at this position, including mutation to Pro, show wild type-like phenotypes. Substituting Arg185 or Arg186 to positively charged residues (Lys or His) does not affect BNIP3 TMD dimerization, but mutations to negatively charged residues Asp or Glu significantly reduce dimerization, probably because these mutants can perturb BNIP3 TMD interactions with the negatively charged head groups of the SDS micelles.





**Figure 3.12 Single substitution mutations at residues 185-188 and their corresponding phenotypes**

Single substitution mutants generated at four positions carboxy terminal of the predicted transmembrane domain sequence (res 185-188) are presented. The residues are listed on top and the substituting amino acids on the left. The same symbols and scores as indicated in Figure 3.7 are used to represent dimerization phenotypes of these mutants. The phenotype descriptions of for the symbols and scores are shown on the right.

### **3.10 Hydrophobic substitutions identify the BNIP3 transmembrane domain dimer interface**

The mutagenesis data have indicated that Pro scanning mutagenesis can map the important region for BNIP3 TMD dimerization, and that average disruption of the pool 1 residues can identify the residues that participate in the association. It would be interesting to know if scanning mutagenesis with the individual residues in pool 1, or pool 2, would be sufficient to identify these interacting residues. A potential problem that may arise from this approach is when the substituting residues are present or enriched in the wild type sequence. For the BNIP3 TMD, the problematic substituting residue is Leu because it is present at seven sites in this sequence. I used a linear regression analysis to determine the correlation between the phenotype score for a given substituting residue at each position in the BNIP3 TMD sequence to the pool 1 or pool 2 mean at that position. The correlation coefficients (Table 3.3) show that fits for Phe, Ile, Leu, and Val have strong correlations to the pool 1 mean ( $R^2 \sim 0.90-0.96$ ). The fits for Met, which show weaker correlation in the chi-squared analysis, also show slightly weaker correlation to the pool 1 mean ( $R^2 = 0.80$ ). Similarly, the fits for Cys, Ala, Thr, Gly, and Ser show strong correlations to pool 2 mean and weak correlations to the pool 1 mean. The fit for Trp suggests a stronger correlation to pool 2 than to pool 1. However, the fit was obtained with limited number of substitutions to Trp, and thus, such a conclusion is not warranted. Because the average disruption of pool 1 residues identifies the residues involved in BNIP3 TMD dimerization, the good correlations between the pool 1 residues and the pool 1 mean indicate that scanning mutagenesis with any of these residues would distinguish the interfacial residues from the non-interfacial residues.

**Table 3.3**      **Correlation coefficients for regression analyses between phenotype for a given substituting residue and the pool phenotype means at each position along the sequence**

<b>residue</b>	<b>N</b>	<b>R<sup>2</sup> versus Pool 1 mean</b>	<b>R<sup>2</sup> versus Pool 2 mean</b>
Phe	20	0.914	0.362
Met	16	0.795	0.454
Ile	18	0.949	0.373
Leu	14	0.958	0.364
Val	20	0.895	0.245
Cys	15	0.322	0.763
Trp	7	0.491	0.813
Ala	18	0.469	0.912
Thr	18	0.132	0.693
Gly	14	0.188	0.840
Ser	16	0.359	0.920
Pro	19	0.115	0.422

(Adapted from Sulistijo and MacKenzie 2006)

### **3.11 Small residues Ala176, Gly180, and Gly184 promote close helix-helix packing through a tandem GxxxG motif**

Three small residues are identified as important for BNIP3 TMD dimerization (Ala176, Gly180, and Gly184), and they are positioned such that they form a tandem motif (McClain et al. 2003; Kim et al. 2004): a GxxxG motif formed by Gly180 and Gly184, and a variant AxxxG motif (Senes et al. 2004) formed by Ala176 and Gly180. The mutagenesis data indicate that only small residues are tolerated at these positions (Figure 3.8). Subtle Gly180Ala and Gly184Ala substitutions that add a methyl group at these positions can completely or significantly disrupt BNIP3 TMD dimerization, suggesting that only residues as small as Gly are tolerated at positions 180 and 184. Replacement of Ala176 with a smaller Gly does not affect dimerization, but substitutions to larger residues show more variations: substitutions to large aliphatic residues completely disrupt dimerization, whereas substitutions to Met or Cys show little or no effects, respectively. In general, these data suggest that small residues are necessary at the dimer interface, probably to allow close packing interactions between the BNIP3 TMD helices. The fact that any residues larger than Gly are not accepted at positions 180 and 184 suggests that the closest point for helix-helix contact is likely located at or between these residues.

### **3.12 Most large aliphatic residues flanking the tandem GxxxG motif are not important for dimerization**

It has been reported previously that large aliphatic amino acids, particularly the  $\beta$ -branched residues, are frequently found flanking the small residues in the GxxxG motif, and they are expected to play a role in the stability of TMD helix-helix associations (Senes et al. 2000). The mutagenesis data for BNIP3 TMD dimerization do not suggest a uniform importance for residues flanking Ala176, Gly180, or Gly184 as substitutions at Leu175, Ile177, and Leu179 show little effects on BNIP3 TMD dimerization, substitutions at Ile181 seem to modulate dimerization to some degree, but substitutions at Ile183 can completely disrupt or significantly reduce BNIP3 TMD dimerization (Figure 3.8). These data suggest that Ile183 can be involved in BNIP3 TMD dimer stability, probably by providing intermonomer van der Waals interactions.

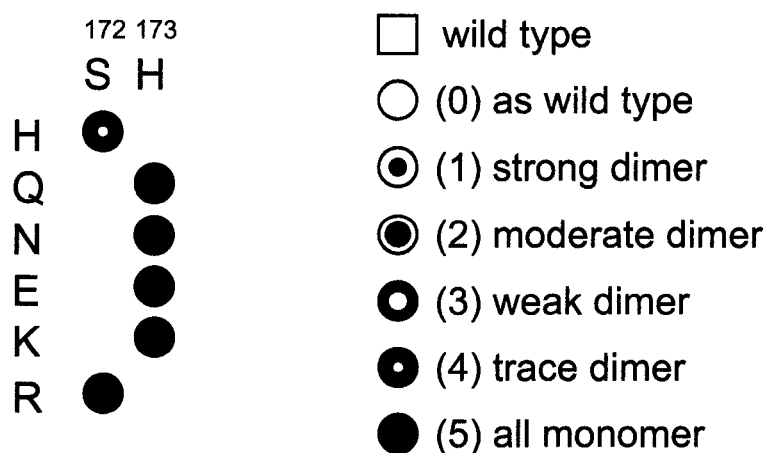
### **3.13 His173 and Ser172 participate in the BNIP3 transmembrane domain dimer stability by making intermonomer hydrogen bonds**

Complete disruption of the SN-BNIP3TMD dimer caused by replacements of His173 to many amino acids, including to nearly isosteric Phe, demonstrates the significance of His173 in dimer stability (Figure 3.8). The involvement of polar residues as mediators for inter-monomer hydrogen bonds in TMD helix-helix associations has been documented previously (Choma et al. 2000; Zhou et al. 2000; Gratkowski et al. 2001; Zhou et al. 2001; Partridge et al. 2002; Dawson et al. 2003), so it is reasonable to propose that His173 participates in the BNIP3 TMD dimer stability by mediating inter-

monomer hydrogen bonds. The fact that His173Trp mutation shows 50% dimer at the protein concentration chosen for this analysis suggests that Trp can form an inter-monomer hydrogen bond through its  $\epsilon$  indole NH in the same manner as His173 interacts with another monomer through its side chain  $\epsilon$  imidazole NH. This also indicates that another side chain, or a backbone carbonyl, takes part in accepting the inter-monomer hydrogen bonding because Trp cannot accept a hydrogen bond (although His can accept a hydrogen bond).

The disruptive phenotypes for Ser172 substitutions to hydrophobic amino acids are consistent with this residue providing a hydrogen bond accepting group for the inter-monomer interaction with His173. Substitution to Thr, which also contains a  $\gamma$  oxygen, shows wild type-like phenotype. The importance of the  $\gamma$  oxygen is confirmed by the substitution to Val, which is isosteric to Thr but abolishes dimerization. Substitution to Ala, which has a smaller side chain and lacks the  $\gamma$  oxygen of the wild type Ser, significantly reduces SN-BNIP3TMD dimerization. Based on the mutagenesis data of Ser172 and His173, I propose that these residues contribute to the stability of the BNIP3 TMD dimer by mediating inter-monomer hydrogen bonds.

To examine whether other polar residues can also mediate inter-monomer hydrogen bonding, I generated polar substitutions at positions 172 and 173 (Figure 3.13). Substitution of His173 to Lys causes a complete disruption of the BNIP3 TMD dimer, probably because the long Lys side chain causes clashes with the other BNIP3 TMD monomer.

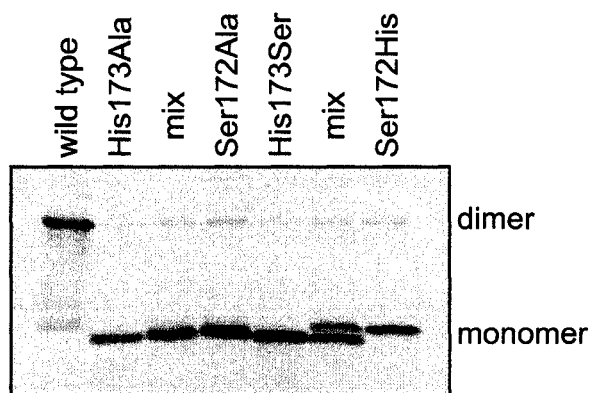


**Figure 3.13 Substitution mutations at residues Ser172 and His173 to polar amino acids and their corresponding phenotypes**

Single substitution mutations to polar amino acids at Ser172 and His173 and their corresponding phenotypes are presented. The replacing residues are listed on the left. The same symbols and scores as indicated in Figure 3.7 are used to represent dimerization phenotypes of these mutants. The phenotype descriptions of for the symbols and scores are shown on the right.

The disruptive phenotypes resulting from substitutions to Gln, Asn, and Glu are somewhat surprising, particularly since these residues have been shown to drive SDS-PAGE resistant oligomerization of designed TMDs (Gratkowski et al. 2001; Zhou et al. 2001). Replacement of Ser172 with a His significantly reduces dimerization, and mixture of His173Ser and Ser172His mutants, which ideally could still provide a pair of hydrogen bonds for the His173Ser-Ser172His heterodimer, does not show an increased dimer band intensity compared to the His173Ser or Ser172His mutants alone (Figure 3.14). In fact, the two mutants migrate separately on SDS-PAGE suggesting that these mutants may interact differently with the SDS micelles, and therefore the disruptive phenotypes caused by mutations to polar residues may be modulated by the variation of the BNIP3 TMD mutants' behavior in the SDS micelles. To test whether the presence of a single inter-monomer hydrogen bond is sufficient to stabilize BNIP3 TMD dimer, I mixed His173Ala mutant with Ser172Ala mutant and resolved the mixture with SDS-PAGE. The resulting Western blot does not display an increase in the apparent dimer-to-monomer ratio on SDS-PAGE compared to the His173Ala or Ser172Ala mutants (Figure 3.14), implying that a pair of inter-monomer hydrogen bonds is necessary to stabilize BNIP3 TMD dimerization in SDS micelles.





**Figure 3.14 Ser172 mutants and His173 mutants do not form heterodimer in SDS micelles**

Mixture of Ser172Ala and His173Ala mutants (1:1) or Ser172His and His173Ser (1:1) does not display an increase in dimer band intensities even though these mutants could facilitate the formation of one or two inter-monomer hydrogen bonds, respectively. The wild type sequence of the SN-BNIP3 TMD and a monomeric mutant His173Ala are included to indicate the positions of the dimer and monomer bands.

### **3.14 Interactions that stabilize BNIP3 transmembrane domain dimerization are inter-dependent**

Saturation mutagenesis data in Figure 3.8 indicate that single substitutions that remove the tandem GxxxG motifs but leave the inter-monomer hydrogen bonding capabilities of His173 and Ser172 intact can abolish BNIP3 TMD dimerization. Small-to-large substitutions at Ala176, Gly180, and Gly184, which can cause steric clashes between monomers, can significantly reduce or cause complete disruption of the BNIP3 TMD dimer even though Ser172 and His173 are intact. These results suggest that the tandem GxxxG motifs not only allow a pair of BNIP3 TMD helices to form close packing interactions, but also provide a context for inter-monomer hydrogen bonds to form between His173 and Ser172. Thus, mutations that remove the proper context for inter-monomer hydrogen bond formation will perturb BNIP3 TMD dimerization.

In a reversed situation, the tandem GxxxG motifs are not sufficient for dimerization because substitutions that eliminate inter-monomer hydrogen bonds but leave the tandem GxxxG motifs intact can also completely disrupt BNIP3 TMD dimerization. Substitutions at Ser172 and His173 that do not introduce steric clashes or disrupt the packing mediated by the GxxxG motif, such as Ser172Ala or His173Ala, are sufficient to abrogate dimerization. In addition, substitutions at Ile183 can significantly reduce BNIP3 TMD dimerization even though neither the tandem GxxxG motifs nor the presume inter-monomer hydrogen bonding partners are not affected. Though the mutagenesis data are not sufficient to confidently define the stabilizing properties of Ile183, together, these data suggest that neither the inter-monomer hydrogen bonds, the close packing interactions, nor the stabilizing properties contributed by Ile183 are

sufficient drive BNIP3 TMD dimerization, rather, these interactions inter-dependently stabilize the BNIP3 TMD self-association.

### **3.15 Disruption at non-interfacial residues correlates with changes in hydrophobicity**

Tables 3.2 and 3.3 show that mutants generated by substitutions to residues in pool 2 (Cys, Trp, Ala, Thr, Gly, and Ser) have statistically higher phenotype scores than pool 1 mutants, and Figure 3.8 indicates that substitutions to these residues can disrupt BNIP3 TMD dimerization significantly at both interfacial and non-interfacial positions. In contrast, substitutions to residues in pool 1 (Phe, Met, Ile, Leu, and Val) are disruptive only at interfacial positions, and mutations at non-interfacial positions have wild type-like phenotypes. Because the pool 2 residues are generally less hydrophobic than the residues in pool 1, it is likely that the variation in disruptive effects of substitutions at non-interfacial positions is due to the hydrophobicity of the substituting residue.

To test this hypothesis, a regression analysis that correlates the change in hydrophobicity with the dimerization score was calculated for each mutation within the sensitive region of the TMD but at non-interfacial positions (residues 169-184, Table 3.4). Mutations to Pro are not included in this analysis. The hydrophobicity scales used were the biological hydrophobicity scale (Hessa et al. 2005) and the interfacial (Wimley and White 1996) and octanol-water scale (Wimley et al. 1996) of Wimley and White (WW).

**Table 3.4**      **Correlation between hydrophobicity change and phenotype for non-interfacial mutations within the sensitive region**

<b>Scale</b>	<b>Pools 1 and 2</b>	<b>Pool 2 only</b>
	(n = 74)	(n = 38)
biological	0.498	0.527
octanol	0.413	0.276
interfacial	0.107	0.029
difference	0.476	0.419

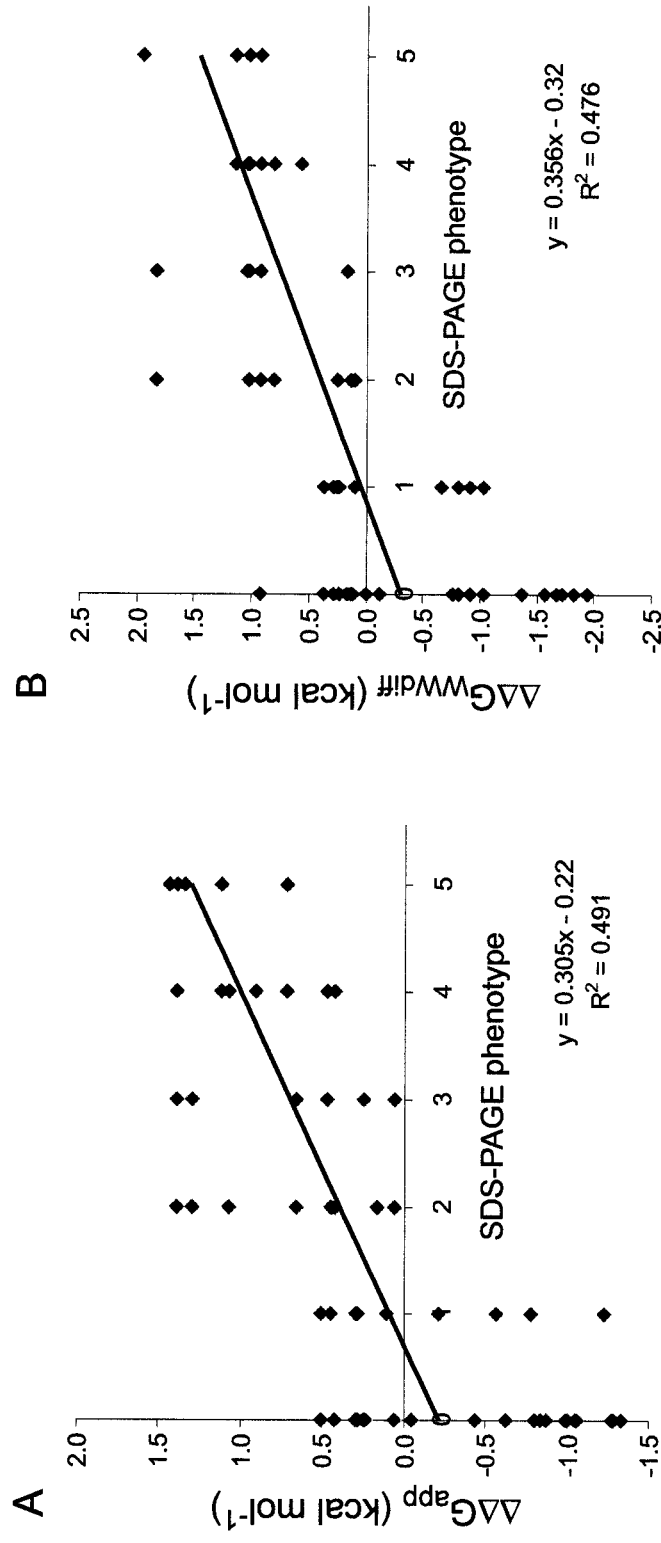
Mutations at non-interfacial positions 170, 171, 174, 175, 177, 178, 179, 181, and 182 were included in this analysis.

(Adapted from Sulistijo and MacKenzie 2006)

The biological scale is able to account for 50% of the variance in the 74 non-interfacial substitutions and 53% of the variance in the 38 pool 2 substitutions (Figure 3.15). The correlations are poor with the WW interfacial scale but reasonably good with the WW octanol-water scale. A new scale constructed by subtracting the WW octanol scale with the WW interfacial scale correlates with the mutagenesis data even better than either scale alone. From these data, I can conclude that changes in hydrophobicity can account for much of the disruptive effects introduced by mutations at non-interfacial positions for BNIP3 TMD dimerization. These data also suggest that unlike substitutions to hydrophobic residues, substitutions to slightly polar and polar amino acids cannot be used to reliably identify the interfacial residues for TMD helix-helix interaction.

### **3.16 Fourier transform infra-red spectroscopy data suggest that mutations to polar residue cause dimer disruption due to differences in solvent exposure**

One possible explanation for how a decrease in hydrophobicity can modulate the amount of observed dimer is the changes in hydrophobicity cause the TMD to unfold. Because polar residues prefer interactions with the micellar surface, introduction of a polar residue in the BNIP3 TMD, either at or away from the dimer interface, can cause the TMD to leave the hydrophobic core of the detergent micelle and expose that portion of the membrane span at the surface of the micelle while unfolding the TMD helix. This unfolding process depletes the folded monomeric state, thereby lowering the amount or observed dimer without directly altering the equilibrium between folded monomer and dimer.

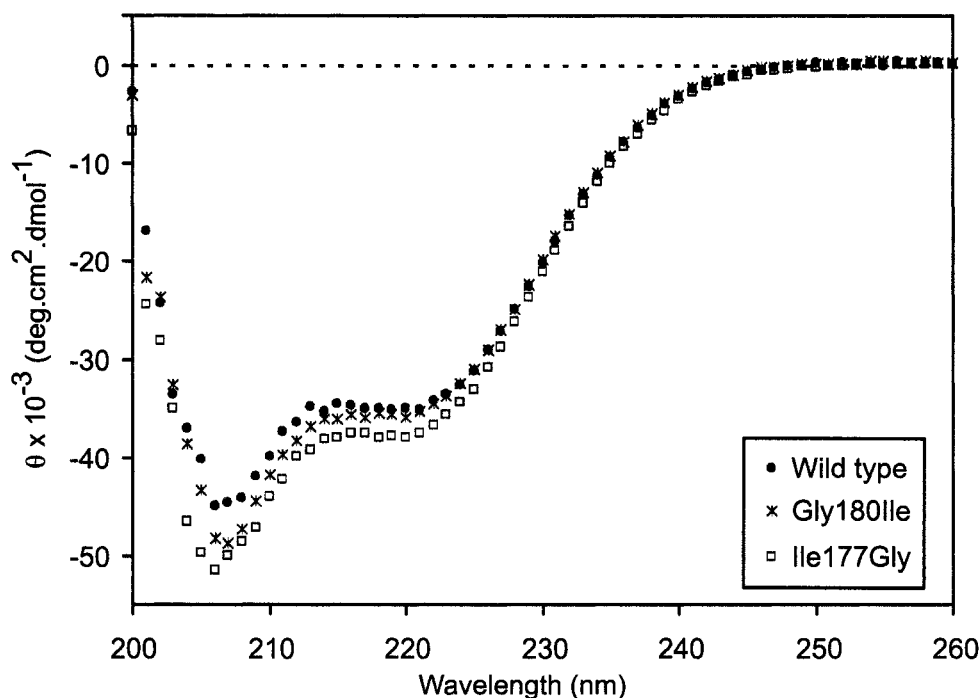


**Figure 3.15 Correlations between disruptive effects of non-interfacial mutations and loss of hydrophobicity**

The disruptive phenotypes observed on 74 non-interfacial mutations are correlated to the loss of hydrophobicity, calculated using the biological hydrophobicity scale (A) or the difference between the Wimley–White octanol and interfacial scales (B). Correlation coefficients for similar regressions are presented in Table 3.4. (Adapted from Sulistijo and MacKenzie 2006).

To confirm this hypothesis, I analyzed the secondary structure of the SN-BNIP3TMD fusion protein with wild type BNIP3 TMD sequence or its mutants using circular dichroism (CD) spectroscopy and Fourier transform infrared (FTIR) spectroscopy. The CD spectra of these fusion proteins show negative ellipticity at 220 and 208 nm, indicating that the wild type, as well as the Ile177Gly and Gly180Ile mutants, is mainly  $\alpha$ -helical (Figure 3.16). The spectra suggest that the mutants may have higher  $\alpha$ -helical content than the wild type SN-BNIP3TMD. However, due to uncertainty in protein concentrations, the CD data do not reliably distinguish the relative differences in the secondary structure between the wild type and the mutants SN-BNIP3TMD.

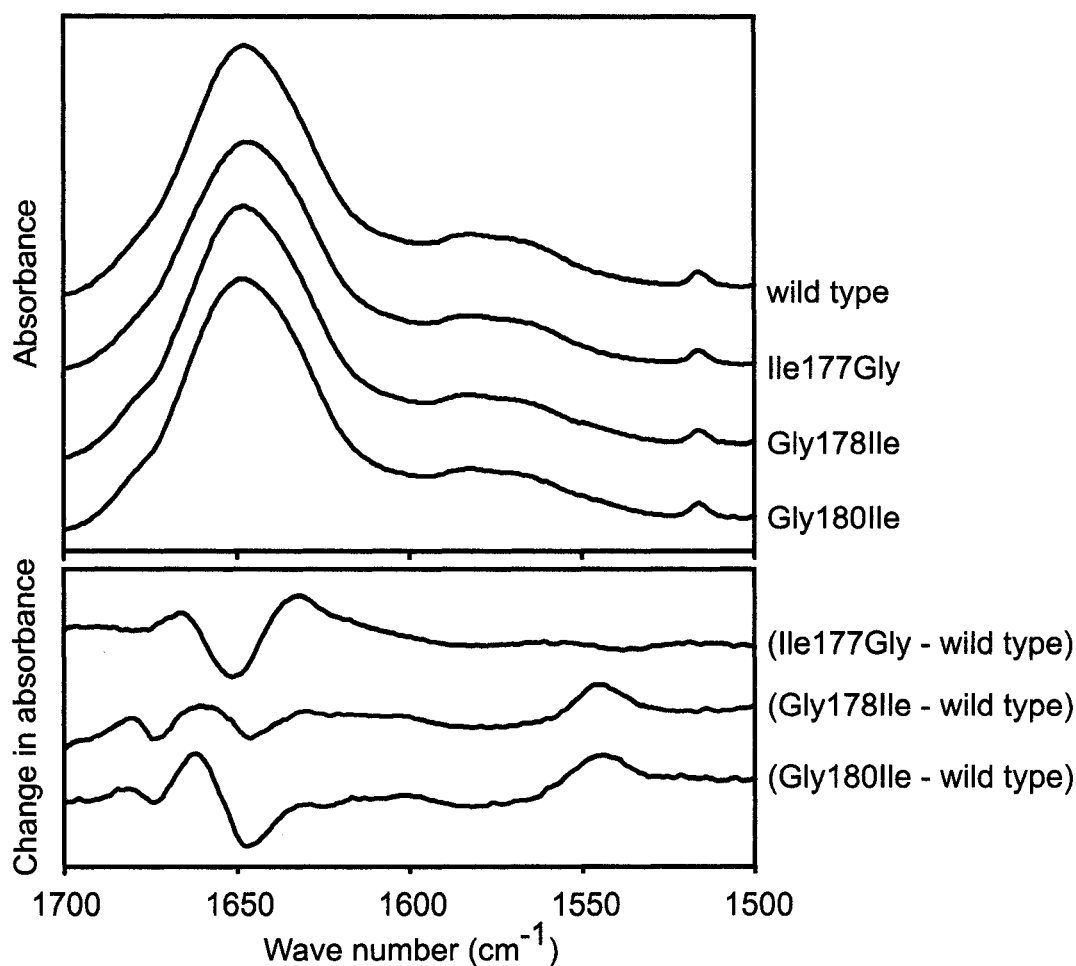
Variations in protein concentrations are not a major issue for analyzing FTIR spectra because certain characteristics, such as the intensity of the Tyr band at 1515  $\text{cm}^{-1}$ , can be used to scale the spectra. Differences in the secondary structure content of the wild type and mutant SN-BNIP3TMD can be detected by examining the differences in absorbance at the amide I region (1600-1700  $\text{cm}^{-1}$ ). However, these differences cannot be easily observed because the wild type and mutants SN-BNIP3TMD have very similar characteristics (Figure 3.17, top panel). Therefore, to allow detection of slight changes in the secondary structure, I generated difference spectra by subtracting the wild type absorbance from the mutant absorbance (Figure 3.17, bottom panel). The (Ile177Gly – wild type) difference spectrum displays an increase of turns (1666  $\text{cm}^{-1}$ ) and  $\beta$ -strand (1635  $\text{cm}^{-1}$ ) content, and a decrease in  $\alpha$ -helix content (1651  $\text{cm}^{-1}$ ), suggesting that introducing a less hydrophobic residue at a position away from the dimer interface can causes a loss of  $\alpha$ -helix in favor of an increase in turns and  $\beta$ -strand content.



**Figure 3.16 Circular dichroism (CD) spectra of wild type and selected mutants of the SN-BNIP3TMD fusion protein**

CD spectra of wild type and mutants SN-BNIP3TMD are presented. The protein samples were prepared in a buffer containing 50 mM Tris pH 8.0, 4% SDS. The spectra were acquired between 260-200 nm at 25 °C. Negative ellipticities at 220 nm and 208 nm indicate that the wild type and mutants SN-BNIP3TMD are mainly  $\alpha$ -helical. Though relative ellipticities can be used to determine changes in the  $\alpha$ -helical contents due to mutations in the BNIP3 transmembrane domain, the differences seen on the spectra may have been influenced by variations in the protein concentrations.





**Figure 3.17 Original and difference FTIR spectra for wild type and selected mutants of SN-BNIP3TMD**

The original transmission FTIR spectra for wild type and mutants SN-BNIP3TMD are presented on top, and the (mutant - wild type) difference spectra at the bottom. Differences in the secondary structure contents caused by mutations are not easily observed on the original spectra. However, the (mutant - wild type) difference spectra, which are not drawn to scale to the original spectra, show slight positive and negative intensities at 1600-1700  $\text{cm}^{-1}$  suggesting that substitutions can modify the secondary structure contents of the BNIP3 transmembrane domain. Relative differences in hydrogen/deuterium exchange rate can also be observed at 1545  $\text{cm}^{-1}$ . (Adapted from Sulistijo and MacKenzie 2006).

On the contrary, the (Gly178Ile – wild type) difference spectrum shows no significant change in  $\alpha$ -helix content even though there is a slight decrease of  $\beta$ -strand ( $1672\text{ cm}^{-1}$ ) and unordered structure ( $1645\text{ cm}^{-1}$ ), and a slight increase of turns ( $1660\text{ cm}^{-1}$ ). The (Gly180Ile – wild type) difference spectrum also shows an increase of turn ( $1662\text{ cm}^{-1}$ ) and a decrease of unordered structure ( $1646\text{ cm}^{-1}$ ), which could mask a very small negative or positive  $\alpha$ -helix intensity ( $1651\text{ cm}^{-1}$ ). Together, neither substitution to Ile indicates significant changes in  $\alpha$ -helical content relative to the wild type.

The difference spectra for the two Gly to Ile mutants show the presence of a band centered at  $1545\text{ cm}^{-1}$ . This band corresponds to amide II N-H stretch, and the intensities can change with the levels of deuterium incorporation in the protein. Because these spectra were acquired in  $\text{D}_2\text{O}$ , the intensity of this band indicates a population that is protected from amide exchange. The positive difference intensities for the more hydrophobic mutants relative to the wild type suggest that the increase in hydrophobicity causes a fraction of the TMD to be protected from hydrogen/deuterium exchange even after 20 hours incubation in  $\text{D}_2\text{O}$ . The absence, or slight negative intensity, of the  $1545\text{ cm}^{-1}$  band for the Ile177Gly mutant relative to the wild type suggests that the TMD undergoes exchange at similar rate, if not faster, than the wild type under these conditions.

## **4 Preparation of the BNIP3 transmembrane domain peptide for solution NMR structure determination**

### **4.1 Introduction**

Determination of protein structures by NMR spectroscopy require samples that are isotopically enriched with  $^{13}\text{C}$  and/or  $^{15}\text{N}$ . Chemical synthesis can provide such samples, but the reactions are not cost-effective in most cases. Protein expression in microorganisms or cultured cells offers a less expensive option to generate isotopically enriched samples, and having already an expression construct that allows high level expression in *E. coli* simplified the process of generating isotopically labeled BNIP3 TMD for solving the structure of the TMD dimer. In this chapter I present the methods for overexpression of isotopically labeled protein, isolation of the BNIP3 TMD peptide from the SN-BNIP3TMD fusion protein construct, and sample preparation for NMR spectroscopy. First, I discuss the optimization strategies employed, and then I provide a detailed description of the protocols used to prepare an NMR sample of the BNIP3 TMD peptide for structural work. These methods yield peptide with excellent purity and quality for NMR studies, suggesting that they may be suitable for the preparation of other TMD peptides.

## **4.2 Optimization of BNIP3 transmembrane domain peptide length**

The complexity in solving structures of biological molecules by NMR is correlated with the size of the molecules being studied. Solving structures of larger molecules is more challenging than solving structures of smaller molecules. Likewise, solving the structure of the BNIP3 TMD peptide dimer (~40 residues) would be less demanding than solving the structure of the BNIP3 protein dimer (~200 residues). As discussed in section 3.9, the carboxy terminal tail of the BNIP3 TMD is not necessary for dimerization, and BNIP3 TMD with truncations after residue 188 or 190 can still mediate strong dimerization of the SN-BNIP3TMD fusion protein in SDS micelles. These data suggest that truncation after residue 188 generates the shortest BNIP3 TMD version that can still mediate dimerization in detergent micelles. Therefore, in determining the structure of the BNIP3 TMD dimer, I used the TMD peptide with truncation after residue 188.

## **4.3 Optimization of protein overexpression**

Optimal conditions for SN-BNIP3TMD overexpression have been discussed in section 2.2, but there is another condition that I optimized to maximize the amount of protein yield. I needed to identify the glycerol stocks that would yield the higher levels of protein because different glycerol stocks made from the same transformation can have varying levels of protein expression though grown under the same conditions. I did this analysis by examining the expression levels of cultures grown from three different glycerol stocks that were made from colonies picked from the same transformation. The

glycerol stocks with the highest protein yield were considered the optimal, and were used to overexpress the SN-BNIP3TMD fusion protein.

## 4.4 Optimization of proteolysis conditions

The SN tag of the fusion protein is rich in positively charged residues (29 arginines and lysines) and negatively charged residues (19 aspartic acids and glutamic acids), making it a good target for proteolysis with trypsin or V8 protease, respectively. Even though recognition sites for these proteases are also present at, or close to, the predicted BNIP3 TMD, they should not be targeted for proteolysis because the TMD is embedded in the Thesit micelles used in protein purification. To confirm that the proteolysis reaction would leave the BNIP3 TMD intact, I performed a competition assay in which I mixed the proteolysis products with SN-BNIP3TMD fusion protein and looked for an appearance of a protein-peptide heterodimer band on SDS-PAGE, similar to what is shown in Figure 3.4. Analysis with products of cleavage with trypsin showed an appearance of a single protein-peptide heterodimer band on SDS-PAGE, indicating that trypsin can be used to isolate BNIP3 TMD peptide from the SN-BNIP3TMD fusion protein.

However, a more detailed analysis with reversed-phase HPLC indicated that trypsin cleavage produces two BNIP3 TMD peptide species. MALDI-TOF mass spectroscopy analysis confirmed the presence of two peptide species that differ by a single amino acid. Sequential analysis of the SN-BNIP3TMD fusion protein construct indicated that the two peptide fragments were caused by the BNIP3 TMD segment

having adjacent trypsin recognition sites: Lys152 and Lys153 (Figure 4.1, top line). Cleavage after Lys153 will produce a peptide fragment of 3831 Da (Figure 4.1, peptide A), and cleavage after Lys152 will produce a peptide fragment of 3982 Da (Figure 4.1, peptide B). Having two BNIP3 TMD peptide species that can be separated with reversed phase-HPLC is not problematic. However, it will reduce the yield for a single TMD peptide. Another issue is that I was not able to optimize the trypsin cleavage conditions to maximize the yield for one of the two fragments. Therefore, to ensure that only a single BNIP3 TMD peptide species is isolated from proteolysis with trypsin, I introduced a mutation in the BNIP3 TMD region to remove Lys152. The sequence of the mutagenesis product is shown on the second line in Figure 4.1. Trypsin cleavage with this SN-BNIP3TMD version produces mainly peptide A fragments. A minor fragment (Figure 4.1, peptide C) can be isolated with analytical reverse-phase HPLC, but the amount of this fragment can be reduced by incubating the reaction at 37 °C, adding higher concentrations of trypsin to the reaction, and letting the reaction go overnight. The reduction of the minor fragment is accompanied with an increased yield of peptide A, which is advantageous.

To determine the length for the proteolysis reaction that would maximize the yield for peptide A, I performed an analytical reaction, during which I drew out samples after 2, 4, 6, 8, 12, and 16 hours of incubation. The relative amount of peptides A and C in the product was determined by reversed-phase HPLC. I found that longer incubation increases the yield for peptide A relative to peptide C, but no significant differences in the peptide A and C ratio were observed after 12 and 16 hours of reaction, suggesting that 12 hours was the minimal incubation period for the proteolysis reaction.



I also examined if the yield for peptide A could be increased by varying the amount of trypsin added (1:40 - 1:10 Trypsin to protein weight ratio). I found that the use of higher trypsin to SN-BNIP3TMD fusion protein ratio would increase the yield for peptide A. However, when the SN-BNIP3TMD concentration in the sample is low, the use of high trypsin concentration, even at 1:4 trypsin to protein weight ratio, did not increase the yield of peptide A significantly. I redid the analysis for the efficiency of trypsin digestion by varying the concentration of trypsin added (0.05 - 0.2 mg ml<sup>-1</sup>), and found that proteolysis with trypsin is efficient at a concentration of 0.1 mg ml<sup>-1</sup>. In a typical SN-BNIP3TMD fusion protein concentration collected from CM52 column purification, which is 1 mg ml<sup>-1</sup>, this concentration corresponds to 1:10 trypsin to protein ratio.

## 4.5 Optimization of NMR sample preparation

An important step in preparing NMR samples of the BNIP3 TMD peptide was to incorporate the TMD peptide into detergent micelles. This is done by titrating detergent solution into the BNIP3 TMD peptide dissolved in organic solvent. The presence of organic solvents can alter the secondary structures of the TMD peptide. For example, alcohols or fluorinated alcohols, such as trifluoroethanol (TFE) and hexafluoroisopropanol (HFIP), can strongly induce  $\alpha$ -helical conformation (Hirota et al. 1997; Kumar et al. 2004), but they can also induce dissociation of TMD oligomers (van den Brink-van der Laan et al. 2004). Because the conformation induced by organic solvents can be artifactual, and because the aim is to solve the BNIP3 TMD dimer structure in detergent micelles, these solvents should be removed completely from an NMR sample,



which can be done by lyophilization. The sample can subsequently be redissolved in H<sub>2</sub>O or D<sub>2</sub>O.

The types of organic solvents used in sample preparation can also influence the quality of NMR sample. Samples prepared in the presence of TFE or TFE plus acetonitrile-isopropanol mixture (3:2 volume ratio), which are the solvents used in reversed phase-HPLC purification, produce <sup>1</sup>H/<sup>15</sup>N-HSQC spectra with well dispersed peaks, suggesting that the BNIP3 TMD peptide is folded under these sample conditions. Similarly, samples prepared in the presence of HFIP generated spectra with well dispersed peaks but with better sensitivity and resolution, indicating that HFIP may be the better organic solvent for use in preparing NMR sample of the BNIP3 TMD peptide. However, HFIP is not as volatile as TFE, so multiple rounds of lyophilization were needed to remove the solvent completely from the peptide-detergent mixture.

## **4.6 Preparation of BNIP3 transmembrane domain peptide for NMR spectroscopy**

### **4.6.1 Overexpression of isotopically labeled SN-BNIP3TMD fusion protein**

The protocol for overexpression of isotopically labeled SN-BNIP3TMD is similar to large scale expression of the fusion protein (described in section 2.5.1). However, instead of growing in LB medium, the culture is grown in M9 minimal medium plus 50  $\mu\text{g ml}^{-1}$  carbenicillin for selection. For overexpression of [ $U$ - $^{15}\text{N}$ ]-labeled SN-BNIP3TMD fusion protein, the medium is prepared with  $^{15}\text{NH}_4\text{Cl}$ , and for overexpression of [ $U$ - $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled fusion protein, the medium is prepared with  $^{13}\text{C}$ -glucose and  $^{15}\text{NH}_4\text{Cl}$ . I started the culture by inoculating an LB glycerol stock of *E. coli* BL21(DE3) cells that had been transformed with plasmid encoding the SN-BNIP3TMD fusion protein (see sections 2.4.3) into 100 ml of M9 minimal medium, prepared as necessary, and allowed the culture to grow overnight at 37 °C with shaking. For the overnight growth, the medium was supplemented with 1:500 culture volume of LB medium to provide nutrients during initial stage of cell growth as the cells adapt from LB (rich medium) to minimal medium. The following day, the overnight culture was diluted 1:200 in 1 L of fresh M9 minimal medium, also supplemented as necessary for isotopic labeling. No LB medium was added for this growth, but 50  $\mu\text{g ml}^{-1}$  carbenicillin was added for selection. The fresh culture was returned to 37 °C and allowed to grow until it reached  $\text{OD}_{600} \sim 1.0$ , at which point protein expression was induced with 0.5 mM IPTG. After 3 hours of incubation at 37 °C, the cells were harvested by centrifugation at 10,000 x *g* for 10 min.

#### **4.6.2 Extraction and purification of the SN-BNIP3TMD fusion protein**

To extract the SN-BNIP3TMD fusion protein, the cell pellet was resuspended in 1:20 culture volume of lysis buffer (20 mM Tris-HCl, 2 mM EDTA, pH 8.0, 0.5 mM PMSF) and then subjected to three rounds of freezing and thawing and probe sonication to lyse the cells. Next, 10 mM  $\text{CaCl}_2$  and 0.1 mg  $\text{ml}^{-1}$  hen egg white lysozyme were added, and the whole cell extract was incubated on ice for 30 min to allow digestion of the chromosomal DNA by SN. The soluble and insoluble fractions were separated with centrifugation at 12,000 x  $g$  for 20 min. The supernatant was removed with decantation, and the pellet, which contains the SN/BNIP3TM fusion protein, was treated with buffers containing different salt and detergent conditions to extract unwanted components. First, the insoluble fraction was resuspended in 1:20 culture volume of lysis buffer plus 1 M ammonium acetate, probe sonicated, and centrifuged at 12,000 x  $g$  for 10 min to separate the soluble and insoluble fractions. Then, the insoluble fraction was treated with lysis buffer containing 2% (w/v) detergent Thesit (Fluka) in the same manner as the previous treatment. To extract the SN-BNIP3TMD fusion protein, the insoluble fraction was resuspended in lysis buffer containing 1 M ammonium acetate and 2% (w/v) Thesit, probe sonicated, and centrifuged at 12,000 x  $g$  for 20 min. At this stage the SN-BNIP3TMD fusion protein is found in the supernatant. To lower the ammonium acetate and Thesit concentrations in the sample, the supernatant was dialyzed twice against 20 sample volumes of dialysis buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 100 mM ammonium acetate, 0.2% (w/v) Thesit, 0.5 mM PMSF). Aggregates resulting from the dialysis were removed by centrifugation at 12,000 x  $g$  for 20 min. Purification of the SN-BNIP3TMD protein to remove residual nucleotides and *E. coli* proteins using DE52 and

CM52 columns (Whatman) followed the same protocol described in section 2.5.1.

Typically, for one liter growth in M9 minimal medium, I would collect ~20 mg of SN-BNIP3TMD fusion protein after CM52 column purification.

#### **4.6.3 Isolation and purification of the BNIP3 transmembrane domain peptide**

The BNIP3 TMD was isolated from the SN-BNIP3TMD fusion protein by treating the protein with 0.1 mg ml<sup>-1</sup> trypsin. The reaction was incubated at 37 °C for at least 12 hours to ensure that the yield for peptide A is maximized (Figure 4.1). Next, the proteolysis product was dialyzed against water to remove soluble SN fragments and excess Thesit and to lower the ammonium acetate concentration. Prior to Thesit removal, the proteolysis product was lyophilized to complete dryness. Thesit was then removed by treating the dried product with two or three rounds of hexane wash, followed by two rounds of hexane-butanol-acetic acid (6:1:0.1) wash. Afterward, the BNIP3 TMD peptide was purified with reversed-phase HPLC, which also removes residual SN peptide fragments. The purification was performed using a phenyl column, and the TMD peptide was eluted with a gradient of acetonitrile/isopropyl alcohol (3:2), 0.1% (v/v) 2,2,2-trifluoroacetic acid (TFAA). The purity and the size of the BNIP3 TM peptide in the pooled HPLC fractions were confirmed with MALDI-TOF mass spectroscopy.

#### 4.6.4 NMR sample preparation

The purified, lyophilized BNIP3 TMD peptide was incorporated into perdeuterated dodecylphosphocholine (DPC- $d_{38}$ , Cambridge Isotope Labs) micelles by first dissolving the peptide in 1:1 mixture of HFIP and water, and then adding a DPC- $d_{38}$  solution dropwise with agitation. The total amount of DPC- $d_{38}$  added would be 128 mM (5% w/v) of monomer concentration in the final NMR sample conditions. Next, water was added to the peptide-detergent mixture to approximately 1:2 HFIP:water volume ratio. To remove organic solvents from the sample, the peptide-detergent mixture was lyophilized to complete dryness, resuspended in water, and re-lyophilized to ensure complete removal of HFIP. Sample that is free of HFIP has a white, fluffy appearance. Afterward, the dried peptide-detergent mixture was dissolved in water with 10% (v/v)  $D_2O$ . Sodium phosphate ( $NaH_2PO_4$ ) and an internal reference 3-(trimethylsilyl)-propionic acid- $d_4$ , sodium salt (TSP) were added to final concentrations of 10 mM and 0.1 mM, respectively, and the sample pH was adjusted to 5.10.

#### 4.6.5 Preparation of phospholipid stocks

Phospholipid stocks were prepared by first removing organic solvents from the solutions of deuterium-labeled 1,2-dipalmitoyl-*sn*-glycero-3-phosphatidylcholine (DPPC- $d_{75}$ , Avanti) or 1,2-dipalmitoyl-*sn*-glycero-3-phosphate (DPPA- $d_{62}$ , Avanti) by blowing argon or nitrogen gas, followed by vacuum drying with a Speedvac. Next, the dried phospholipids were resuspended in 5% (w/v) DPC- $d_{38}$  solution prepared in 10% (v/v) D<sub>2</sub>O. To ensure that the phospholipids are fully incorporated in the DPC micelles, the phospholipid-detergent mixture was incubated at 50 °C and sonicated briefly in a waterbath sonicator.

## **5 Determination of the BNIP3 transmembrane domain dimer structure by solution NMR spectroscopy**

### **5.1 Introduction**

Structure determination using solution NMR spectroscopy relies on measurements of  $^1\text{H}$ - $^1\text{H}$  nuclear Overhauser effects (NOE), which are relaxation events mediated by through-space spin-spin dipolar interactions. The magnitude of these effects is correlated to the inverse of the sixth power of the separation between two nuclei, and can provide distance measurements between atoms. Other experimentally derived restraints, such as torsion angles, residual dipolar couplings, and hydrogen bonds, can be specified to complement the distance restraints provided by NOE measurements.

In this chapter I discuss how I used data from NMR experiments to determine the structure of the BNIP3 transmembrane domain (TMD) dimer. First, I explain the method to screen optimal sample and experimental conditions, and the approach to collect data for sequential assignments of the backbone atoms, identification of the side chain resonances, and determination of the NOE correlations. Then, I discuss the process of calculating the structure of the BNIP3 TMD dimer and show how the resolution of the BNIP3 TMD dimer structure can be improved by adding supplemental information to the structure calculation.

## **5.2 Initial assessment of sample and experimental conditions for structure determination of the BNIP3 transmembrane domain**

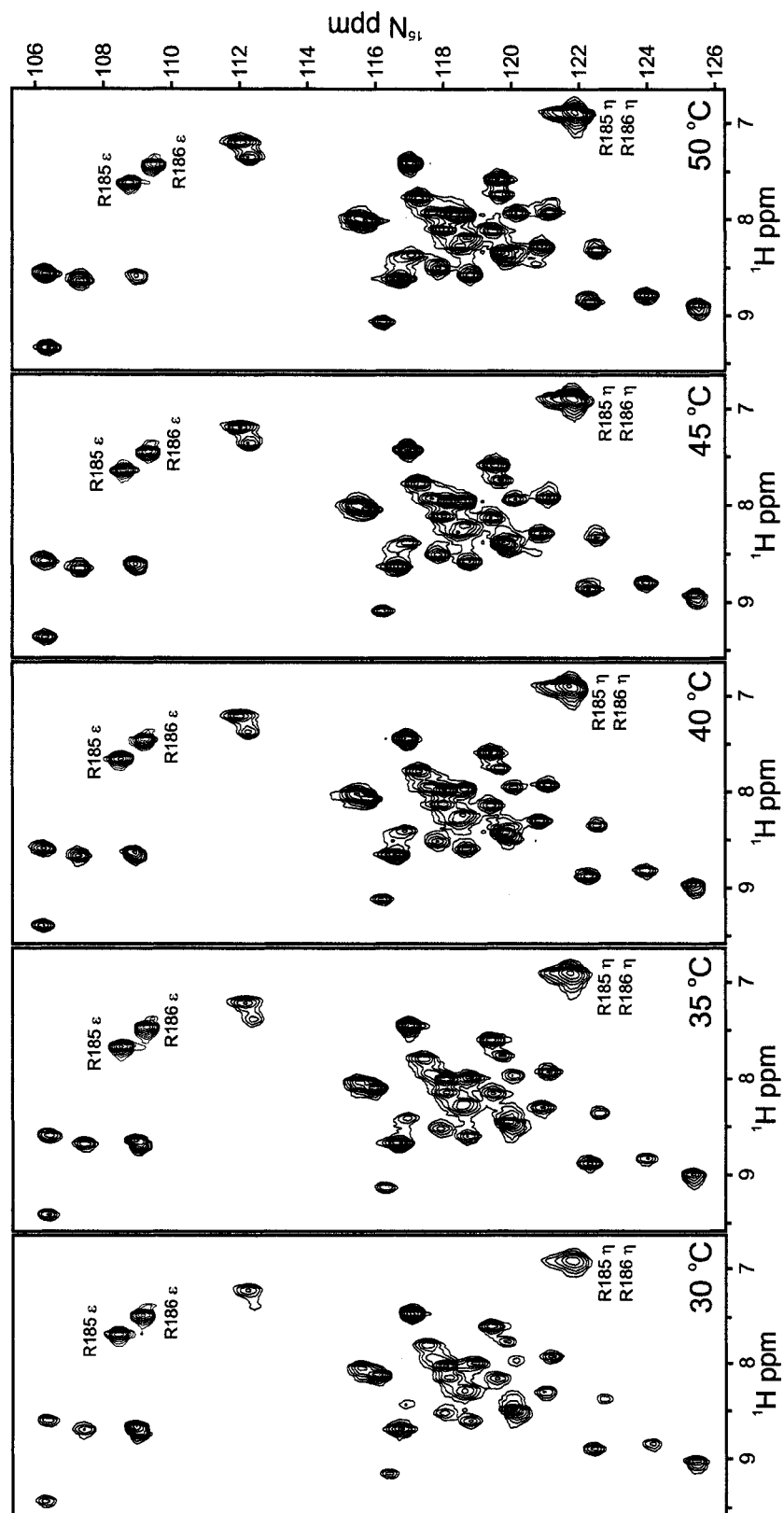
The ability to perform sequential assignment of the backbone atoms requires well resolved amide proton ( $H^N$ )-amide nitrogen (N) correlations. A 2D  $^1H/^{15}N$  heteronuclear correlation (HSQC) experiment can be utilized to search for sample conditions that meet this requirement. In addition, the  $^1H/^{15}N$ -HSQC experiment can be used to examine whether the protein of interest is properly folded because the  $H^N$  and N resonances are sensitive to local environmental conditions. The fact that the experiment can be performed on inexpensive  $^{15}N$ -labeled protein makes the  $^1H/^{15}N$ -HSQC an attractive approach to screen for optimal conditions and to evaluate whether the protein of interest is suitable for structural study with solution NMR spectroscopy.

The  $^1H/^{15}N$ -HSQC spectra that I analyzed when searching for optimal sample conditions were acquired with samples of 0.4-1.0 mM BNIP3 TMD peptide on an INOVA Varian 800 MHz, 600 MHz, or 500 MHz spectrometer. The spectra were obtained using the gradient-selected sensitivity-enhanced  $^1H/^{15}N$ -HSQC experiment of Kay and co-workers (Kay et al. 1992) as provided in the BioPack suite or as coded by my advisor, Kevin MacKenzie. Acquired data were processed with NMRPipe (Delaglio et al. 1995) and visualized with NMRDraw (Delaglio et al. 1995). The qualities that I was looking for were well resolved spectra, in which all amide protons are represented, and overlap between peaks is minimal.



To find the optimal experimental temperature for structural work, I analyzed  $^1\text{H}/^{15}\text{N}$ -HSQC spectra that were acquired at temperatures ranging between 30-50 °C. As shown in Figure 5.1, there are little changes in peak dispersion at various temperatures, indicating that there is no large structural transition over the temperature range. However, the spectra show temperature dependent line width and intensity. At elevated temperature, a few peaks show weaker intensities, probably due to exchange with solvent and not to conformational change because other peaks do not experience significant chemical shift changes. Chemical shift dispersion and relative peak intensities on the  $^1\text{H}/^{15}\text{N}$ -HSQC are maximal in spectra acquired at 40 °C. Similar to the observations seen with temperature analysis, various pH conditions, ranging from 4.0 to 7.0, caused little changes in peak dispersion, but line width showed pH dependence. As the pH is closer to neutral, some peaks show weaker intensities, which may be due to exchange with solvent. Peaks with relatively similar intensities were observed when the sample was prepared at pH ~5.10. Therefore, subsequent samples were prepared at pH 5.10, and NMR spectra were acquired at 40 °C.

To find optimal detergent conditions, two types of detergent were examined: the negatively charged sodium dodecylsulfate (SDS) which was used in the mutagenesis analysis (Chapter 3) and the zwitterionic dodecylphosphocholine (DPC). Both detergent conditions yield  $^1\text{H}/^{15}\text{N}$ -HSQC spectra with well dispersed peaks, but slight variations in the  $^1\text{H}/^{15}\text{N}$  resonances were observed, probably because of differences in the micellar head group conditions. The decision as to which detergent is the better came to the fact that DPC forms smaller micelles (50-60 molecules per micelle) than SDS (~120 molecules per micelle). Even though the two detergent conditions produce spectra with similar peak dispersion, the larger size of the SDS micelles increases relaxation rates,



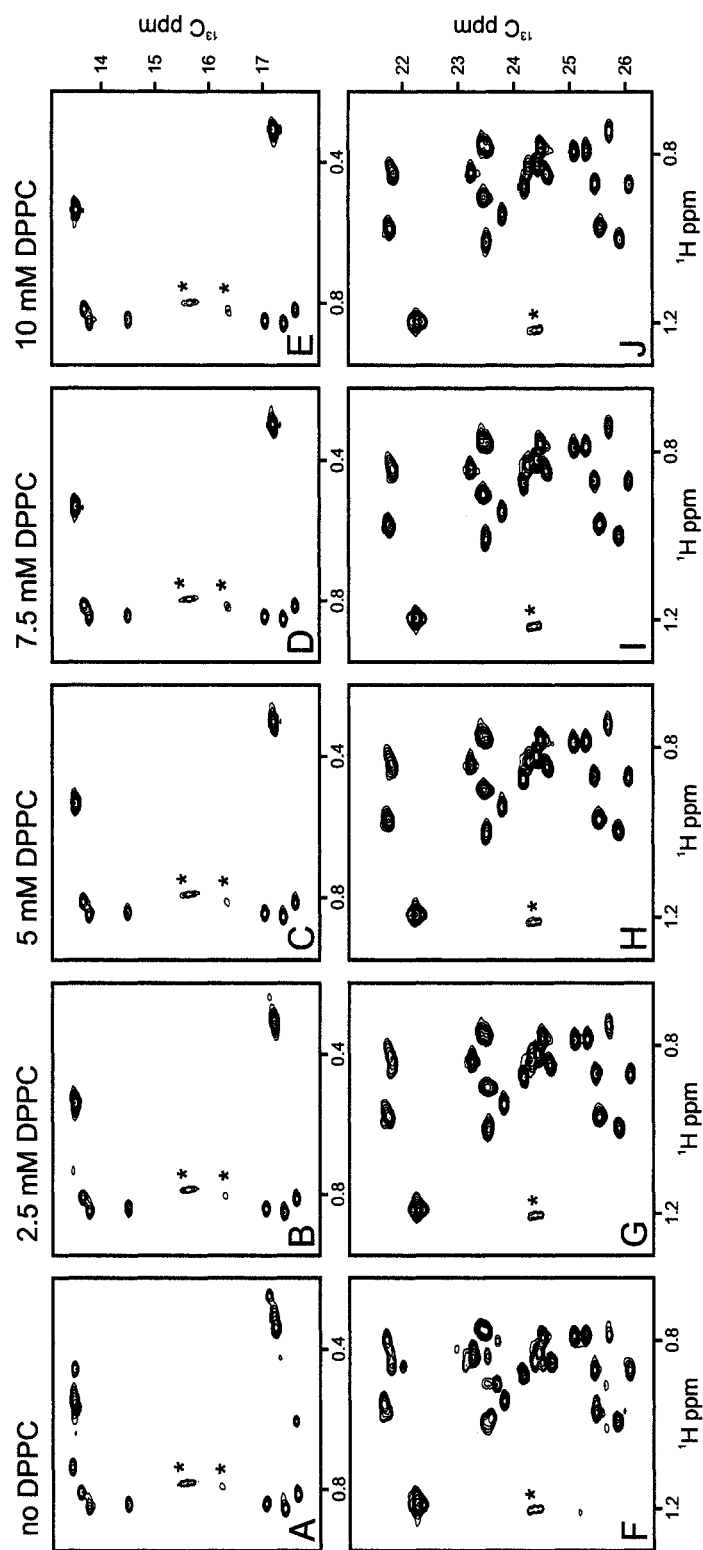
**Figure 5.1  $^1\text{H}/^{15}\text{N}$ -HSQC temperature series of the BNIP3 transmembrane domain in 5% dodecylphosphocholine**

The backbone amide region of a series of  $^1\text{H}/^{15}\text{N}$ -HSQC spectra acquired at temperatures 30-50 °C are presented. The spectra were acquired on a 500 MHz spectrometer with  $^1\text{H}$  carrier placed on the water resonance and  $^{15}\text{N}$  carrier at 116.8 ppm. The spectra were processed identically with NMRPipe: cosine-bell was applied in both direct and indirect dimension. The  $^1\text{H}/^{15}\text{N}$  cross-peaks for the side chains are indicated. Side chain  $^1\text{H}/^{15}\text{N}$  correlations can be observed on the spectra. Depending on the dwell time, these correlations are folded once (in red) or folded twice (in black).

Thus causing linewidth broadening, which is less desirable for structural study using solution NMR spectroscopy. Additionally, unlike SDS, DPC provides a head-group environment similar to the phosphatidylcholine-rich mitochondrial outer membrane, making it a better detergent of choice. Hence, the samples used for determination of the BNIP3 TMD dimer structure were prepared in DPC.

### **5.3 Added phospholipid affects BNIP3 transmembrane peptide chemical shifts in detergent micelles**

Though it requires acquisition with the more expensive  $^{13}\text{C}$ -labeled materials, the 2D  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum of the methyl region can also be used to examine the quality of the BNIP3 TMD sample and to confirm whether the conditions identified by preliminary analyses are indeed optimal for determining the structure of the TMD dimer. Figures 5.2 A and F show that a BNIP3 TMD sample prepared in the conditions chosen based on data from  $^1\text{H}/^{15}\text{N}$ -HSQC analyses may not be suitable for structural work because even though the peaks are well resolved, extra methyl peaks are observed in the spectra (see Figure 4.1 for the BNIP3 TMD peptide sequence). For example, seven peaks are seen in both the Ile  $\delta$  methyl region (13.5-14.5 ppm in  $^{13}\text{C}$  chemical shift) and the Ile  $\gamma$  methyl region (17-18 ppm) although the BNIP3 TMD peptide contains only four Ile residues (Figure 5.2 A). More than two Val  $\gamma$  methyl peaks are present at 21.7 ppm for the single Val164 (Figure 5.2 F), and more than eighteen Leu  $\delta$  methyl peaks are observed between 23-26 ppm for nine leucine residues (Figure 5.2 F).



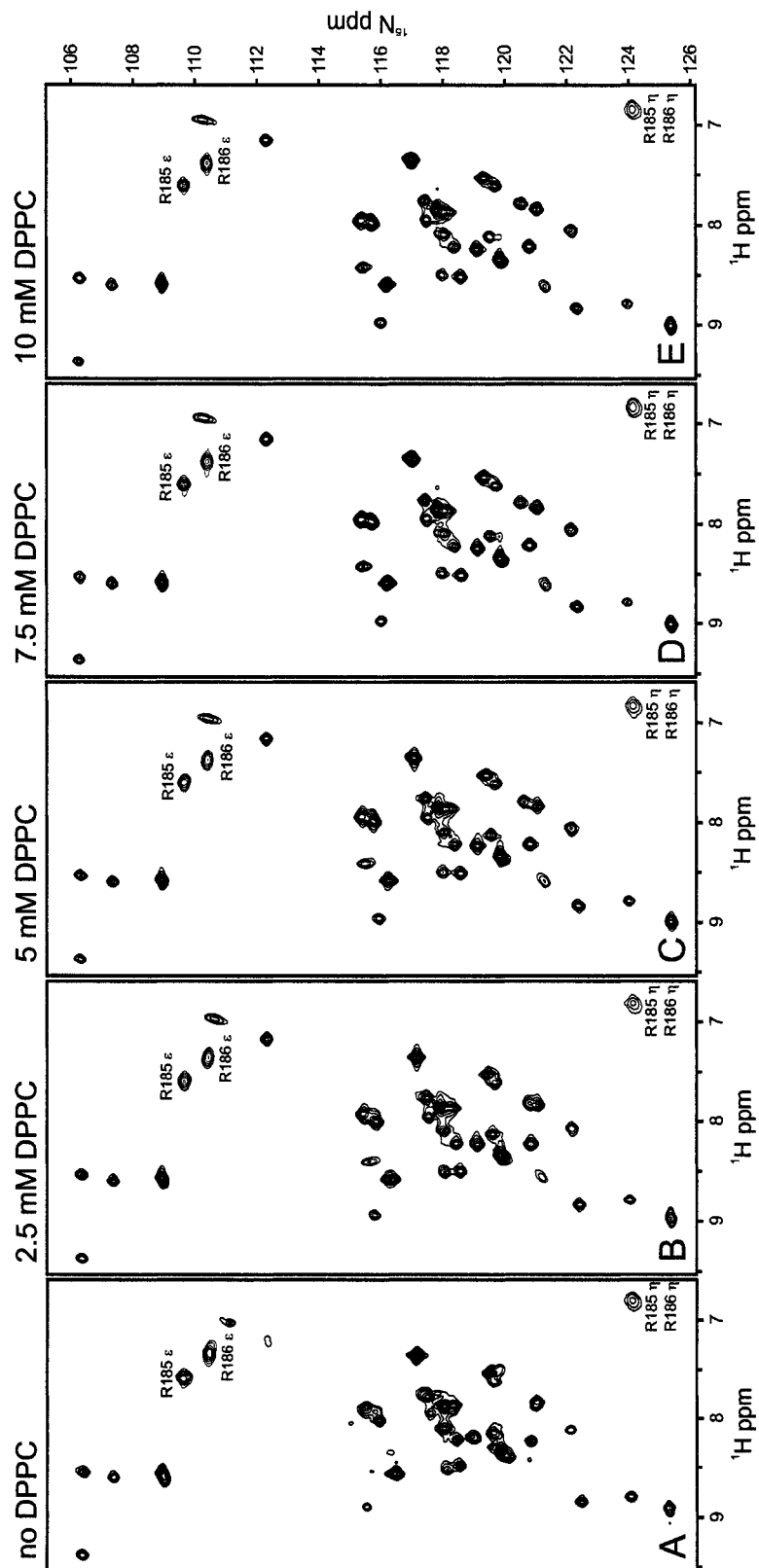
**Figure 5.2 Effects of added dipalmitoylphosphatidylcholine (DPPC) on peak dispersion and resolution on the  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum at the methyl region**

The  $^1\text{H}/^{13}\text{C}$ -HSQC spectra at the methyl region acquired with sample containing various concentrations of DPPC (0-10 mM) are presented. Panels A-E show the region for Ile  $\delta$  and  $\gamma$  methyls, and panels F-G show the region for Val  $\gamma$  methyl and Leu  $\delta$  methyl. The spectra were acquired on a 800 MHz spectrometer with  $^1\text{H}$  carrier placed on the water resonance and  $^{13}\text{C}$  carrier at 10 ppm. The spectra were processed identically with NMRPipe: cosine-bell squared on both the direct and indirect dimensions. The peaks marked with asterisks correspond to the  $^1\text{H}/^{13}\text{C}$  correlations of the DPC micelles.

The disproportionate number of methyl peaks suggests that the TMD peptide adopts multiple conformations under this sample condition, and that the rate for conversion between these conformations is slow enough that they can be observed on the NMR timescale.

I found that this multiple conformation issue can be alleviated by the addition of the lipid dipalmitoyl-phosphatidylcholine (DPPC) to the sample. Adding 2.5 mM DPPC, which is roughly 1:50 lipid-to-detergent molar ratio, is sufficient to eliminate the complexity seen in the  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum. Under these conditions, the number of  $^1\text{H}/^{13}\text{C}$  peaks for Ile, Leu, and Val methyls observed on the spectrum are as expected (four Ile  $\delta$  methyls and four  $\gamma$  methyls, a pair of Val  $\gamma$  methyls, and nine pairs of Leu  $\delta$  methyls) (Figure 5.2 B and G), indicating that addition of DPPC stabilizes a single conformation of the BNIP3 TMD peptide. One of the Val164  $\gamma$  methyl peaks still displays more than one conformation in this sample conditions though not to the extent observed in the absent of DPPC (Figure 5.2 G). At higher DPPC concentrations further improvements in the spectrum can be observed: linewidths become sharper and the peaks are better resolved (Figure 5.2 C, D, E, G, I, and J). However, no significant changes are observed at DPPC concentrations of >10 mM, suggesting that adding 10 mM DPPC to the BNIP3 TMD sample may be optimal.

Addition of DPPC to the NMR sample also increases the quality of the  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum. In the absence of DPPC, though the peaks are well resolved, peak intensities vary, probably due to differences in local peptide dynamics or exchange. Moreover, fewer than 33  $^1\text{H}/^{15}\text{N}$  cross-peaks can be observed in the spectrum (Figure 5.3 A). For a 35 amino acid peptide with an internal Pro, I expected to see 33  $^1\text{H}/^{15}\text{N}$  cross-peaks in the  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum. Adding 2.5 mM DPPC to the sample yielded a spectrum in which peak intensities are more uniform, linewidths are sharper, and peaks that were previously absent appear (Figure 5.3 B). Moreover, under these conditions, all 33 expected  $^1\text{H}/^{15}\text{N}$  cross-peaks are detected. At higher DPPC concentrations, the  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum quality improves similarly to what is observed in the  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum (Figures 5.3 C, D, E), but no significant improvements are seen with addition of >10 mM DPPC. Overall, these results indicate that addition of DPPC to the BNIP3 TMD sample can resolve the multiple conformation issue because added lipids appear to stabilize a single TMD peptide conformation. It is not clear how DPPC induces this stabilization, but the excellent peak dispersion and resolution of the  $^1\text{H}/^{13}\text{C}$ -HSQC and  $^1\text{H}/^{15}\text{N}$ -HSQC spectra suggests that structure determination of the BNIP3 TMD dimer can be performed under these sample conditions.



**Figure 5.3 Effects of dipalmitoylphosphatidylcholine (DPPC) addition on peak dispersion and resolution on the  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum**

Panels A-E show the backbone amide region of a series of  $^1\text{H}/^{15}\text{N}$ -HSQC spectra acquired with samples containing various concentrations of DPPC (0-10 mM). The spectra were acquired on a 800 MHz spectrometer with  $^1\text{H}$  carrier placed on the water resonance and  $^{15}\text{N}$  carrier at 116 ppm. The spectra were processed identically with NMRPipe: cosine-bell and line broadening of 5 Hz was applied in the direct dimension, and cosine-bell in the indirect dimension.

## 5.4 NMR data collection

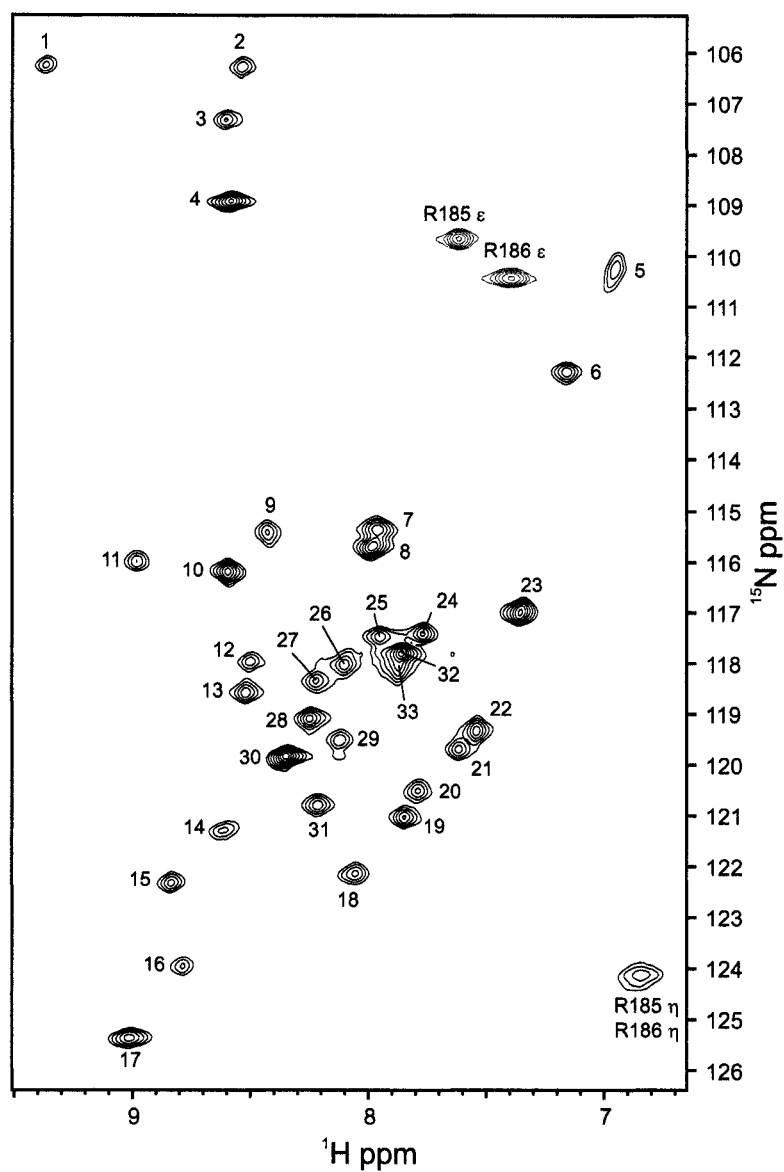
Identifying optimal sample and experimental conditions allowed me to start collecting data needed for determining the structure of the BNIP3 TMD dimer. Experiments needed for this data collection were performed with a single sample, which was prepared by mixing 0.4 mM [ $U$ - $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled BNIP3 TMD peptide with 0.6 mM unlabeled peptide in 10 mM phosphate buffer at pH 5.10, in the presence of 128 DPC and 10 mM DPPC. The use of mixed labeled and unlabeled peptide enabled identification of inter-monomer NOEs, which will be discussed in section 5.4.5. Most of the experiments were acquired on the 800 MHz spectrometer, with the exception of the  $^{13}\text{C}$  TOCSY and HCCH-COSY spectra, which are needed for assignment of the side chain chemical shifts, and the  $^3J$  coupling experiments, which are used to determine the side chain torsion angles. These experiments were acquired on the 600 MHz spectrometer. NMR data were processed with NMRPipe (Delaglio et al. 1995) and visualized and analyzed with either NMRDraw (Delaglio et al. 1995) or Sparky (Goddard and Kneller, UCSF).



### 5.4.1 Sequential assignment of the BNIP3 transmembrane domain peptide backbone atoms

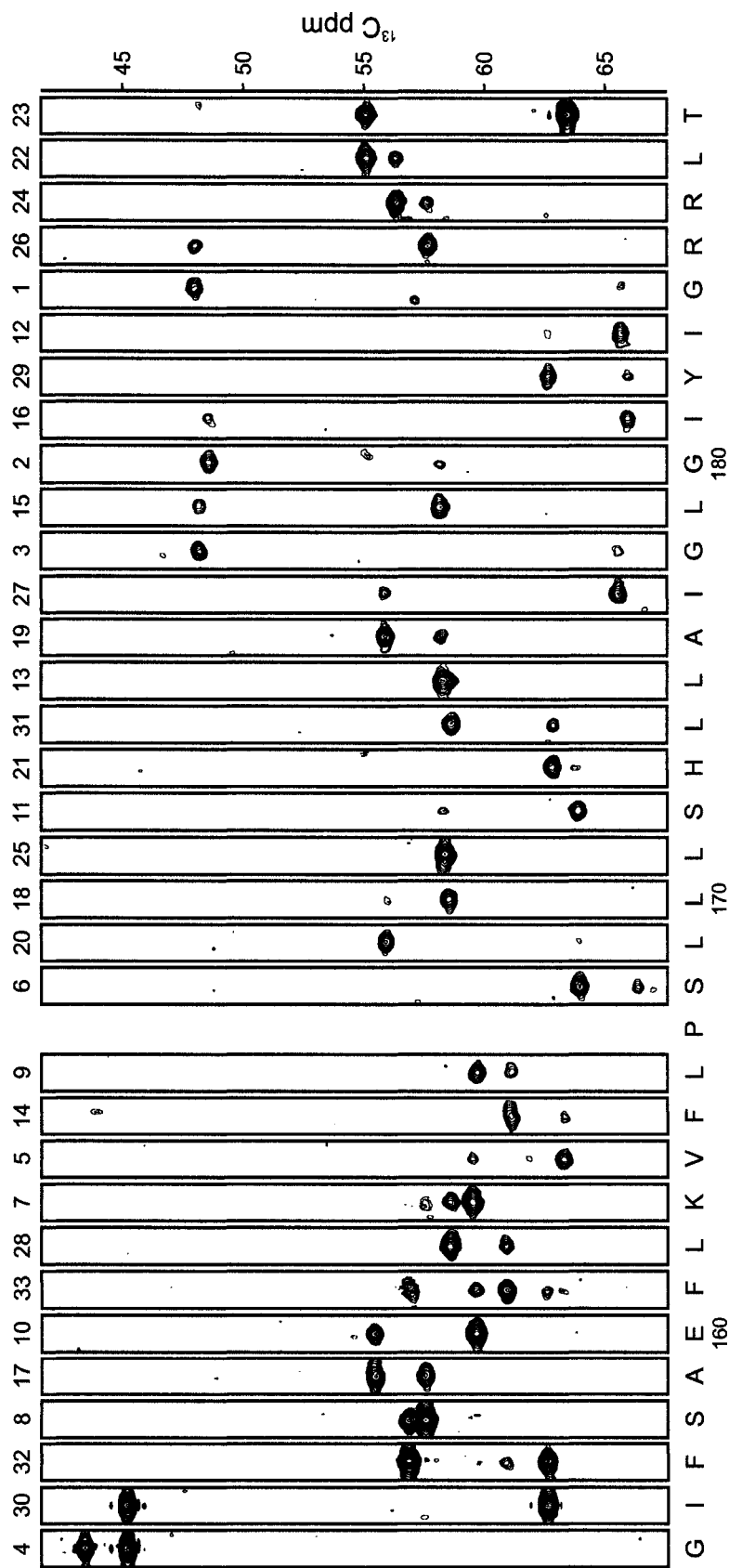
The first step in obtaining information for structure determination of the BNIP3 TMD dimer is to assign the chemical shifts of the backbone atoms. Assignments of the amide proton ( $H^N$ ), amide nitrogen (N), and alpha carbon ( $C^\alpha$ ) were performed with a standard triple resonance 3D HNCA experiment. In this experiment, each  $H^N/N$  pair is correlated to the alpha carbon of its own residue ( $C_i^\alpha$ ,  $^1J_{NC\alpha} \approx 12$  Hz) and of the preceding residue ( $C_{i-1}^\alpha$ ,  $^2J_{NC\alpha} \approx 7$  Hz). Because the HNCA experiment is based on the 2D  $^1H/^{15}N$ -HSQC experiment, every  $^1H^N/^{15}N$  cross-peak in the  $^1H/^{15}N$ -HSQC spectrum (Figure 5.4) can be correlated to two alpha carbon chemical shifts. Figure 5.5 shows 2D slices from the HNCA spectrum. Each slice represents a picked  $H^N/N$  pair from Figure 5.4, and their correlations with the  $C_i^\alpha$  (stronger intensity) and the  $C_{i-1}^\alpha$  (weaker intensity) atoms are indicated on the slices. Arranging the HNCA slices using the  $C_i^\alpha$  and  $C_{i-1}^\alpha$  chemical shift data produces a sequence that corresponds to the sequence of the BNIP3 TMD peptide. From this arrangement, I was able to assign the  $H^N$ , N, and  $C^\alpha$  chemical shifts for each residue in the TMD peptide.

Most of the HNCA slices provide unambiguous chemical shift assignments for the  $C_i^\alpha$  and  $C_{i-1}^\alpha$ . However, some ambiguity was encountered when the  $C_i^\alpha$  and  $C_{i-1}^\alpha$  correlations are overlapping as shown in strip 13 in Figure 5.5, or when there is a degeneracy in the  $C^\alpha$  chemical shifts, such as strips 13 and 25 in Figure 5.5.



**Figure 5.4** The  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum of BNIP3 transmembrane domain peptide prepared in 128 mM DPC, 10 mM DPPC, 10 mM phosphate, pH 5.10

Shown is the  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum of the sample that was used for structure determination of the BNIP3 transmembrane domain dimer. The spectrum was acquired and process as described in Figure 5.3. The labels on the  $^1\text{H}/^{15}\text{N}$  cross-peaks indicate the numbering of the picked peaks in the HNCA data set.

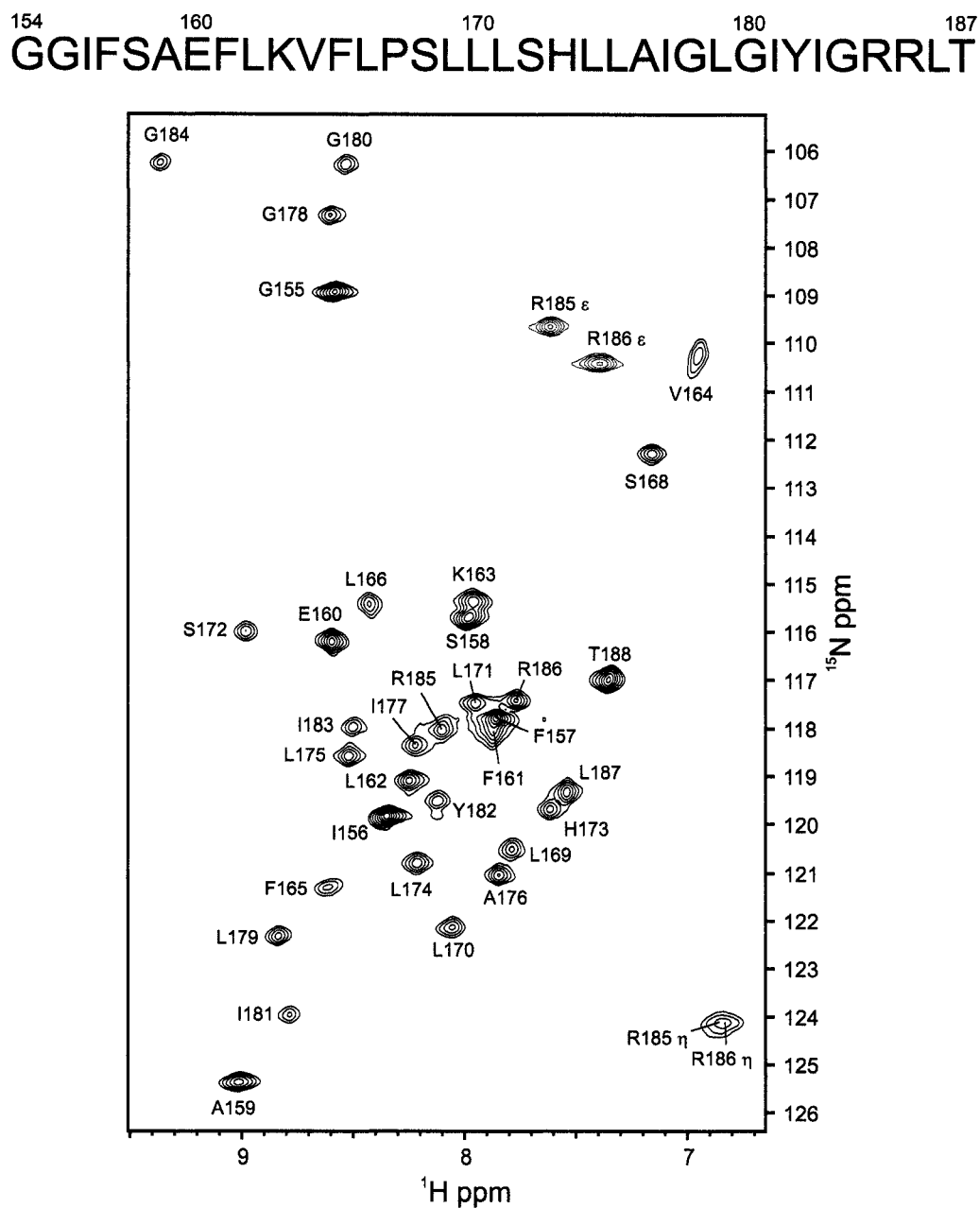


**Figure 5.5 Slices from 3D HNCA spectrum of the BNIP3 transmembrane domain**

Each HNCA slice represents a  $^1\text{H}/^{15}\text{N}$  pair that was picked from the 2D  $^1\text{H}/^{15}\text{N}$ -HSQC spectrum (see Figure 5.4), with the peak numbers shown on top. Shown on the HNCA slices are the  $\text{C}^\alpha$  and  $\text{C}-1^\alpha$  chemical shifts correlated to the  $^1\text{H}/^{15}\text{N}$  pair. These correlations can be used to arrange the slices in a sequence that correspond to the sequence of the BNIP3 transmembrane domain peptide (shown at the bottom). The HNCA spectrum was acquired on the 800 MHz spectrometer using the BioPack version of the HNCA pulse sequence.

I resolved such ambiguities by comparing the HNCA strips against strips from another triple resonance experiment, 3D CBCA(CO)NNH, which correlates each  $H^N/N$  pair with the alpha and beta carbons of the preceding residue ( $C_{i-1}^\alpha$  and  $C_{i-1}^\beta$ ). Other triple resonance 3D experiments, such as HN(CO)CA and HNCACB, can also be employed for the backbone chemical shift assignments. I used information from the HN(CO)CA spectrum at the initial stage of the structural study, but not in the stage that led to solving the structure of the BNIP3 TMD dimer. I did not acquire the HNCACB spectrum because this experiment is not sensitive enough to detect the weaker  $C_{i-1}^\alpha$  and  $C_{i-1}^\beta$  peaks for my system. Nevertheless, sequential assignment of the backbone atoms using HNCA and CBCA(CO)NNH allowed me to assign the  $H^N/N$  peaks in the  $^1H/^{15}N$ -HSQC spectrum to their respective residues (Figure 5.6).

Chemical shifts for carbonyl carbon (C) of each residue were assigned from a 3D HNCO spectrum. This experiment correlates the  $H^N/N$  pair with the chemical shift of carbonyl carbon of the preceding residue ( $C_{i-1}$ ). With the knowledge of  $H^N$  and N chemical shifts of each residue from the HNCA data, I was able to assign the carbonyl carbon resonances to their corresponding residues. The HNCO spectrum also enabled assignments of the Arg  $H_\epsilon$  and  $N_\epsilon$  resonances which will be discussed further in section 5.4.2.



**Figure 5.6 Assigned resonances of the BNIP3 transmembrane domain <sup>1</sup>H/<sup>15</sup>N-HSQC**

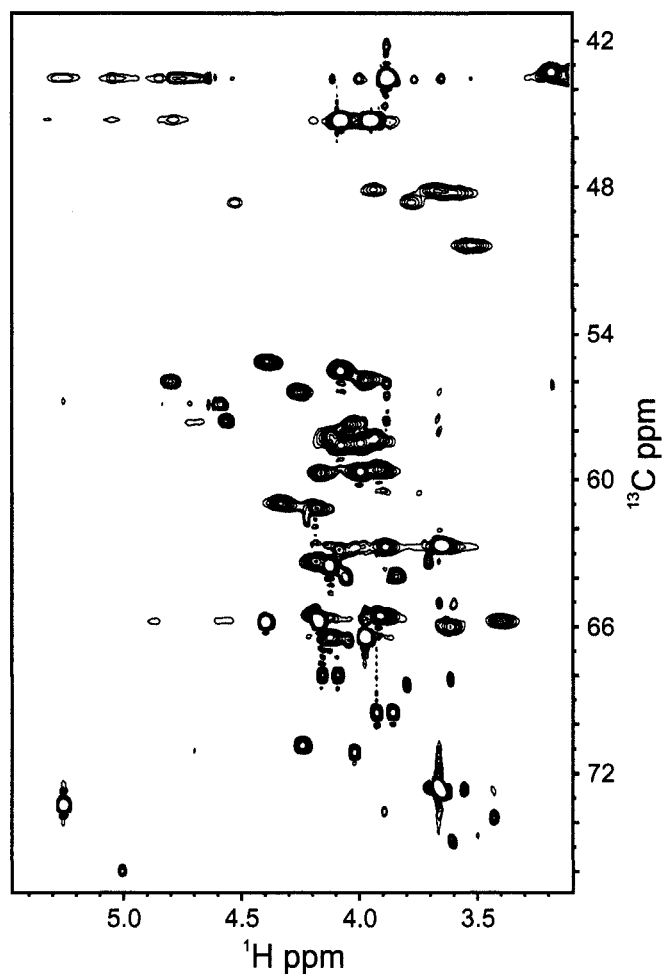
The <sup>1</sup>H/<sup>15</sup>N-HSQC spectrum is identical to that shown in Figure 5.4. The BNIP3 transmembrane domain peptide sequence is shown on top. All assigned backbone correlation are labeled on the spectrum. The weak intensity observed downfield of Y182 in the <sup>15</sup>N chemical shift dimension is an artifact of the I156 peak.

Though knowledge of the  $C^\alpha$  chemical shifts for each residue in the BNIP3 TMD peptide should enable assignments of the  $H^\alpha$  resonances from the  $^1H/^{13}C$ -HSQC spectrum in the  $H^\alpha/C^\alpha$  region, this approach is not straight forward because many residues in the BNIP3 TMD peptide have overlapping  $C^\alpha$  chemical shifts, thus preventing unambiguous assignment of the  $H^\alpha$  resonances. Moreover, cross peaks that correspond to the choline head group of the detergent DPC or the phospholipid DPPC are also detected in this region and overlap with some of the peptide  $H^\alpha/C^\alpha$  cross-peaks, hence obscuring the assignments (Figure 5.7).

Nevertheless, data from 2D  $^1H/^{13}C$  CT-HSQC and 3D triple-resonance experiments HNCA, CBCA(CO)NNH, and HNCO enabled assignment of  $H^N$ , N,  $C^\alpha$ ,  $C^\beta$ , and C chemical shifts, and these resonances are summarized in Table 5.1.

#### **5.4.2 Assignment of the side chain chemical shifts**

The second step in data collection is to assign the side chain chemical shifts which was performed by analyzing 3D H(C)CH-COSY and (H)CCH-TOCSY spectra. These experiments utilize through bond magnetization transfer to establish connectivities between the sequentially assigned  $C^\alpha$  resonances from the HNCA and the side chain methyl and methylene carbon and proton resonances. The (H)CCH-TOCSY spectrum was acquired on the 600 MHz spectrometer using BioPack pulse sequence, and the H(C)CH-COSY was also acquired on the 600 MHz spectrometer using the pulse sequence written by my advisor Kevin MacKenzie.



**Figure 5.7**  $\text{H}^\alpha/\text{C}^\alpha$  region of the  $^1\text{H}/^{13}\text{C}$ -HSQC of the BNIP3 transmembrane domain

The spectrum indicates that in addition to the peptide  $\text{H}^\alpha/\text{C}^\alpha$  pairs, correlations on the detergent or lipid head groups, marked with sharper line widths, can be detected in this regions. The spectrum was acquired on the 800 MHz spectrometer using a gradient selected constant-time  $^1\text{H}/^{13}\text{C}$ -HSQC pulse sequence (constant time = 10 msec), written by my advisor, Kevin MacKenzie.

**Table 5.1** List of backbone atom chemical shifts for the BNIP3 transmembrane domain peptide obtained from HNCA, CBCA(CO)NNH, and HNCB spectra

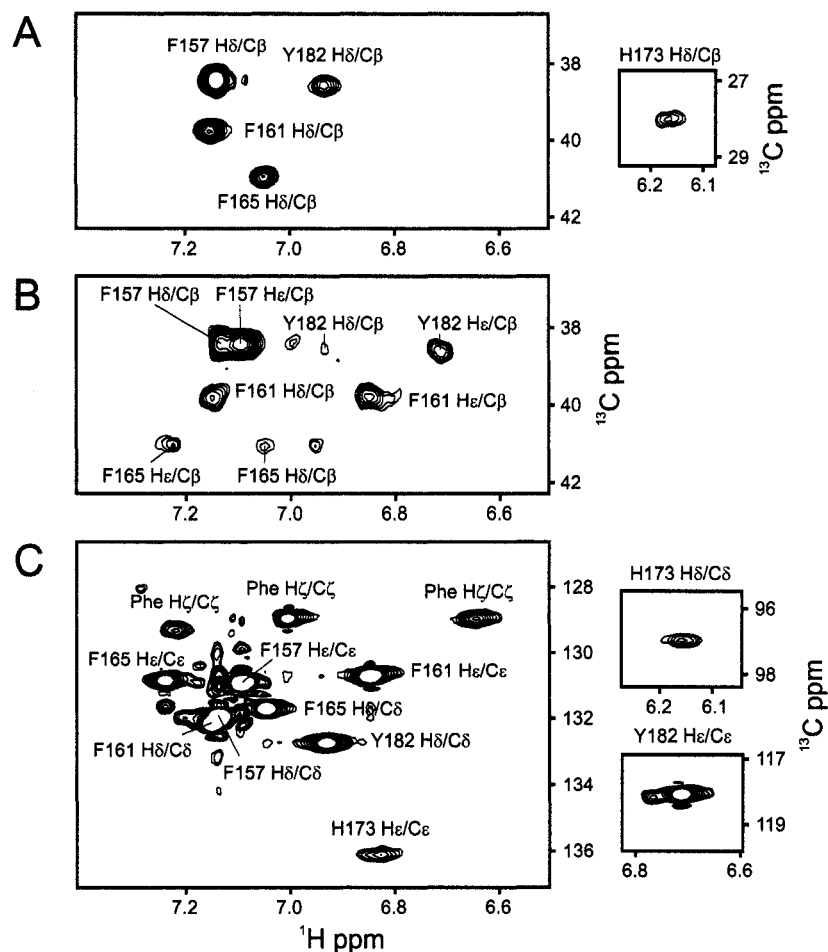
Residue	Peak #	H <sup>N</sup>	N	C <sup>α</sup>	C <sup>β</sup>	C
Gly154		-	-	42.77	-	169.94
Gly155	4	8.57	108.94	44.56	-	173.69
Ile156	30	8.34	119.87	62.06	37.68	174.47
Phe157	32	7.85	117.79	56.31	37.84	174.32
Ser158	8	7.97	115.67	57.01	64.87	174.16
Ala159	17	9.00	125.39	54.85	18.07	178.29
Glu160	10	8.59	116.18	59.01	28.22	177.44
Phe161	38	7.86	118.05	60.33	39.16	177.13
Leu162	28	8.24	119.11	57.99	41.51	177.13
Lys163	7	7.95	115.32	58.85	32.29	174.01
Val164	5	6.94	110.25	62.68	32.68	177.29
Phe165	14	8.61	121.29	60.41	40.33	175.72
Leu166	9	8.42	115.42	59.09	38.38	-
Pro167		-	-	65.80	30.34	177.15
Ser168	6	7.15	112.33	63.33	62.06	174.16
Leu169	20	7.77	120.48	55.34	41.98	180.10
Leu170	18	8.05	122.15	57.91	40.80	176.97
Leu171	25	7.94	117.49	57.76	41.12	177.91
Ser172	11	8.97	115.97	63.19	59.73	176.82
His173	21	7.60	119.72	62.22	27.44	176.97
Leu174	31	8.20	120.78	57.91	40.96	178.22
Leu175	13	8.51	118.60	57.76	40.80	177.44
Ala176	19	7.83	121.03	55.26	18.07	178.22
Ile177	27	8.21	118.33	64.87	37.37	178.38
Gly178	3	8.59	107.32	47.53	-	173.85
Leu179	15	8.83	122.35	57.60	41.04	178.22
Gly180	2	8.52	106.26	47.92	-	174.16
Ile181	16	8.78	123.97	65.34	36.90	177.13
Tyr182	29	8.11	119.57	61.90	37.91	176.97
Ile183	12	8.49	118.00	65.02	37.60	177.76
Gly184	1	9.36	106.26	47.38	-	174.32
Arg185	26	8.09	118.05	57.05	29.63	177.29
Arg186	24	7.76	117.39	55.67	29.71	176.19
Leu187	22	7.53	119.36	54.44	41.59	175.41
Thr188	23	7.34	116.99	62.78	70.18	-



Both experiments were setup such that the  $^{13}\text{C}$  chemical shift dimension was folded to permit acquisition in a reasonable amount of time.

Side chain assignment of the aliphatic residues can be performed by first classifying the methyl  $^1\text{H}/^{13}\text{C}$  correlations in the  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum into residue types based on the  $^{13}\text{C}$  chemical shifts as was discussed in section 5.2 for Ile, Leu, and Val methyl groups. Then, the  $\text{C}^\alpha$  shifts on the H(C)CH-COSY and (H)CCH-TOCSY columns of these methyls were compared to the  $\text{C}^\alpha$  shifts obtained from the HNCA experiment. For long chain amino acids Arg and Lys, this assignment was done through connections to  $\text{C}^\alpha$  in columns that arose from the  $\gamma$  or  $\delta$  methylenes, and for Phe, Tyr, and His, the assignment used the connections in the columns that arose from the  $\beta$  methylenes.

The same connectivities from the aromatic carbons and protons to the  $\text{C}^\alpha$  could not be detected with the H(C)CH-COSY or (H)CCH-TOCSY experiments, but I could identify the aromatic carbon resonances of Phe, Tyr, and His by first assigning the  $\text{H}^\delta$  and  $\text{H}^\epsilon$  resonances through correlations to the  $\text{C}^\beta$  chemical shifts from the HBCBCGCDHD and HBCBCGCDCEHE spectra (Figure 5.8) (Yamazaki et al. 1993). Then, I assigned the  $\text{C}^\delta$  and  $\text{C}^\epsilon$  chemical shifts from the  $\text{H}^\delta/\text{C}^\delta$  and  $\text{H}^\epsilon/\text{C}^\epsilon$  cross-peaks in the  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum for the aromatic region.



**Figure 5.8 2D spectra for the aromatic residues of the BNIP3 transmembrane domain**

Presented are spectra acquired with the experiments HBCBCGCDHD (A), HBCBCGCDCEHE (B) (Yamazaki et al., 1993), and constant time  $^1\text{H}/^{13}\text{C}$ -HSQC of the aromatic region (C). The HBCBCGCDHD and HBCBCGCDCEHE spectra enable assignment of the H $\delta$  and H $\epsilon$  chemical shifts through correlations with the C $\beta$  resonances. This information then allows assignment of the C $\delta$  and C $\epsilon$  chemical shifts of the  $^1\text{H}/^{13}\text{C}$ -HSQC spectrum.

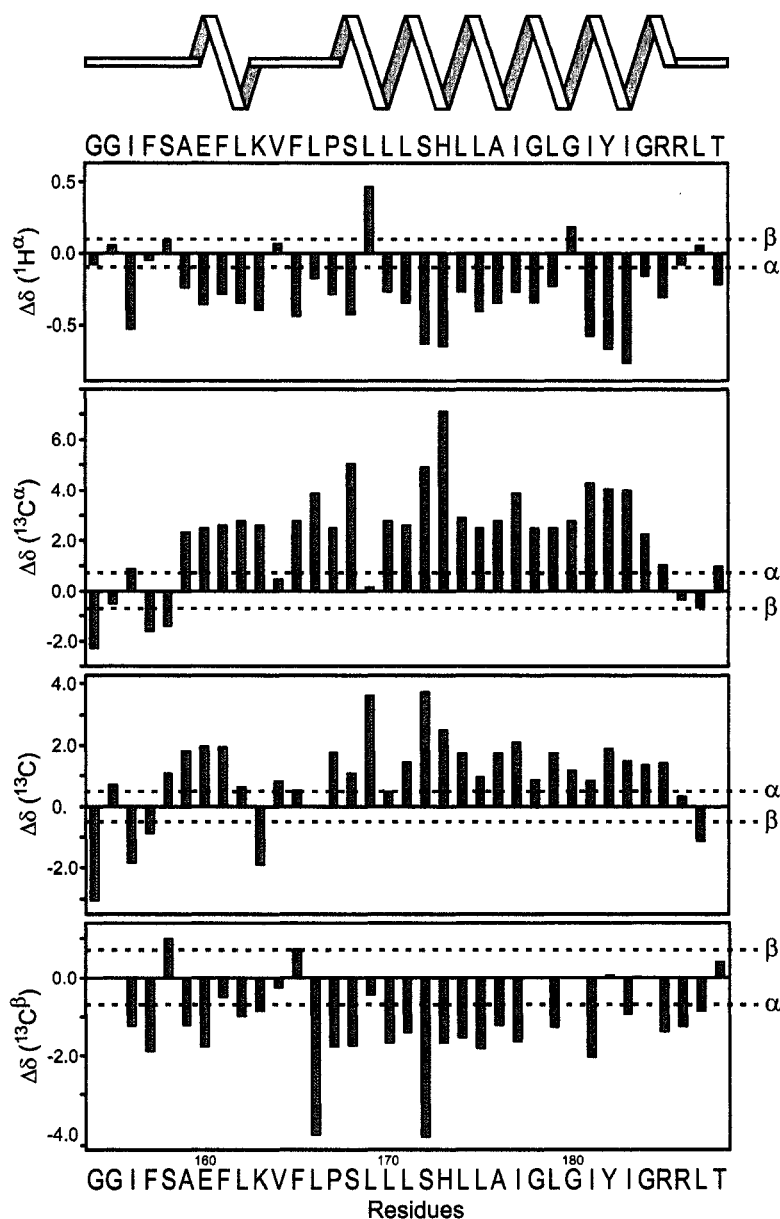
The folded side chain peaks observed in the  $^1\text{H}/^{15}\text{N}$ -HSQC spectra correspond to the Arg  $\text{H}^\epsilon/\text{N}^\epsilon$  and  $\text{H}^\eta/\text{N}^\eta$  resonances. The  $\text{H}^\epsilon/\text{N}^\epsilon$  peaks can be assigned to specific residues by matching the  $\text{C}^\delta$  chemical shifts, obtained from the (H)CCH-TOCSY, to the  $\text{C}^\delta$  chemical shift from the HNCA experiment because the HNCA spectrum shows the correlations between Arg  $\text{H}^\epsilon/\text{N}^\epsilon$  chemical shifts and the  $\text{C}^\delta$ . A similar approach using data from the HNCO spectrum can be used to assign the resonances for Arg  $\text{H}^\eta/\text{N}^\eta$ . However, because the two  $\text{H}^\eta/\text{N}^\eta$  pairs in each Arg residue exchange rapidly, I could not detect these correlations in the HNCO spectrum. Hence, I could not unambiguously assign the chemical shifts of these atoms.

Even though I could not determine the stereospecific assignment of the methylene protons or the methyl carbons of Val  $\gamma$  and Leu  $\delta$  using the backbone and side chain assignment protocols, this approach allowed me to assign roughly 93% of the BNIP3 TMD chemical shifts (Appendix C). Each residue has only a single set of chemical shifts, suggesting that the monomers in the BNIP3 TMD dimer are symmetrical. Therefore, all restraints used for structure calculation can be doubled for each monomer.

### 5.4.3 Determination of backbone dihedral angles

Secondary structures of polypeptides have characteristic backbone torsion angles  $\phi$  and  $\psi$ . An  $\alpha$ -helix has  $\phi$  and  $\psi$  angles of  $-57^\circ$  and  $-47^\circ$ , respectively, and a  $\beta$ -sheet has  $\phi$  and  $\psi$  angles of  $-120^\circ$  and  $140^\circ$ . Where supported by experimental data, restraining these torsion angles can complement the distance restraints provided by NOE correlations for structure calculation. The chemical shifts of  $H^\alpha$ ,  $C^\alpha$ ,  $C^\beta$ , and carbonyl C, obtained from sequential assignment of the backbone atoms, can be used to evaluate secondary structure components of a protein or a polypeptide because these shifts are strongly dependent on the protein secondary structure (Wishart et al. 1991; Wishart et al. 1992; Wishart and Sykes 1994). For example, the  $H^\alpha$  chemical shifts of all 20 amino acids experience an upfield shift relative to the random coil value when in  $\alpha$ -helical configuration, and a downfield shift when in  $\beta$ -strand configuration.

I evaluated the secondary structure of the residues in the BNIP3 TMD by comparing the chemical shifts listed in Table 5.1 against the chemical shift indices (CSI) of Wishart and Sykes (Wishart and Sykes 1994). Differences in the  $H^\alpha$ ,  $C^\alpha$ ,  $C^\beta$ , and C resonances of each residue from the CSI of Wishart and co-workers are presented in Figure 5.9, and the data suggest that most of the BNIP3 TMD peptide is  $\alpha$ -helical under conditions chosen for structural study. A consensus of the four chemical shift data suggests that the  $\alpha$ -helical regions cover residues 159-163 and 167-185. Therefore,  $\phi$  and  $\psi$  torsion angle restraints for  $\alpha$ -helix, as well as intra-molecular hydrogen bond restraints, were applied for residues in these regions for structure calculation.



**Figure 5.9 CSI analysis of the chemical shifts of backbone resonances identifies  $\alpha$ -helical region of the BNIP3 transmembrane domain**

The  $\text{H}^\alpha$ ,  $\text{C}^\alpha$ ,  $\text{C}$ , and  $\text{C}^\beta$  chemical shift data were analyzed using the methods of Wishart and Sykes (1994), Wishart *et al.* (1992), and Wishart *et al.* (1991). The chemical shift difference ( $\delta\Delta$ ) range for random coil is indicated between the dashed lines. The secondary structure of the BNIP3 transmembrane domain peptide was determined using the consensus of the four shift indices. Region A159-K163 and P167-R185 are determined to be  $\alpha$ -helix (shown on top).

Three-bond coupling between  $H^N$  and  $H^\alpha$  ( $^3J_{HNH\alpha}$ ) provides an alternative method to define the backbone  $\phi$  angle (Vuister and Bax 1993). The magnitude of the  $^3J_{HNH\alpha}$  couplings can be related to the  $\phi$  angle using Karplus equation:

$$J(\phi) = A \cos^2(\phi - 60) + B \cos(\phi - 60) + C$$

where  $A = 6.98$ ,  $B = -1.38$ , and  $C = 1.72$ . For BNIP3 TMD, the  $^3J_{HNH\alpha}$  measurements may be not reliable for residues in the middle region because of the broad linewidths (~40 Hz), but the sharper linewidths of amino terminal residues imply that the measurements for eight residues in this region can be used to define the  $\phi$  angles (Table 5.2). These results indicate that residues 159-163 are in  $\alpha$ -helical configuration ( $^3J_{HNH\alpha} < 6$  Hz), supporting the conclusions obtained from CSI analyses.

#### 5.4.4 Determination of side chain torsion angles

Side chain torsion angles  $\chi_1$  of Ile and Val residues were identified by measuring  $^3J$  couplings between the  $\gamma$  methyl carbons and the backbone C and N nuclei ( $^3J_{CC}$  and  $^3J_{CN}$ , respectively). These experiments were acquired using gradient-selected versions of Bax and co-workers methods (Grzesiek et al. 1993; Vuister et al. 1993) that have been written by my advisor, Kevin MacKenzie. The reference and difference spectra obtained from the  $^3J_{CC}$  and  $^3J_{CN}$  experiments are shown in Figure 5.10, and the  $^3J$ -coupling values were calculated from the cross peak intensities observed on these spectra (Table 5.3). For a gauche configuration, the coupling values are expected to be 0.9 and 0.4 Hz for  $^3J_{CC}$  and  $^3J_{CN}$ , respectively, and the coupling values for the trans configuration are expected to be 4.0 and 2.1 Hz.

**Table 5.2**  $^3J_{\text{HNNH}\alpha}$  coupling for residues at the amino terminus of the BNIP3 transmembrane domain peptide

Residue	$S_{\text{cross}}$	$S_{\text{diagonal}}$	$^3J_{\text{HNNH}\alpha}$ (Hz)
G155	$-2200900 \pm 86150$	$6673700 \pm 86150$	$6.6 \pm 0.1$
I156	$-5907200 \pm 89332$	$26384000 \pm 89332$	$5.6 \pm 0.1$
F157	$-7376200 \pm 88395$	$24386000 \pm 88395$	$6.4 \pm 0.1$
S158	$-1909000 \pm 86445$	$7266300 \pm 86445$	$6.0 \pm 0.1$
A159	$-216450 \pm 93712$	$2282700 \pm 93712$	$3.8 \pm 0.6$
E160	$-728010 \pm 91637$	$7759200 \pm 91637$	$3.8 \pm 0.2$
L162	$-319600 \pm 92418$	$2360300 \pm 92418$	$4.5 \pm 0.5$
K163	$-728010 \pm 91637$	$7759200 \pm 91637$	$3.8 \pm 0.2$

$^3J_{\text{HNNH}\alpha}$  is calculated using the formula described by Vuister and Bax (Vuister and Bax 1993):

$$^3J_{\text{HNNH}\alpha} = \frac{1}{2t\pi} \tan^{-1} \left( \sqrt{-\frac{S_{\text{cross}}}{S_{\text{diagonal}}}} \right)$$

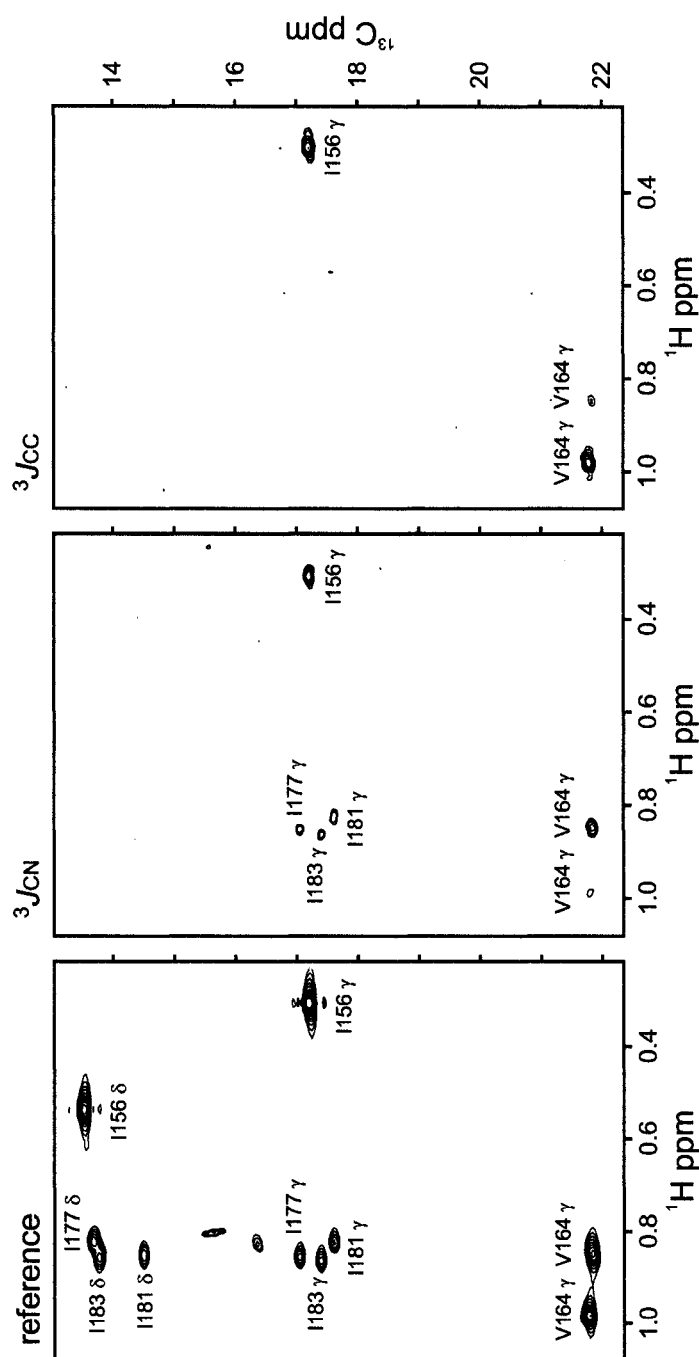
where  $t = \frac{1}{2} (del1 + del2)$

$S_{\text{cross}}$  = intensity of  $\text{H}^\alpha$  peak

$S_{\text{diagonal}}$  = intensity of  $\text{H}^N$  peak

$(del1 + del2) = 0.025057$  (indicated by BioPack)

A  $^3J_{\text{HNNH}\alpha}$  of  $< 6.0$  Hz is an indicative of local  $\alpha$ -helix configuration.



**Figure 5.10 Reference and spin-echo difference spectra used to determine coupling between  $\gamma$  methyl C and backbone N or C**

The methyl region of the CT-HSQC reference spectrum (left) and the spin-echo difference spectra acquired with  $^3J_{\text{CN}}$  (middle) and  $^3J_{\text{CC}}$  (right) experiments. The  $^3J_{\text{CN}}$  difference spectrum is plotted at a contour level 7 times lower than the reference spectrum and the  $^3J_{\text{CC}}$  is plotted 4 times lower than the reference spectrum. The data were used to identify the  $\chi^1$  torsion angles of Ile and Val  $\gamma$ -methyls. These spectra were acquired with the 600 MHz spectrometer using pulse sequences written by my advisor and based on the methods of Bax and co-worker (Grzesiek et al. 1993, Vuister et al. 1993).



**Table 5.3** Three bond heteronuclear  $J$  couplings between  $\gamma$  carbons and backbone C and N nuclei

Residue	Resonance ( $^1\text{H}$ , ppm)	$^3J_{\text{CC}}$ (Hz)	$\phi_{\text{trans}}(\text{C})$	$^3J_{\text{CN}}$ (Hz)	$\phi_{\text{trans}}(\text{N})$	Torsion angle for structure calculation
Val164	(0.83)	$0.62 \pm 0.03$	$\sim 0.0$	$0.82 \pm 0.02$	$0.25 \pm 0.01$	$\chi 1$ is not determined
	(0.97)	$2.25 \pm 0.01$	$0.44 \pm 0.01$	$0.34 \pm 0.03$	$\sim 0.0$	
Thr188		$1.05 \pm 0.01$	$0.05 \pm 0.01$	$< 0.4$	$\sim 0.0$	$\chi 1$ is not determined
Ile156		$1.35 \pm 0.01$	$0.14 \pm 0.01$	$0.62 \pm 0.01$	$0.13 \pm 0.01$	$\chi 1$ is not determined
Ile177		$0.75 \pm 0.08$	$\sim 0.0$	$1.34 \pm 0.03$	$0.55 \pm 0.02$	$\chi 1 = 180^\circ$
Ile181		$< 0.60$	$\sim 0.0$	$1.15 \pm 0.05$	$0.44 \pm 0.03$	$\chi 1 = 180^\circ$
Ile183		$< 0.60$	$\sim 0.0$	$0.43 \pm 0.10$	$0.02 \pm 0.05$	$\chi 1$ is not determined

$^3J_{\text{CX}}$  were calculated using the formula (Grzesiek et al. 1993; Vuister et al. 1993):

$$^3J_{\text{CX}} = \frac{1}{\pi T} \sin^{-1} \left( \sqrt{\frac{I_{\text{diff}}}{2I_{\text{ref}}}} \right)$$

Where  $^3J_{\text{CX}}$  = three bond coupling between  $\gamma$  carbon and the backbone C or N

$T$  = constant time period

$I_{\text{diff}}$  = intensity of a correlation in the difference spectrum

$I_{\text{ref}}$  = intensity of a correlation in the reference spectrum

The fraction of the trans rotamer ( $\phi_{\text{trans}}$ ) were calculated with the formula:

$$^3J_{\text{CX}} = ^3J_{\text{trans}} \phi_{\text{trans}} + ^3J_{\text{gauche}} (1 - \phi_{\text{trans}})$$

Where  $^3J_{\text{trans}} = 4$  or  $2.1$  for  $^3J_{\text{CC}}$  or  $^3J_{\text{CN}}$ , respectively

$^3J_{\text{gauche}} = 0.9$  or  $0.4$  for  $^3J_{\text{CC}}$  or  $^3J_{\text{CN}}$ , respectively

Many of the values shown in Table 5.3 do not strongly indicate either trans or gauche configuration, suggesting that the side chains are experiencing rotamer averaging, and the fraction of time spent in the trans configuration ( $\phi_{\text{trans}}$ ) can be calculated from the  $^3J$ -coupling data. However, the  $^3J$  coupling values may be affected by systematic errors: incomplete  $^{13}\text{C}$  labeling, fast relaxation, and incomplete inversion of the spins (Grzesiek et al. 1993). The sum of these errors can result in a 1-5% reduction of the  $^3J_{\text{CC}}$  and  $^3J_{\text{CN}}$  couplings, as well as the  $^3J_{\text{CC}\alpha}$  coupling which is presented in Table 5.4.

The  $^3J$ -couplings for Ile177 and Ile181 indicate that their  $\gamma$  methyls are trans from the N nuclei 55% and 44% of the time, respectively. Though these values are low, a  $\chi_1$  of  $180^\circ$  can be assigned for these residues. The  $\chi_1$  of Ile183 may average between  $60^\circ$  and  $180^\circ$  because its  $^3J_{\text{CC}}$  of  $\sim 0$  suggests that the  $\gamma$  methyl cannot be trans from the carbonyl C ( $\chi_1$  of  $-60^\circ$ ). A specific  $\chi_1$  torsion angle cannot be assigned to Ile156, because the  $^3J_{\text{CC}}$  and  $^3J_{\text{CN}}$  values imply that the  $\gamma$  methyl spends almost equal time among the  $-60^\circ$ ,  $60^\circ$ , and  $180^\circ$  rotamers. In some instances, the  $^3J_{\text{CC}}$  and  $^3J_{\text{CN}}$  couplings may be used to identify stereospecific assignment of Val164  $\gamma$  methyls. Yet, because these Val 164  $\gamma$ -methyls are averaging (the  $\gamma$  methyl on 0.83 ppm spends 25% of the time trans from the amide N, and the  $\gamma$  methyl on 0.98 ppm spends 44% of the time trans from the carboxyl C, I could not assign the stereochemistry of these  $\gamma$  methyls or their  $\chi_1$  torsion angles with certainty. However, the NOE correlations data imply that the  $\gamma$  methyl at 0.83 ppm is likely the  $\gamma_1$  methyl, and the one at 0.98 ppm is likely the  $\gamma_2$  methyl. Based on these NOE data, I was able to assign specific chirality for Val164  $\gamma$  methyls, but I did not include the  $\chi_1$  torsion angle restraint in structure calculation.

My advisor developed a method similar to the  $^3J_{CC}$  and  $^3J_{CN}$  experiments to identify  $^3J$ -coupling between  $\delta$  methyls of Ile and Leu and the  $C^\alpha$  (MacKenzie et al. 1996). Data collection and analysis of the  $^3J_{CC\alpha}$  coupling are the same as those described for  $^3J_{CC}$  and  $^3J_{CN}$  (Figure 5.11 and Table 5.4), and I could use these data to determine the  $\chi_2$  torsion angles and stereospecific assignments for Leu  $\delta$  methyls. However, because the  $\delta$ -methyls of some Leu residues are averaging, similar what observed with the Val164  $\gamma$ -methyls, I was able to determine with certainty only the  $\chi_2$  torsion angles and stereospecific assignments for Leu171, Leu174, Leu175, and Leu179 (Table 5.4). NOE correlations data were employed to assign the chirality of the  $\delta$ -methyls for Leu166 and Leu169, but these data were not sufficient to resolve the stereospecific assignments of the  $\delta$ -methyls for Leu162, Leu170, and Leu187. Therefore, for structure calculation I included only the  $\chi_2$  torsion angles for Leu171, Leu174, Leu175, and Leu179. In addition, I specified the chirality, but not the  $\chi_2$  torsion angles, of Leu166 and Leu169  $\delta$ -methyls.

The reference spectrum (left) is plotted at a contour level 6 times higher than the difference spectrum (right). The data were used to determine the  $\chi^2$  torsion angles of Ile and Leu  $\delta$ -methyls. In combination with NOE correlations data, the  $^3J_{CC\alpha}$  couplings can be used to identify the stereospecific assignment of Leu  $\delta$ -methyls. The experiment was performed on the 600 MHz spectrometer using pulse sequences written by my advisor (MacKenzie et al 1996).

**Table 5.4**  $^3J_{CC\alpha}$  couplings between Leucine and Isoleucine  $\delta$  methyls and their  $C^\alpha$  carbons

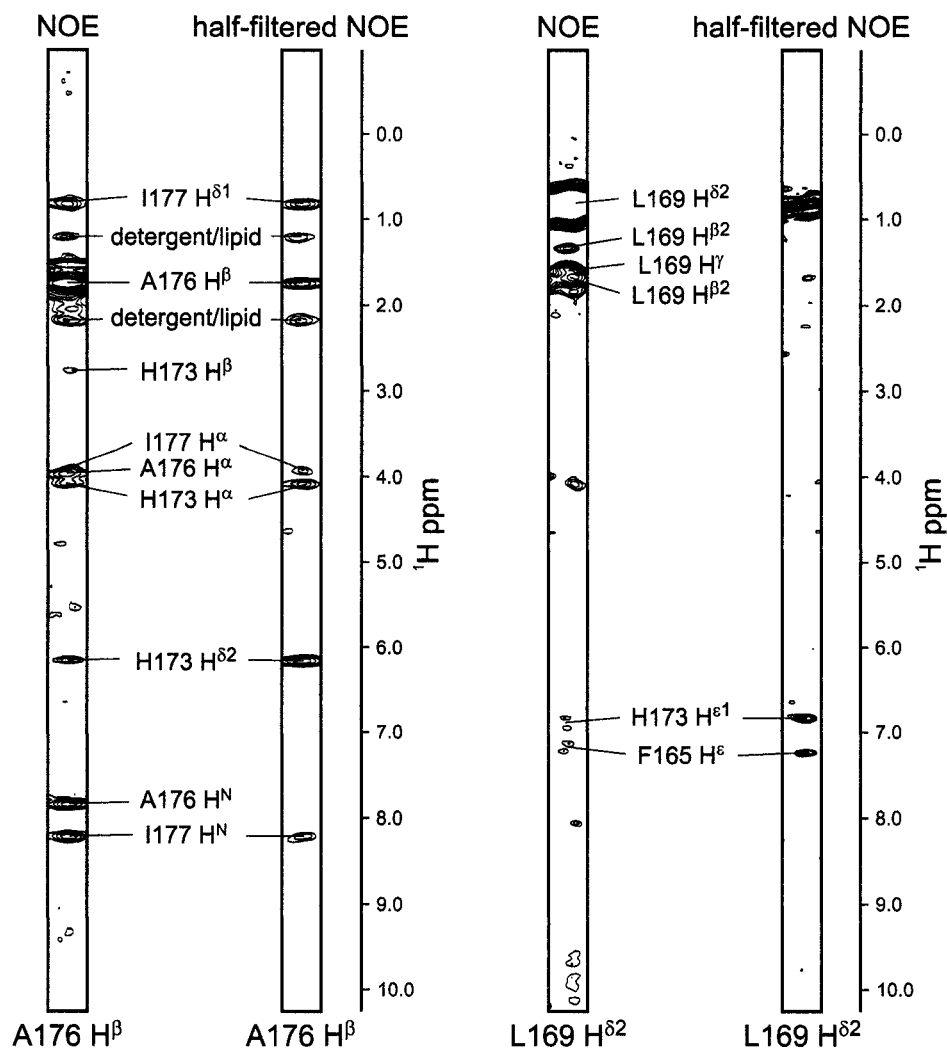
Residue	Resonance ( $^{13}C$ , ppm)	$^3J_{CC\alpha}$ (Hz)	$\phi_{trans}$	Stereospecific assignment (from NOE data)	Torsion angle for structure calculation
Leu162	(25.26)	$1.62 \pm 0.01$	$0.23 \pm 0.01$	Not determined	$\chi_2$ is not determined
	(22.84)	$0.76 \pm 0.02$	$\sim 0.0$	Not determined	
Leu166	(24.87)	$1.54 \pm 0.01$	$0.21 \pm 0.01$	$\delta_2$	$\chi_2$ is not determined
	(22.84)	$1.09 \pm 0.01$	$0.06 \pm 0.01$	$\delta_1$	
Leu169	(25.10)	$1.52 \pm 0.02$	$0.20 \pm 0.01$	$\delta_2$	$\chi_2$ is not determined
	(23.62)	$0.93 \pm 0.02$	$0.01 \pm 0.01$	$\delta_1$	
Leu170	(23.77)	$1.55 \pm 0.01$	$0.21 \pm 0.01$	Not determined	$\chi_2$ is not determined
	(23.85)	$1.40 \pm 0.01$	$0.16 \pm 0.01$	Not determined	
Leu171	(23.96)	$1.48 \pm 0.01$	$0.19 \pm 0.01$	$\delta_2$	$\chi_2 = 180^\circ$
	(23.56)	$0.46 \pm 0.01$	$\sim 0.0$	$\delta_1$	
Leu174	(24.48)	$1.54 \pm 0.01$	$0.21 \pm 0.01$	$\delta_2$	$\chi_2 = 180^\circ$
	(22.91)	$<0.10$	$\sim 0.0$	$\delta_1$	
Leu175	(24.63)	$1.56 \pm 0.01$	$0.21 \pm 0.01$	$\delta_2$	$\chi_2 = 180^\circ$
	(22.84)	$<0.10$	$\sim 0.0$	$\delta_1$	
Leu179	(24.79)	$1.98 \pm 0.01$	$0.35 \pm 0.01$	$\delta_2$	$\chi_2 = 180^\circ$
	(23.07)	$0.62 \pm 0.02$	$\sim 0.0$	$\delta_1$	
Leu187	(25.49)	$1.78 \pm 0.01$	$0.29 \pm 0.01$	Not determined	$\chi_2$ is not determined
	(22.60)	$0.97 \pm 0.01$	$0.02 \pm 0.01$	Not determined	
Ile156		$1.45 \pm 0.01$	$0.18 \pm 0.01$		$\chi_2$ is not determined
Ile177		$1.76 \pm 0.02$	$0.28 \pm 0.01$		$\chi_2$ is not determined
Ile181		$1.84 \pm 0.03$	$0.30 \pm 0.01$		$\chi_2$ is not determined
Ile183		$1.60 \pm 0.03$	$0.23 \pm 0.01$		$\chi_2$ is not determined

The formulae used to calculate  $^3J_{CC\alpha}$  and  $\phi_{trans}$  are the same as those listed on the footnote for Table 5.3.

#### 5.4.5 Assignment of intra- and inter-monomer distance restraints

Distance restraints for BNIP3 TMD were determined from the intensities of the  $^1\text{H}$ - $^1\text{H}$  nuclear Overhauser effects (NOE), and these correlations were recorded from 3D transient NOE experiments (NOESY). One of the experiments performed was  $^{13}\text{C}$ ,  $^{15}\text{N}$ -resolved NOESY-HSQC, which resolves the NOE correlations in the  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shift dimensions. The  $^{13}\text{C}$ ,  $^{15}\text{N}$ -NOESY-HSQC provides a source to record NOE observed to both aliphatic and amide protons. However, the sensitivity and resolution are particularly poor on the amide region, so I acquired an additional experiment, 3D  $^{15}\text{N}$ -resolved NOESY-HSQC. The two versions of the NOESY-HSQC experiments were collected on the 800 MHz spectrometer using the BioPack pulse sequences. The mixing time was set to be 50 msec, which reduces the possibility of having contaminations from spin diffusion.

For molecules in dimeric state, the NOE correlations observed in the 3D NOESY-HSQC spectra can arise from intra- or inter-monomer cross relaxations, or from a combination of the two, and the ability to discriminate between these correlations is needed as these NOE can provide a basis to define distances between the monomers for structure calculation using a simulated-annealing protocol. A half-filtered version of 3D NOESY-HSQC experiment which is performed on a mixture of isotopically labeled and unlabeled BNIP3 TMD peptides can provide such information (Zwahlen et al. 1997). This experiment was set such that only NOE from protons of the unlabeled monomer to protons of the labeled monomer were selected, thus producing a set of NOE cross-peaks that could only have arisen from inter-monomer contact (Figure 5.12).

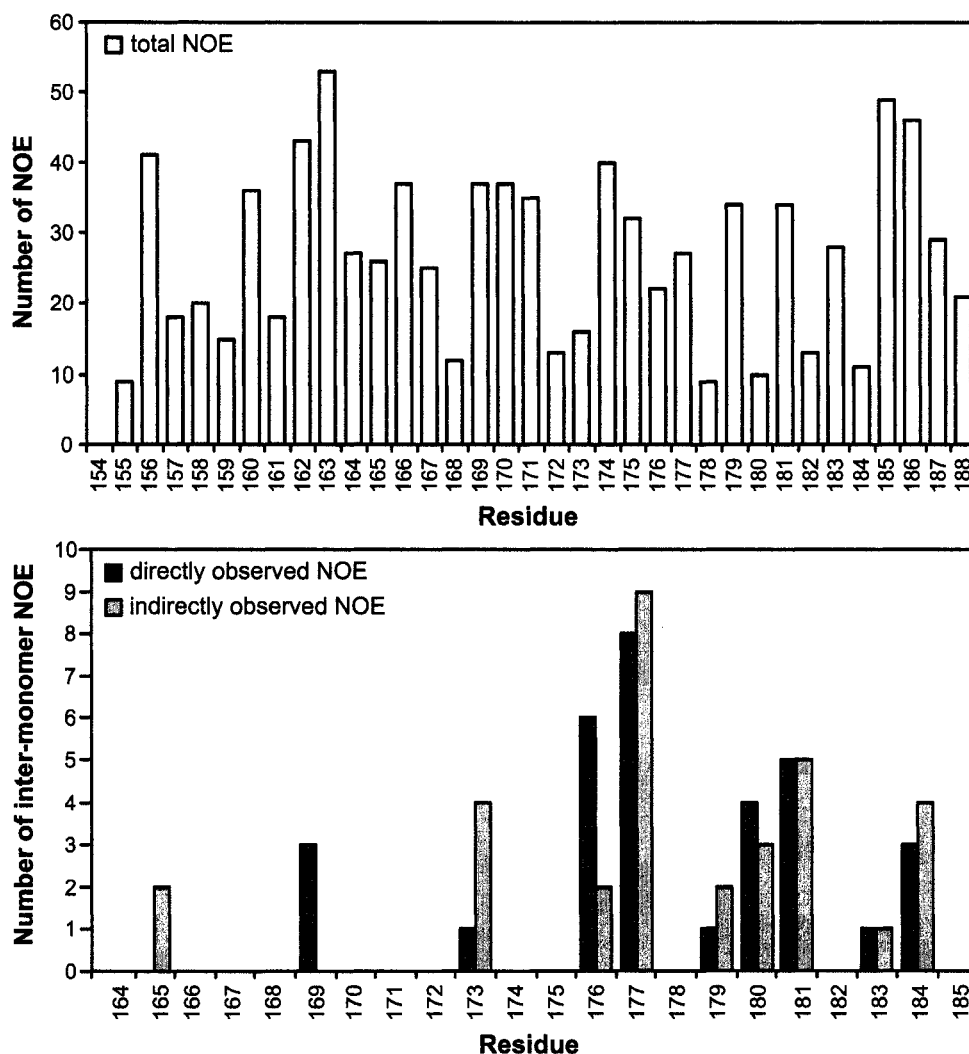


**Figure 5.12 Slices from 3D NOESY-HSQC and half-filtered NOESY-HSQC**

Slices of 3D NOESY-HSQC spectra provide NOE correlations that can arise from both intra- and inter-monomer cross relaxation. The inter-monomer NOEs can be distinguished by analyzing the correlations in the half-filtered NOESY-HSQC spectrum. H173 and F165 peaks are observed as singlets in the half-filtered NOESY-HSQC spectrum but as doublets in the regular NOESY-HSQC spectrum because protons from the  $^{13}\text{C}$ -labeled peptide, whose cross-peaks can be observed in the regular but not in the half-filtered spectrum, experience intra-residue nuclear spin-spin coupling with attached  $^{13}\text{C}$  atoms.

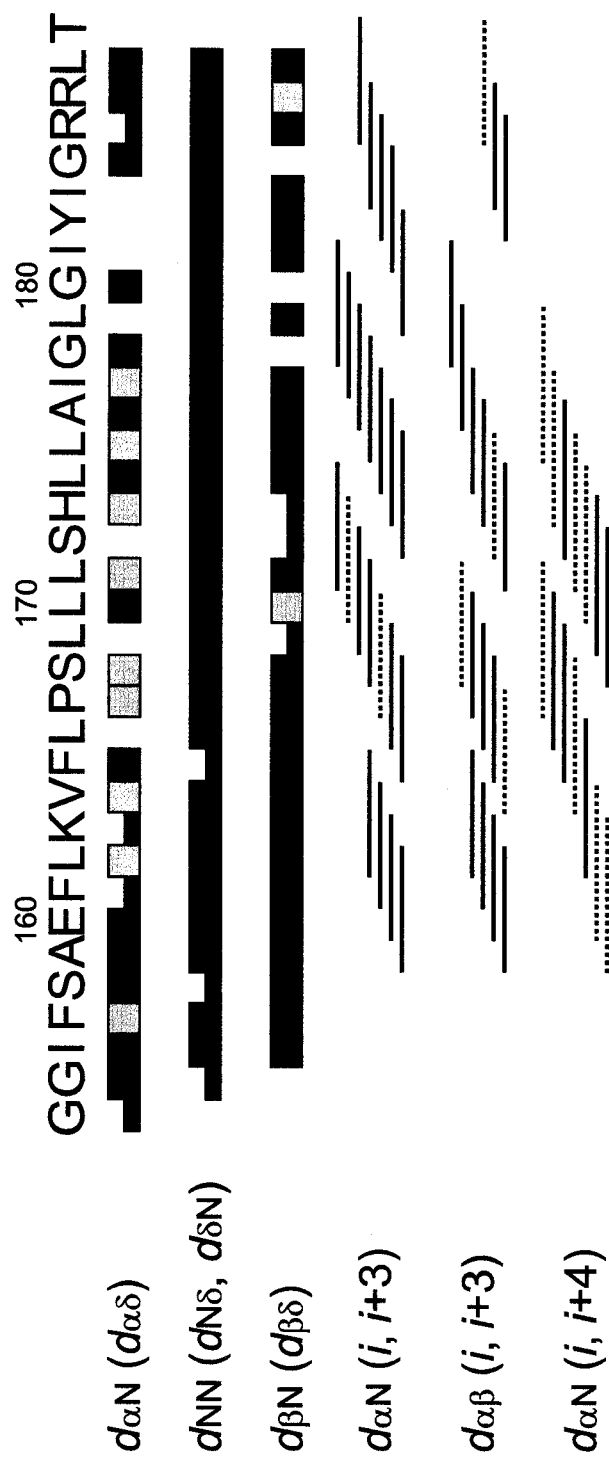
In assigning the NOE correlations, the identity of the direct  $^1\text{H}$  can be determined through a combination of directly observed  $^1\text{H}$  chemical shift and the indirect heteronuclear chemical shift. However, the generally poor resolution in the indirect  $^1\text{H}$  chemical shift dimension adds a challenge to assigning the NOE cross-peaks unambiguously. Because of that, incorporation of some structural information was needed to resolve this uncertainty and to guide the NOE assignment procedure. This approach allowed me to characterize 923 unambiguous NOE correlations and 32 inter-monomer NOE correlations. Of the 32 observed inter-monomer NOEs, 26 correlations are unique. Figure 5.13 shows the distribution of the NOE constraints per residue. Both the inter- and intra-NOE correlations were used to calculate the structure of the BNIP3 TMD dimer, and they were also used to confirm chemical shift assignments and the secondary structure components of the BNIP3 TMD during the refinement stages of structure calculation (Figure 5.14, see section 5.5).





**Figure 5.13 The number of NOE constraints per residue**

The histogram on the top shows the total number of NOE constraints, both intra- and inter-monomer, per residue. The histogram on the bottom shows the number of inter-monomer NOE constraints per residues. These inter-monomer constraints can be identified as correlations directly observed on the protons of the indicated residues (dark grey) or indirect correlations from protons of the indicated residues (light grey).



**Figure 5.14 Summary of NOE correlation data of the BNIP3 transmembrane domain peptide**

Sequential NOE correlations (first three rows) confirm chemical shift assignments, and medium range NOE correlations (the bottom three rows) indicate the BNIP3 transmembrane domain region that is in an  $\alpha$ -helical configuration. The sequential NOEs are represented as thick and thin bars to indicate strong and weak NOE intensities. The medium range NOEs are represented as lines connecting the two residues. The grey bars and the broken lines indicate ambiguous NOE assignments.

## 5.5 Calculation of the BNIP3 transmembrane domain dimer structure

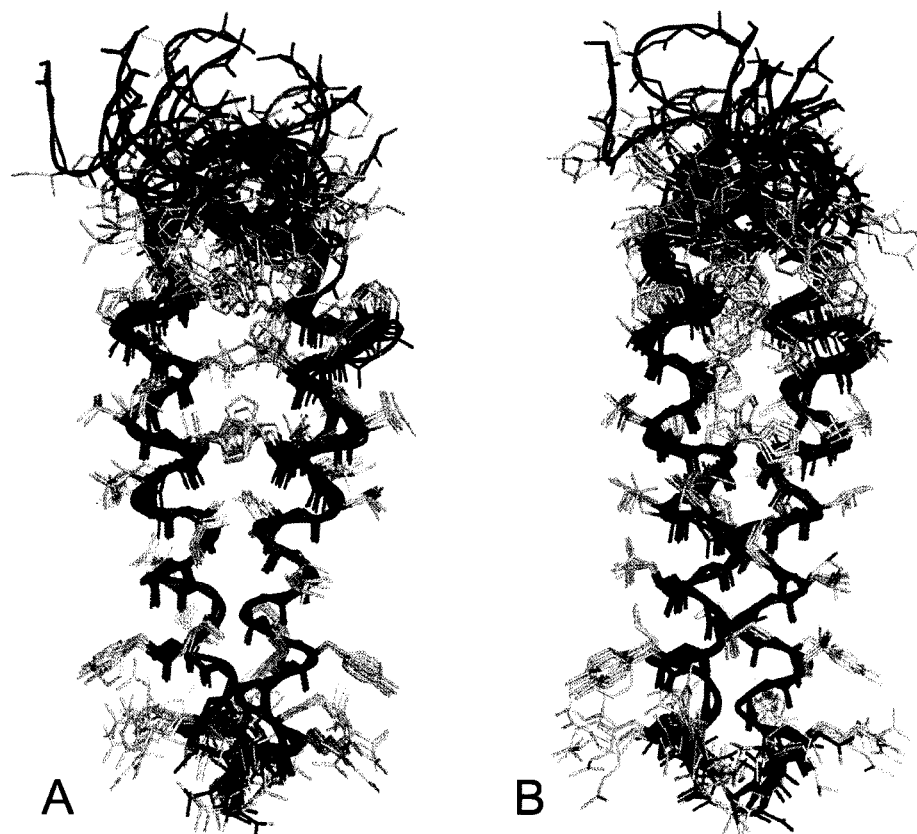
Calculation of the BNIP3 TMD dimer structure was performed with the software ARIA (Ambiguous Restraints for Iterative Assignment) version 1.2 (Nilges 1995; Linge and Nilges 1999; Linge et al. 2001; Linge et al. 2003). This software generates structure ensembles using the simulated annealing strategy codes of CNS version 1.1 (Brunger et al. 1998), and validates the geometry and restraints violation of the final structure ensemble using PROCHECK (Morris et al. 1992; Laskowski et al. 1993). Structure calculation with ARIA requires a list of assigned chemical shifts and a list of assigned, partially assigned, or uninterpreted NOE cross-peaks. Additional restraints, such as *J*-couplings, hydrogen bonds, residual dipolar couplings, torsion angles, disulfide bonds, and stereospecific assignments, can be specified to improve the quality of the structure.

The presence of a single set of chemical shifts for each atom in the BNIP3 TMD suggests that the monomers in the BNIP3 TMD dimer are symmetric, and that I can double all restraints each monomer for structure calculation. The process to calculate the structure of the BNIP3 TMD dimer can be divided into three phases: preliminary structure calculation, initial refinement, and final refinement.

### 5.5.1 Preliminary structure calculation

The aims of this phase are to get familiar with the ARIA code used for structure calculation and to collect additional restraints for subsequent refinement process. The calculations in this stage were done using the following restraints: an initial set of intra- and inter-monomer NOE restraints, intra-molecular hydrogen bond and backbone torsion angles ( $\phi$  and  $\psi$ ) defined from CSI analyses, and the  $\chi_1$  torsion angle restraints for Ile  $\gamma_2$ -methyls from  $^3J$ -coupling data. Inclusion of hydrogen bond and torsion angle restraints in these calculations helped generate structure ensembles with somewhat converged secondary structure (Figure 5.15).

The BNIP3 TMD dimer structures generated in this phase had low resolution because even though the general fold for the secondary structure is likely correct, the side chains may be placed with incorrect rotamers. Nevertheless, these structures provided structural information that could be used to collect additional intra- and inter-monomer NOE assignments, to evaluate the reliability of correlations in the intra- and inter-monomer NOE lists, to determine whether these assignments were consistent with distance restraints, and to identify whether the cross-peaks were assigned to the correct atoms. The structural data and the updated NOE lists were then used to make stereospecific assignments of methylene protons and a few of the Leu  $\delta$ -methyls and to define their respective  $\chi_1$  and  $\chi_2$  torsion angle restraints.



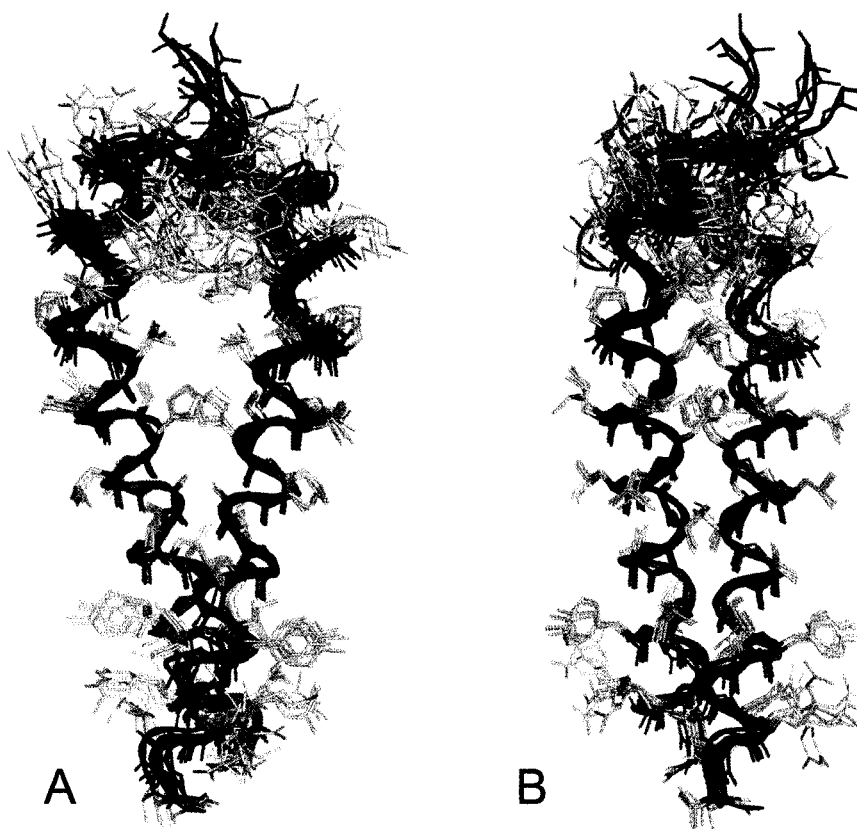
**Figure 5.15 Initial structure ensemble of the BNIP3 transmembrane domain dimer**

Shown is structure ensemble of the BNIP3 transmembrane domain dimer viewed from the front (A) and from the side (B). The amino terminus of the peptide is located on the top and the carboxy terminus at the bottom. The backbone atoms are in dark blue, and the side chains are in grey. A worm is drawn along the helix backbone to emphasize the secondary structure of the peptide. This ensemble represents a family of seven BNIP3 transmembrane domain dimer structures with the lowest energy, which were generated using preliminary distance (NOE) restraints and limited hydrogen bond and torsion angle restraints.

## 5.5.2 Initial structure refinement

Additional restraints collected from the preliminary structure calculation were incorporated in this phase to refine the BNIP3 TMD dimer structure. As shown in Figure 5.16, the structure ensemble generated in this phase is more converged, but the side chain atoms are still somewhat divergent. I also calculated BNIP3 TMD dimer structure with symmetry restraints applied (see Appendix E), but structures calculated with these restraints were not significantly more improved than those calculated without the restraints, suggesting that the use of symmetry restraints may not be necessary to calculate the structure of the BNIP3 TMD dimer.

Similar to the previous calculation stage, the structures generated in this refinement step were used to evaluate NOE peak assignments. I used the ambiguous peak list generated by ARIA to identify NOE peaks that were incorrectly assigned and to correct these assignments; to determine whether NOE peaks that I included in previous structure calculation were actual NOE peaks or noise; to identify NOE peaks that should have been assigned as ambiguous peaks due to overlapping chemical shift resonances with other atoms; and to assigned peaks that were specified as ambiguous assignments to the proper atoms. This evaluation also allowed me to identify NOE peaks that correspond to correlations with lipids. For example, two NOE peaks in the Ala176  $\beta$ -methyl can be assigned as correlation to Thr188  $\gamma$ -methyl (1.2 ppm) and Ile181  $H^{\beta}$  (2.1 ppm) (see Figure 5.12).



**Figure 5.16 Structure ensemble of the BNIP3 transmembrane domain dimer after initial refinement**

Shown is structure ensemble of seven BNIP3 transmembrane domain dimer structures generated using ARIA 1.2 with additional distance and torsion angle restraints incorporated. The ensemble is viewed from the front (A) and from the side (B).

Because the calculated structure indicates that Ala176 is not close to these residues, I knew that these NOE peaks could not represent correlations to Thr188 and Ile181, and that they must be correlations to lipid (see the peaks with asterisks in Figure 5.2). Therefore, I excluded these assignments from subsequent structure calculations. Even though the generated structure helped me identify correlations to lipids, I could not easily identify NOE peaks that correspond to correlations to the methyl group lipid or detergent micelles because these chemical shifts overlap with the chemical shifts of Ile and Leu methyl protons (see the peaks at 0.8 ppm on the  $^1\text{H}$  dimension in Figure 5.2). Initially, I specified these NOEs as ambiguous, but when structural data strongly suggest that such NOEs represent correlations to lipid, I flagged these assignments in the NOE peak lists, so that they are excluded in subsequent calculations.

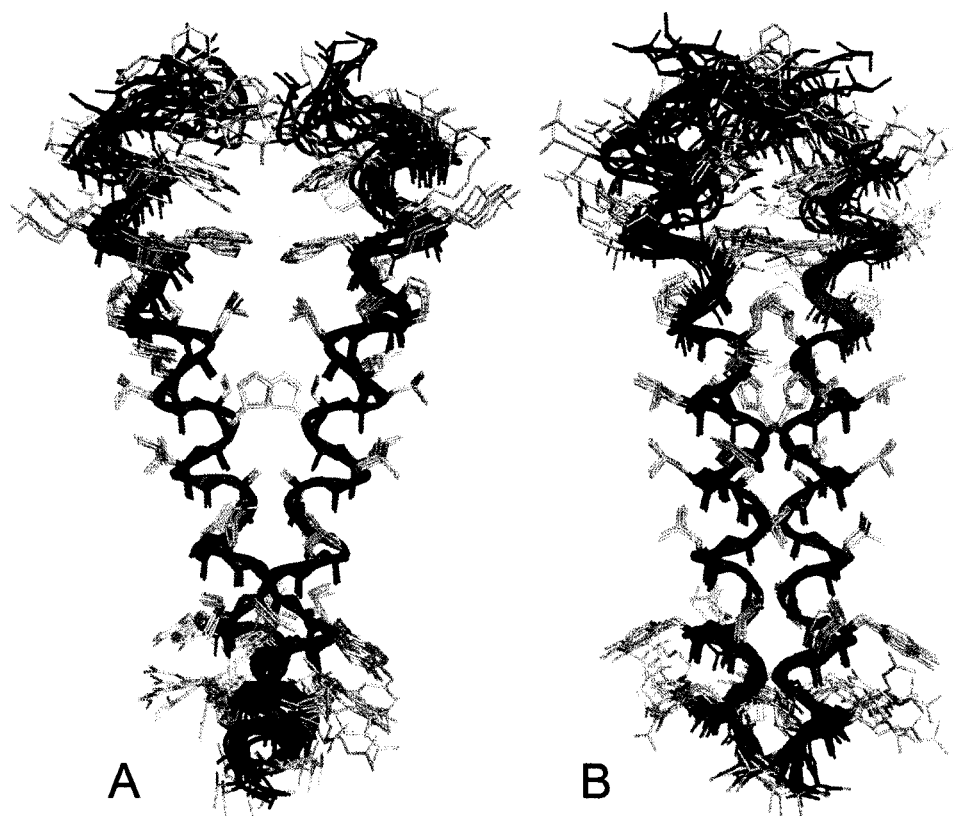
During NOE peak evaluation, I found that I had incorrectly assigned the chemical shift resonances for Leu170 and Leu174. These residues have very similar chemical shifts for  $\text{C}^\alpha$  and  $\text{C}^\beta$ , and because of that I confused Leu170 and Leu174 when performing assignment of the side chain atoms (section 5.4.2). This error required that I not only correct the chemical shift resonances of atoms in these residues, but also determine whether NOE correlations to these residues were properly assigned.

Structural information obtain from this phase also allowed me to expand intra- and inter-monomer NOE peak list, to define additional stereochemical assignments, and to determine additional  $\chi_1$  and  $\chi_2$  torsion angles for methylene protons and Leu  $\delta$ -methyls. In addition, I examined the structures and the data sets to ensure that an appropriate scaling factor was used for the inter-monomer data, which were comparatively scarce and lack meaningful auto peak intensities.



### 5.5.3 Final structure refinement

Structure calculation in the final refinement stage used the restraints collected in the initial refinement process, and the BNIP3 TMD dimer structure generated with these restraints has high resolution as the structure ensemble shows that both secondary structure and side chain rotamers are converged (Figure 5.17). During this process PROCHECK analysis was incorporated to evaluate final structure ensemble (Table 5.5). I defined a stable  $\alpha$ -helix region for RMS difference calculation by ARIA. At this stage, no additional restraints were collected, and modifications on the available restraints were minor. The lists of restraints used to calculate the final structure are listed in Appendix D, and ARIA code that I have modified from the original to suit the requirements for calculating the structure of the BNIP3 TMD dimer is presented in Appendix E.



**Figure 5.17 Final structure ensemble of the BNIP3 transmembrane domain dimer**

The final refinement process generated high resolution structures of the BNIP3 transmembrane domain dimer. The ensemble shows that structures in this family agree closely, and the middle region of the transmembrane helices (res 168-185), which is identified as  $\alpha$ -helical based on CSI data (see Figure 5.9), is well-defined.

**Table 5.5. NMR structural statistics and atomic RMS differences**

<b>Distance and dihedral restraints</b>	
Unambiguous NOE restraints	
Intra-residue	590
Inter-residue	
Sequential	183
Medium range	105
Long range	45
Inter-monomer	32
Hydrogen bond restraints	31
Torsion angle restraints	
backbone $\phi$	17
backbone $\psi$	17
side chain $\chi_1$	12
side chain $\chi_2$	4
<b>Structure statistics <sup>(1)</sup></b>	
Restraints violations	
distance ( $>0.3$ Å)	0
dihedral ( $>5^\circ$ )	0
Deviation from idealized geometry	
Bonds (Å)	$0.001 \pm 0.0001$
Angles ( $^\circ$ )	$0.307 \pm 0.005$
Impropers ( $^\circ$ )	$0.141 \pm 0.008$
Dihedrals ( $^\circ$ )	$42.210 \pm 0.222$
VDW	$1.900 \pm 0.394$
Deviation from experimental restraints	
NOE	$0.008 \pm 0.001$
Dihedrals ( $^\circ$ )	$0.189 \pm 0.040$
RMS deviation, res 168-185	
backbone atoms	$0.36 \pm 0.09$
all heavy atoms	$0.60 \pm 0.10$
Ramachandran analysis <sup>(2)</sup>	
%residues in most favored region	80.9
%residues in allowed region	17.6
%residues in generously allowed region	1.2
%residues in disallowed region	0.3

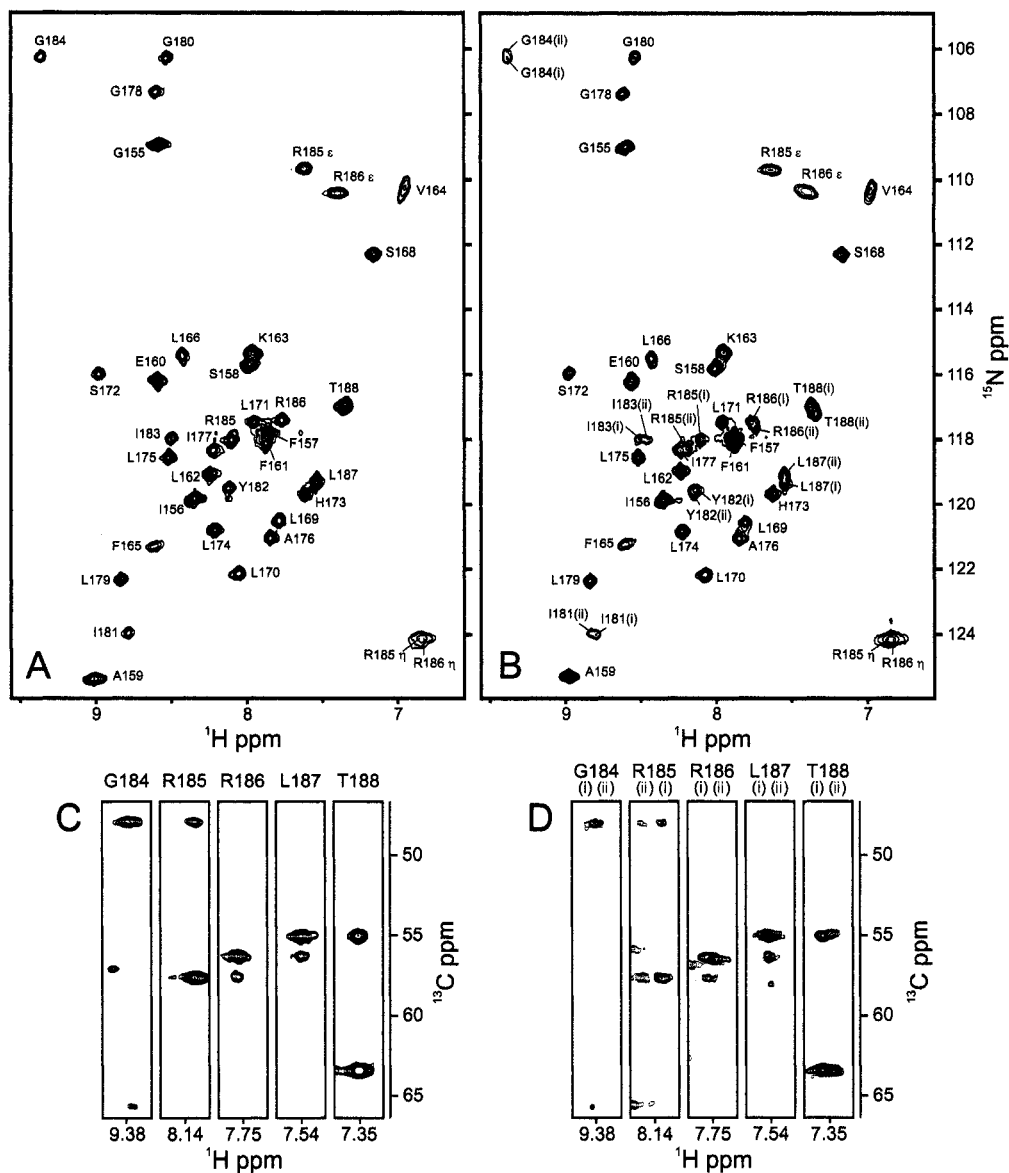
<sup>(1)</sup> Statistics are calculated and averaged over 20 lowest energy dimer structures out of a total of 50 calculated structures.

<sup>(2)</sup> Ramachandran statistics were determined using PROCHECK (Morris et al. 1992; Laskowski et al. 1993)

## 5.6 Different phospholipids can also modulate structure or stability of the BNIP3 transmembrane dimer

The improvements seen in NMR spectra with addition of DPPC to the BNIP3 TMD sample and the ability to solve the BNIP3 TMD dimer structure with a sample prepared in mixed DPC-DPPC micelles raise a question whether other types of added phospholipid have the same effects on the quality of BNIP3 TMD sample. Figure 5.18 shows that addition of dipalmitoylphosphatidic acid (DPPA), a similar phospholipid to DPPC but lacking the choline head group, can also improve the quality of NMR spectra similarly to addition of DPPC. Backbone assignment indicates that the  $H^N$  and N resonances of sample prepared in 10 mM DPPA are almost identical to those of sample prepared in 10 mM DPPC, suggesting that the BNIP3 TMD peptide may adopt similar conformation when prepared in the presence of DPPC or DPPA. This idea is supported by the  $^1H/^{13}C$ -HSQC spectra for the methyl region which show that peak dispersion and relative peak intensities for samples in DPPC and DPPA are similar (Figure 5.19).

However, a closer examination of the  $^1H/^{15}N$ -HSQC shows that there are two sets of chemical shifts for residues 181-188 when the BNIP3 TMD peptide is prepared in DPPA, and data from HNCA spectrum can distinguish the chemical shifts of these populations (Figure 5.18 D). A similar doubling can also be detected in the  $^1H/^{13}C$ -HSQC spectrum for the peaks that have similar chemical shifts to the Thr188  $\gamma$  methyl and Leu187  $\delta$  methyl (Figure 5.19).



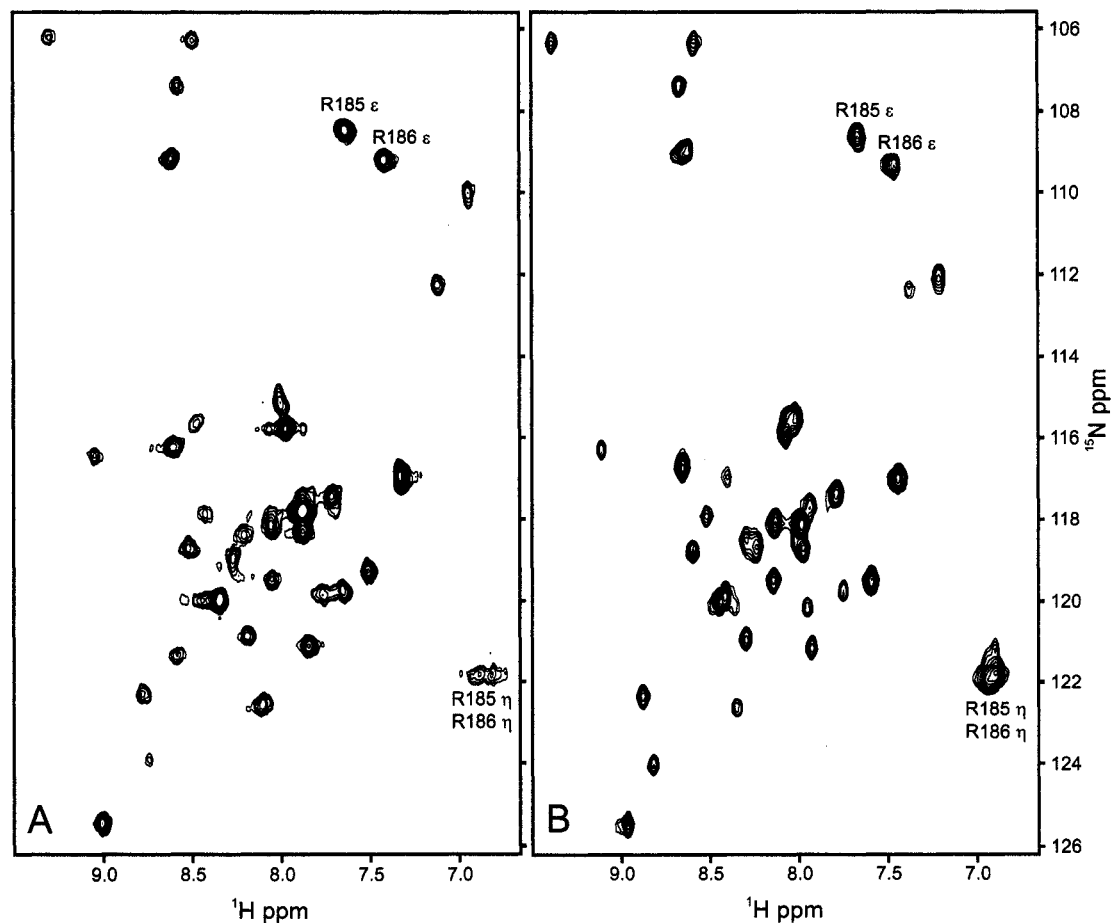
**Figure 5.18 The effects of adding DPPC or DPPA to a sample of BNIP3 transmembrane domain peptide**

The different effects resulting from adding 10 mM DPPC (A) or DPPA (B) are presented. HNCA experiments (C and D) allowed assignment of the backbone atoms in either condition. Peak doubling seen in panel B suggests that carboxy terminal of the BNIP3 transmembrane domain (res 181-188) may be in two conformations or two different environments in the sample prepared with DPPA. The population to which these conformations belong can be identified from the HNCA data.

**Figure 5.19**  $^1\text{H}/^{13}\text{C}$ -HSQC spectra acquired with samples prepared in 10 mM DPPC and 10 mM DPPA

This peak doubling implies that residues 181-188 are in two different conformations, and that the exchange rate between the two conformations or environments is slow enough in the NMR timescale that they can both be observed in the spectra. One possible explanation for this doubling may be the carboxy terminal residues are in different detergent/lipid environments, but with the data that I currently have, it is not possible to discriminate between this and many other possibilities.

Improvements were also seen in the  $^1\text{H}/^{15}\text{N}$ -HSQC spectra acquired with a sample prepared in the presence of palmitoyl-oleoylphosphatidic acid (POPA) or with a sample containing residual *Escherichia coli* lipids, which are composed of phosphatidylethanolamine (PE), phosphatidylglycerol (PG), and cardiolipin (Figure 5.20). Though an HNCA experiment was not performed on these samples, the similarities in chemical shift dispersion make it appealing to assign the  $^1\text{H}/^{15}\text{N}$  resonances by analogy to the spectra in Figure 5.18. The resemblance of the  $^1\text{H}/^{15}\text{N}$ -HSQC spectra of samples prepared in four different lipid conditions may suggest that the BNIP3 TMD dimer adopt the same, or at least similar, conformation regardless the type of phospholipid present in the sample, and that the BNIP3 TMD dimer structure could be determined with samples prepared in various lipid conditions. The presence of a single set of  $\text{H}^{\text{N}}/\text{N}$  peaks in the  $^1\text{H}/^{15}\text{N}$ -HSQC spectra indicates that addition of phospholipids stabilize a single conformation of the BNIP3 TMD dimer. However, the current data cannot explain how these phospholipids can induce such stabilization. Additional investigations, such as which part of the lipid interacts with the BNIP3 TMD peptide, may provide information about peptide-lipid interactions and enable elucidation of the stabilizing effects of added phospholipids.



**Figure 5.20**  $^1\text{H}/^{15}\text{N}$ -HSQC spectra of BNIP3 transmembrane domain peptide prepared in the presence of POPA or *E. coli* lipids

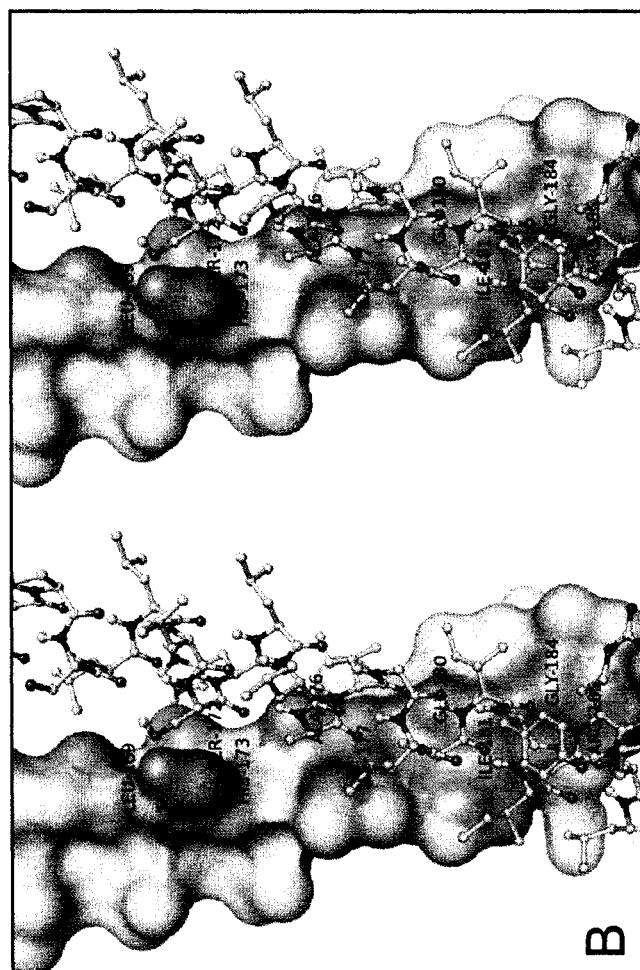
Improvements of the  $^1\text{H}/^{15}\text{N}$ -HSQC spectra with addition of 10 mM POPA (A) or when residual *E. coli* lipids are present (B) are presented. Spectrum A was acquired with the 600 MHz spectrometer, and spectrum B was acquired with the 500 MHz spectrometer. The similarities in the  $\text{H}^{\text{N}}/\text{N}$  cross-peak resonances with those presented in Figure 5.18 imply that the dimer structure may not alter when prepared in different phospholipid conditions.



## **6 Interpretation of the BNIP3 transmembrane domain dimer structure**

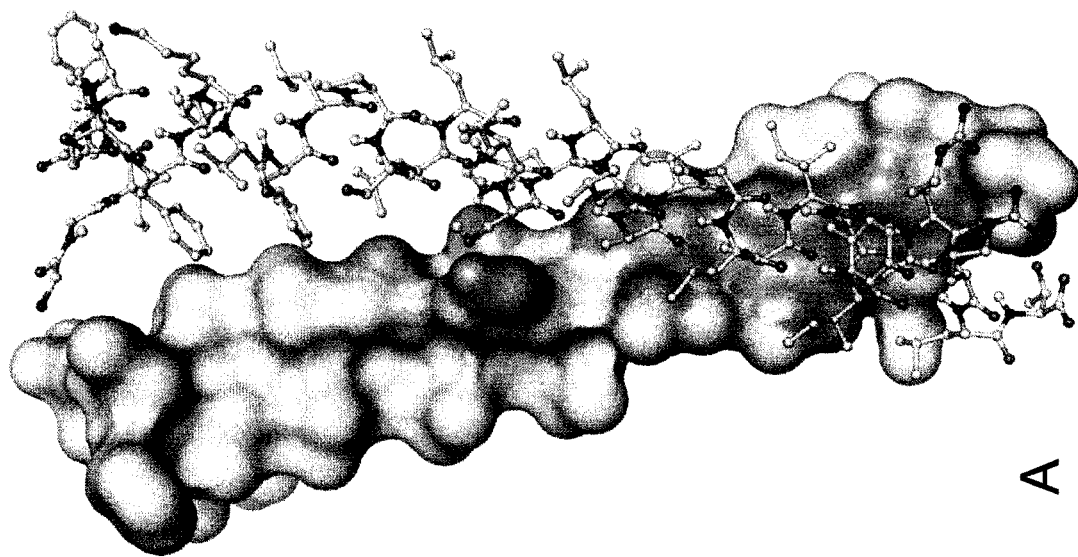
### **6.1 The NMR structure identifies the same interfacial residues as those predicted from saturation mutagenesis**

Consistent with the conclusions drawn from saturation mutagenesis data (Chapter 3), the structure of the BNIP3 transmembrane domain (TMD) dimer shows that the TMD dimerizes in parallel with a right handed crossing angle ( $\Omega$  of  $-32^\circ \pm 2^\circ$  for the well-defined region, res 168-185), and that the helix axes are separated by  $5.8 \pm 0.3 \text{ \AA}$ . The helices pack more closely at the carboxy-terminal half of the TMD (Figure 6.1), and this conformation explains why substitutions at residues 164-168 have no or very little effects on BNIP3 TMD dimerization on SDS-PAGE (section 3.6, Figure 3.8). The residues that are closely packed between the helices and form an interacting face for BNIP3 TMD dimerization are Leu169, Ser172, His173, Ala176, Ile177, Gly180, Ile181, Gly184, and Arg185 (Figure 6.1 B). Five of these interfacial residues (Ser172, His173, Ala176, Gly180, and Gly184) are the same as those identified as important for BNIP3 TMD dimerization from saturation mutagenesis analysis. Leu169, Ile177, Ile181, and Arg185 are classified as not important for dimerization based on saturation mutagenesis data, but the structure indicates that these residues participate in the stability of the BNIP3 TMD dimer. On the other hand, Ile183, which has been suggested to play a role in BNIP3 TMD dimerization based on mutagenesis data, is not part of the dimer interface, yet the structure alone is not sufficient to explain the importance of Ile183 for BNIP3 TMD dimerization.



**Figure 6.1 Structure of the BNIP3 transmembrane domain dimer**

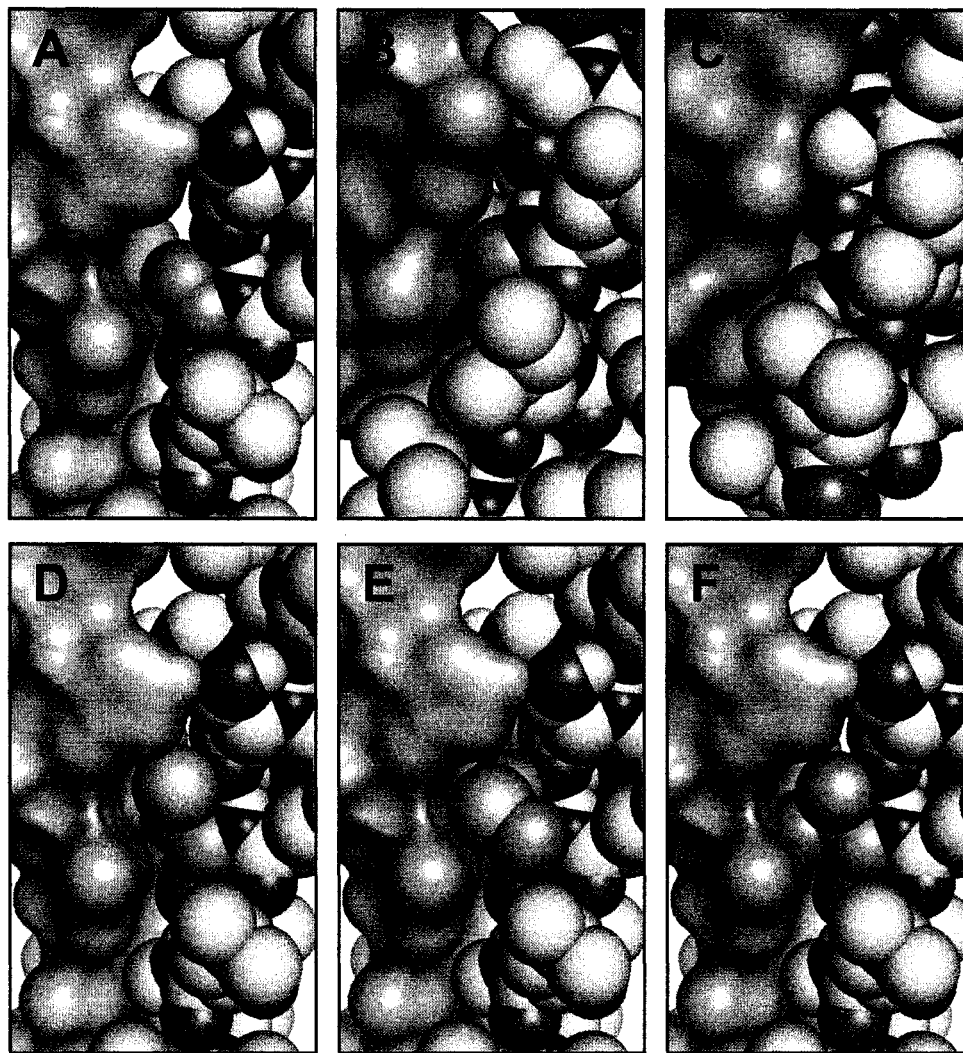
Presented is the structure of the BNIP3 transmembrane domain dimer visualized with Pymol (DeLano Scientific). The amino termini of the monomers are drawn on the top, and the carboxy termini are at the bottom. One of the monomer is presented in surface representation and the other as ball-and-stick to distinguish between the two. Highlighted in cyan on the surface representation are atoms located within 4 Å of atoms on the opposite monomer, creating an interacting face for dimerization. Panel A shows the full view of the dimer, and panel B shows stereo representation of the interacting region (res 169-185). The residues highlighted in cyan are labeled.



Because neither mutagenesis data nor BNIP3 TMD dimer structure provide enough data to confidently explain how Ile183 is involved in dimerization, additional information, such as how hydrophobicity and helicity of the BNIP3 TMD monomer can effect dimerization, is needed to better define the significance of Ile183 in the BNIP3 TMD helix-helix association.

## **6.2 Small residues provide a context for close packing between BNIP3 transmembrane domain helices**

Three of the interfacial residues observed in the BNIP3 TMD structure are small residues Ala176, Gly180, and Gly184, the same as those identified from the mutagenesis data. Close up views at these positions demonstrate that Ala176, Gly180, and Gly184 mediate close packing between BNIP3 TMD helices (Figure 6.2 A, B, C). Gly180 and Gly184 are packed very tightly between the two monomers. The space between helices is extremely narrow at these positions, such that an increase in side chain size, for example by replacing Gly with Ala or Val, cannot be accommodated at these positions. These data agree with the saturation mutagenesis results, where substitutions at Gly180 and Gly184 to any large amino acids disrupt BNIP3 TMD dimerization. Saturation mutagenesis data also indicate that substitutions of Gly184 to other small residues (Ala, Ser, Thr) significantly reduce dimerization, but the same substitutions at Gly180 completely disrupt the association, and these are consistent with the helix-helix packing being the tightest at position 180 and a little less tight at residue 184 (Figure 6.2 B and C).



**Figure 6.2 Close up view of the BNIP3 transmembrane domain dimer interface at the region for the tandem GxxxG motif**

Shown are close up views of the BNIP3 transmembrane domain dimer interface with one monomer presented in surface representation and the other in sphere representation. The top row shows the views centered at residues Ala176 (A), Gly180 (B), and Gly184 (C) (highlighted in cyan). The views are taken at different angles to emphasize differences in gap sizes between the monomers at these positions. The panels on the bottom row display the same view as panel A but with Ala176 substituted to Cys (D), Met (E), and Val (F). BNIP3 transmembrane domain structures with substitutions at residue 176 were generated with the program O (Jones et al. 1991).

The BNIP3 TMD dimer structure also shows that Ala176 is not packed as closely against the opposite monomer as Gly180 and Gly184 because the gap between monomers is wider at this position (Figure 6.2 A). Thus, the structure indicates that replacing Ala176 with a smaller Gly will not interfere with helix-helix packing, as has been observed from saturation mutagenesis (Figure 3.8), but replacing Ala176 with a larger residue may or may not be accommodated for BNIP3 TMD dimerization. Figures 6.2 D, E, and F show possible outcomes for replacing Ala176 with larger residues. These replacements were generated using the program O (Jones et al. 1991) from the BNIP3 TMD dimer structure containing the wild type sequence. Replacing Ala176 with a Cys, which adds a  $\gamma$ -thiol group at this position, can be accommodated without causing clashes that would interfere with dimerization (Figure 6.2 D), which is in agreement with the mutagenesis data where the Ala176Cys mutant dimerizes as strongly as the wild type (Figure 3.8). The structure also argues that replacing Ala176 with Ser can be accommodated for dimerization sterically, but the mutagenesis data demonstrate that this mutant can abrogate BNIP3 TMD dimer. Thus, the combination of structural and mutagenesis data implies that there are additional factors that can influence how this mutant dimerizes in SDS micelles, as has been discussed in section 3.11.

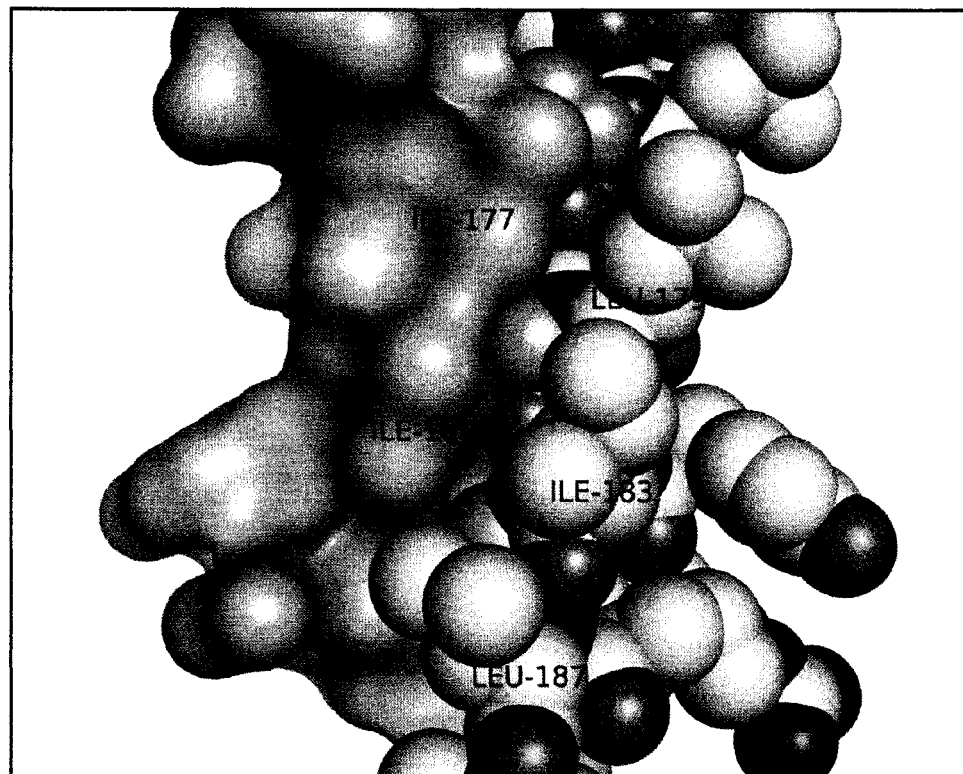
The effect of replacing Ala176 with Met depends on the side chain  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  rotamers used. Met side chain rotamers of  $(-60^\circ, -60^\circ, +60^\circ)$  allows the residue to fit snugly in the cleft that is formed between residues His173 and Ile177 on the opposite monomer (Figure 6.2 E). By comparison, other common Met rotamers,  $(-60^\circ, -60^\circ, -60^\circ)$  and  $(-60^\circ, 180^\circ, +60^\circ)$  (Ponder and Richards 1987; Chamberlain and Bowie 2004), will cause the Met side chain to clash with the other monomer. This side chain clash may explain why the Ala176Met mutation is slightly disruptive in SDS-PAGE (Figure 3.8).

Unlike replacements with Cys or Met, replacement of Ala176 with Val, which adds two  $\gamma$ -methyl groups, always causes clashes. Even though side chain rotamer places the  $\gamma_2$  methyl at the same position as the  $\gamma$ -thiol of Cys, the  $\gamma_1$  methyl collides with the backbone of the opposite monomer (Figure 6.2 F). Similarly, other  $\beta$ -branched amino acids (Ile and Thr) cannot be tolerated at position 176. Bulkier amino acids (Leu or Phe) will not be tolerated either because they will not fit in the crevice formed by the BNIP3 TMD dimer at position 176. Overall, the structural data imply that size, shape, and side chain rotamers determine the types of residue can be accommodated at position 176 for BNIP3 TMD dimerization.

### **6.3 BNIP3 TMD dimer exhibits a ridges-to-grooves helix-helix packing**

A pair of helices can pack in a knobs-to-holes fashion or in a ridges-to-grooves manner (Crick 1953; Richmond and Richards 1978; Chothia et al. 1981; Walther et al. 1996; Bowie 1997). The knobs-to-holes interaction entails having the side chain of a residue from one helix (the knob) protruding into the space between side chains from the other helix (the hole), and the ridges-to-grooves packing implicates a ridge of side chains on one helix fitting into the groove on the other helix. Helix-helix interactions involving these packing motifs have been observed in the M1 and M2 helices of the potassium channel KcsA and in the transmembrane domain of  $\text{Ca}^{2+}$ -ATPase (Lee 2002), as well as in the transmembrane domains of photosynthetic reaction centers, bacteriorhodopsin, and cytochrome C oxidase (Langosch and Heringa 1998).

The structure of the BNIP3 TMD dimer shows that helix-helix packing at Ala176 may fall into the knobs-to-holes category because the  $\beta$ -methyl packs well at the cavity between His173, Ala176, and Ile177 on the opposite monomer (Figure 6.2 A). In contrast, packing at Gly180 and Gly184 cannot be classified into the same category, because in the absence of a side chain, these two Gly residues do not form knobs. However, a closer examination of the BNIP3 TMD dimer structure indicates that the small interfacial residues adopt the other helix-helix packing motif, the ridges-to-grooves packing (Chothia et al. 1981; Chothia 1984; Walters and DeGrado 2006). In this helix-helix packing motif, Ala176, Gly180, and Gly184 form a small ridge on the surface of the BNIP3 TMD helix (Figure 6.1 B) that fits into a groove on the opposite helix, created by the Ala176-Gly180-Gly184 ridge and the Ile177-Ile181 ridge. The structure also shows that part of the Ile177-Ile181 ridge can fit into the groove created by the Ala176-Gly180-Gly184 ridge and the Leu179-Ile183 ridge (Figure 6.3). With helix-helix crossing angle ( $\Omega$ ) of  $-32^\circ$ , the geometry for the ridges-to-grooves helix-packing motif in the BNIP3 TMD dimer deviates to some degree from the ideal geometry ( $\Omega$  of  $-50^\circ$ ) (Chothia et al. 1981; Chothia 1984). However, the BNIP3 TMD dimer structure clearly indicates that ridges-to-grooves helix-packing is involved in this association, and this packing may promote van der Waals interactions between the two helices.



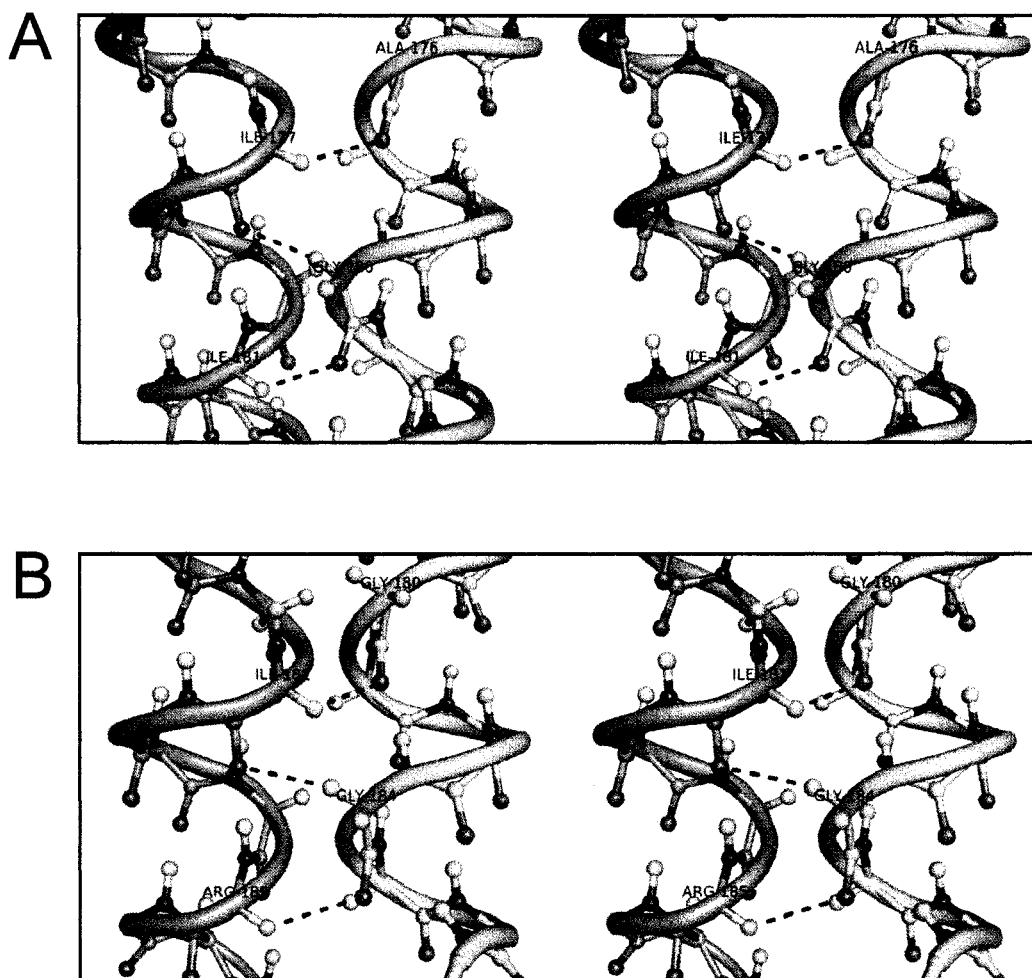
**Figure 6.3 Ridges-to-grooves helix-packing between the monomers of the BNIP3 transmembrane domain dimer**

A close-up view of the BNIP3 transmembrane domain dimer showing residues 176-187. One of the monomer is in surface representation and the other in sphere representation. Ridges on the BNIP3 transmembrane domain surface are formed by the small interfacial residues Ala176, Gly180, and Gly184 (highlighted in cyan) and by large aliphatic residues that flank the small residues (labeled). These ridges pack into the grooves formed between two ridges on the opposite helix. This helix-helix packing motif may promote inter-monomer van der Waals interactions for the stability of the BNIP3 transmembrane domain dimer.



## 6.4 The AxxxGxxxG motif promotes a network of inter-monomer backbone-to-backbone $C^\alpha-H^\alpha\cdots O=C$ hydrogen bonds

The GxxxG motif has been associated with the formation of inter-monomer  $C^\alpha-H^\alpha\cdots O$  hydrogen bonds between the helix backbones or between the helix backbone and the side chains, and this interaction can contribute to the stability of TMD helix-helix association (Senes et al. 2001; Kleiger et al. 2002). It has also been reported that the  $C^\alpha-H^\alpha\cdots O$  hydrogen bond is strongly biased for parallel, right-handed helix-helix packing (Senes et al. 2001), and because the BNIP3 TMD dimer falls into this category, this inter-monomer interaction may be involved in the stability of this self-association. Detailed examination of the BNIP3 TMD dimer structure indicates that residues that make up the tandem GxxxG motifs, as well as Ile177, Ile181, and Arg185, indeed participate in a network of backbone-to-backbone  $C^\alpha-H^\alpha\cdots O=C$  contacts (Figures 6.4, Table 6.1). Despite having the  $d_H$  and  $d$  values deviate somewhat from the ideal values, the inter-monomer  $C^\alpha-H^\alpha\cdots O=C$  hydrogen bonds between Gly184-Ile181 and Arg185-Gly184 can be classified into this category because they meet the criteria used by Senes and co-workers for identification of potential inter-monomer hydrogen bonds in interacting TMD helices ( $d_H < 3.5$  Å and  $\zeta > 120^\circ$ , or  $\zeta > 90^\circ$  when  $d_H < 3.0$  Å) (Senes et al. 2001) and the geometry range observed from crystal structures ( $\zeta$  between  $90^\circ$  and  $180^\circ$ ,  $\xi$  between  $100^\circ$  and  $170^\circ$ , and  $\theta$  between  $0^\circ$  and  $90^\circ$ ) (Vargas et al. 2000). Based on these criteria, each BNIP3 TMD monomer donates five  $C^\alpha-H^\alpha\cdots O=C$  bonds to the other monomer.



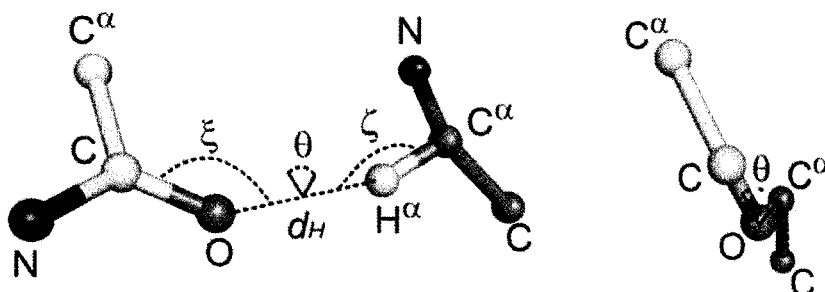
**Figure 6.4 The network of inter-monomer  $C^{\alpha}-H^{\alpha} \cdots O=C$  hydrogen bonds in BNIP3 transmembrane domain dimer**

Presented are stereo image representations of the BNIP3 transmembrane domain dimer showing residues 176-181 (A) and 180-185 (B). The close packing between the monomers promote inter-monomer  $C^{\alpha}-H^{\alpha} \cdots O=C$  contacts mediated by the interfacial residues and residues  $i+1$  or  $i+3$  on the opposite monomer (dashed lines). Shown on panels A and B are five of the ten backbone-to-backbone hydrogen bonds, observed on one side of the BNIP3 transmembrane domain dimer. Table 6.1 lists the average geometry of these contacts calculated from 20 BNIP3 transmembrane domain dimer structures in the ensemble of structures calculated by ARIA.

**Table 6.1**      **Geometry of inter-monomer  $C^\alpha-H^\alpha \cdots O=C$  contacts in the BNIP3 transmembrane domain dimer**

Donor	Acceptor	$\zeta$ (°)	$\xi$ (°)	$\theta$ (°)	$d_H$ (Å)	$d$ (Å)
Ideal values (Vargas et al. 2000)						
$C^\alpha - H^\alpha$	$O = C$	180	120	0	$\leq 2.7$	$\leq 3.8$
Average of all 40 interactions in the 20 lowest energy structures						
Ile177 – $H^\alpha$	Ala176 – O	$149 \pm 4$	$114 \pm 4$	$-54 \pm 10$	$2.5 \pm 0.2$	$3.5 \pm 0.2$
Gly180 – $H^\alpha$	Ile177 – O	$140 \pm 6$	$132 \pm 5$	$92 \pm 8$	$2.5 \pm 0.1$	$3.4 \pm 0.1$
Ile181 – $H^\alpha$	Gly180 – O	$134 \pm 6$	$116 \pm 6$	$-24 \pm 9$	$2.3 \pm 0.1$	$3.1 \pm 0.1$
Gly184 – $H^\alpha$	Ile181 – O	$144 \pm 8$	$115 \pm 7$	$96 \pm 13$	$3.1 \pm 0.3$	$4.1 \pm 0.3$
Arg185 – $H^\alpha$	Gly184 – O	$134 \pm 8$	$102 \pm 9$	$-19 \pm 18$	$3.4 \pm 0.4$	$4.2 \pm 0.4$

The nomenclature used to define the geometry parameters are as described previously (Derewenda et al. 1995; Senes et al. 2001):



$\zeta$  is the  $C^\alpha-H^\alpha-O$  angle

$\xi$  is the  $H^\alpha-O-C$  angle

$\theta$  is the dihedral angle between the  $C^\alpha-H^\alpha$   
bond and the  $O=C$  bond

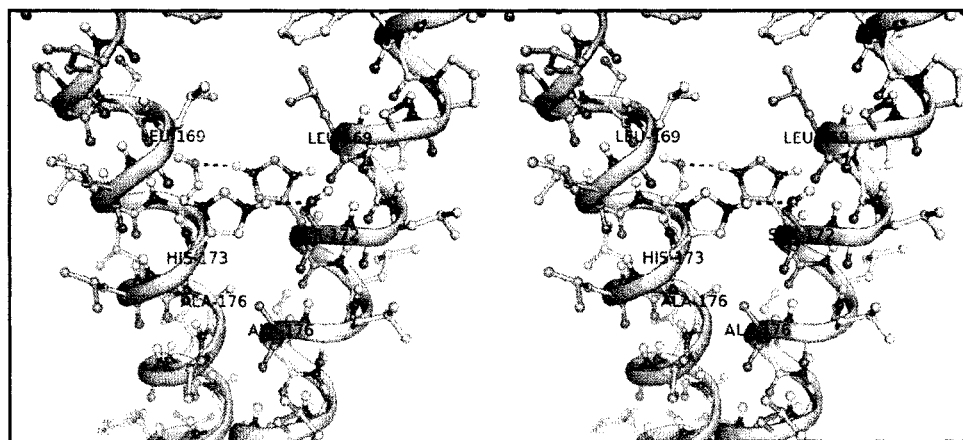
$d_H$  is the distance between  $H^\alpha$  and O

$d$  is the distance between  $C^\alpha$  and O

The  $C^\alpha-H^\alpha \cdots O=C$  hydrogen bond has an estimated energy between 1.9 – 3 kcal mol<sup>-1</sup>, or about half of the conventional  $N-H \cdots O$  hydrogen bond (Vargas et al. 2000; Scheiner et al. 2001). Alone, this interaction contributes significantly to the stability of TMD helix-helix associations, and in a network that coordinates multiple inter-monomer  $C^\alpha-H^\alpha \cdots O$  hydrogen bonds, this contribution is even greater. Because helix-helix packing of the BNIP3 TMD dimer can promote the formation of up to ten inter-monomer  $C^\alpha-H^\alpha \cdots O$  hydrogen bonds, these interactions can provide as much as 19 – 30 kcal mol<sup>-1</sup> for the stability of the TMD dimer. In addition, because these inter-monomer hydrogen bonds involves the tandem GxxxG motifs, as well as residues flanking these motifs, one can argue that the A176xxxG80xxxG184 motif contributes to the stability of the BNIP3 TMD dimer by mediating close packing interactions between the TMD helices and promoting the formation of inter-monomer  $C^\alpha-H^\alpha \cdots O$  hydrogen bonds, and that substitutions that replace Ala176, Gly180, or Gly184 will disrupt BNIP3 TMD dimerization not only by causing a loss of favorable close packing interactions, but also by breaking the network of inter-monomer  $C^\alpha-H^\alpha \cdots O$  hydrogen bonds.

## 6.5 The structure of BNIP3 transmembrane domain dimer confirms inter-monomer interaction between Ser172 and His173 side chains

The A176xxxG180xxxG184 motif also provides a context for inter-monomer hydrogen bonds between His173 and Ser172. Being two helix turns away from the closest contact point, the distance between His173 and Ser172 does not allow the formation of backbone-to-backbone  $C^{\alpha}-H^{\alpha}\cdots O=C$  hydrogen bonds ( $d_H \geq 4.4$  Å). However, the BNIP3 TMD dimer structure shows that inter-monomer hydrogen bonds can form between His173  $H^{\epsilon 2}$  and Ser712  $O^{\gamma}$  (Figure 6.5). It is quite remarkable that even though the BNIP3 TMD dimer structure was calculated in the absence of NOE correlations involving the His173 exchangeable side chain protons or Ser172  $H^{\beta 2}$  and  $H^{\beta 3}$  due to low sensitivity, correlations to His173 observed on Leu  $\delta$ -methyls and Ala176  $\beta$ -methyl (see Figure 5.12), as well as other experimentally derived restraints, were sufficient to define the side chain rotamer for His173 and to position this residue within distance for hydrogen bond formation with Ser172 from the opposite monomer. Moreover, the structure of the BNIP3 TMD dimer generated with these data confirms the inferences from saturation mutagenesis data regarding inter-monomer hydrogen bonds between His173  $H^{\epsilon 2}$  and Ser712  $O^{\gamma}$  (section 3.11).



**Figure 6.5 His173 and Ser172 form inter-monomer hydrogen bonds**

Stereo image representation of the BNIP3 transmembrane domain dimer interface showing residues 169-176. Both monomers are in ball-and-stick representation. Highlighted and labeled are the interfacial residues Leu169, Ser172, His173, and Ala176. The dash lines indicate inter-monomer hydrogen bonds that formed between His173 H $\epsilon^2$  and Ser172 O $\gamma$ .

## **6.6 Leu169 is at the dimer interface but contributes minimally to the stability of the BNIP3 transmembrane domain dimer**

Even though it is located at the dimer interface and within distance for close contact with residues on the other monomer, Leu169 does not seem to contribute significantly in the stability of the BNIP3 TMD dimer. The separation between Leu169 and Ser168 on the opposite monomer does not satisfy backbone-to-backbone  $C^\alpha-H^\alpha \cdots O=C$  hydrogen bond ( $d_H > 9 \text{ \AA}$ ), and helix-helix distance also prevents backbone-to-side chain hydrogen bonds between Leu169  $C^\alpha-H^\alpha$  and Ser172  $O^\gamma$  ( $d_H \geq 7.4 \text{ \AA}$ ). Moreover, helix-helix packing at position 169 is not comparable to those seen at residues 176, 180, and 184 (see Figure 6.1), suggesting that packing interactions may not be crucial at this position. These observations correspond well with saturation mutagenesis data, which indicate that Leu169 is not very important for BNIP3 TMD dimerization even though it is located on the interacting face of the transmembrane helix (section 3.7).

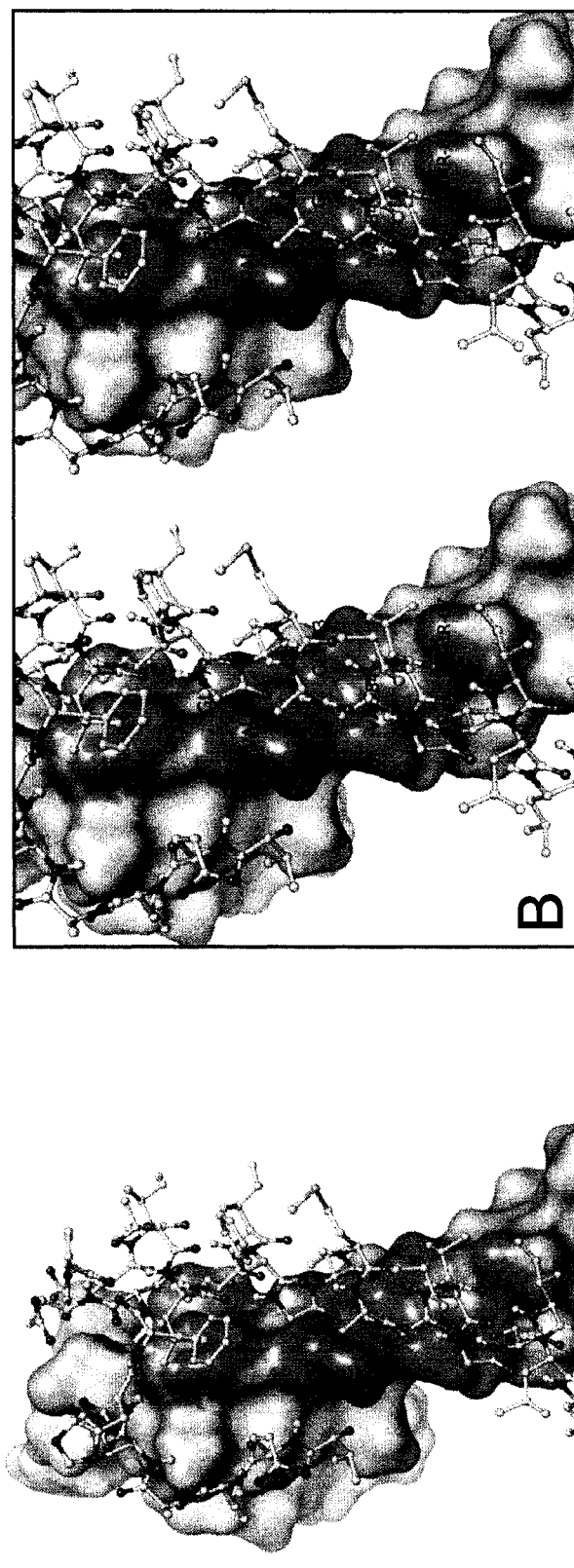
## **6.7 The structure is not sufficient explain the importance of Ile183 for BNIP3 transmembrane domain dimerization**

Thus far, the BNIP3 TMD dimer structure is able to confirm structural inferences from the mutagenesis data, such as the orientation of the dimer crossing, the packing interactions mediated by the A176xxxG180xxxG184 motif, and the inter-monomer hydrogen bonds between His173 and Ser172. The structure also can be used to explain why some Ala176 mutants have wild type-like phenotypes on SDS-PAGE despite close helix-helix packing at this position, and why Leu169 is less significant for dimerization even though it is located at the interacting face. However, the structure is not sufficient to explain why mutations at Ile183 can perturb BNIP3 TMD dimerization in SDS micelles. Examination of the BNIP3 TMD dimer structure indicates that the disruptive phenotype of Ile183Phe mutant in SDS-PAGE could be caused by destabilization of the BNIP3 TMD helix: replacing Ile183 with Phe can cause intra-molecular clashes with Leu179 or Leu187 depending on the side chain rotamers ( $[\chi_1, \chi_2]$  of  $[-60, -60]$  or  $[180, -60]$ , respectively). In contrast, the structure shows that replacing Ile183 with Met, Leu, or Val would not cause clashes, so these replacements would not be expected to affect dimerization. Clearly, because Met, Leu, and Val are hydrophobic, polarity should not be a factor that influences the disruptive phenotypes of these substitutions. However, the BNIP3 TMD structure cannot explain why these mutants can perturb BNIP3 TMD dimer, suggesting that there may be certain properties required at position 183 for the stability of the BNIP3 TMD dimer that cannot be easily identified from the structural data alone.



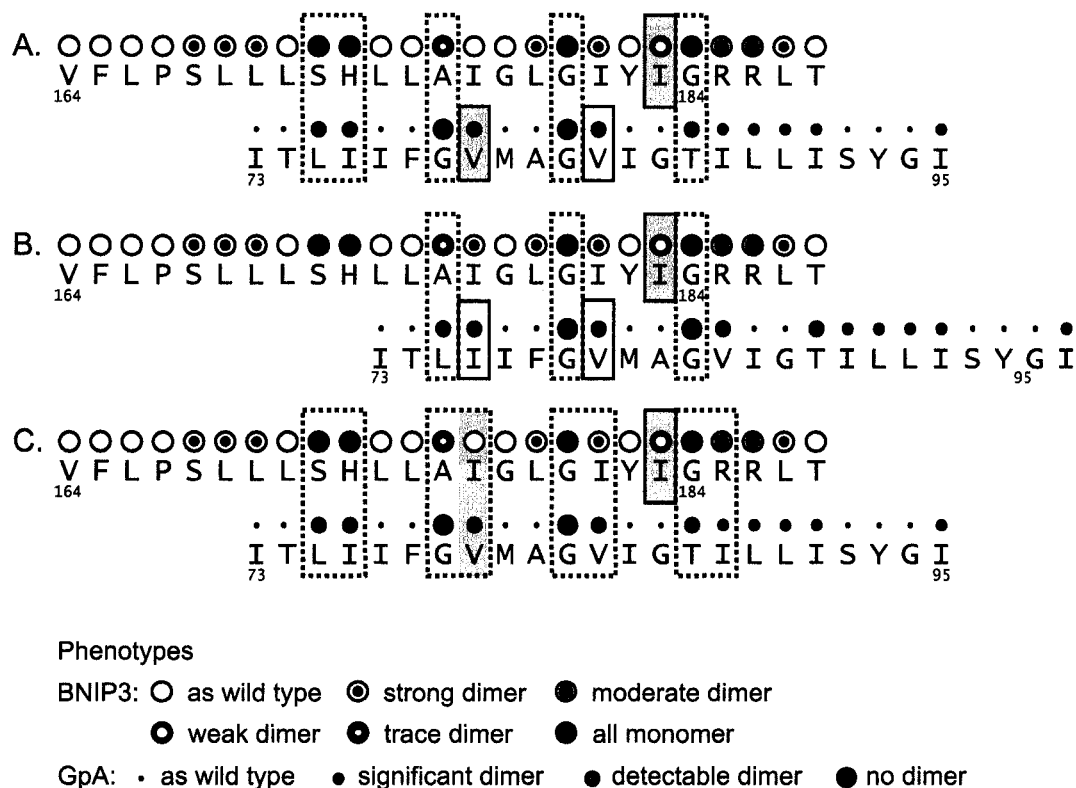
## 6.8 Comparison of BNIP3 and GpA transmembrane domain dimer structures

Another TMD dimer whose structure has been solved using solution NMR spectroscopy is that of glycophorin A (GpA; Figure 6.6) (MacKenzie et al. 1997). Similar to the BNIP3 TMD dimer, the GpA TMD dimer has a right-handed crossing ( $\Omega = -39^\circ \pm 5^\circ$ ), but the helix-helix contacts are the closest at the amino terminus of the TMD. Both BNIP3 and GpA TMD sequences contain the GxxxG motif, and this motif participates in the stability of the TMD helix-helix association in both systems. Yet, because the BNIP3 TMD sequence consists of two GxxxG motifs, the A176xxxG180 and G180xxxG184 motifs, there are two ways to align the two TMD sequences to compare the similarities between the BNIP3 TMD and the GpA TMD dimerization: 1) by aligning the GpA GxxxG motif with the BNIP3 A176xxxG180 motif (Figure 6.7 A and Figures 6.8), and 2) by aligning the GpA GxxxG motif with the BNIP3 G180xxxG184 motif (Figure 6.7 B and Figure 6.9). Relying only on saturation mutagenesis does not allow to distinguish the better TMD sequence alignment because both indicate that a few of the residues that are important for dimerization in one system do not have a match in the other system (Figures 6.7 A and B) (Lemmon et al. 1992; Sulistijo and MacKenzie 2006). However, when the residues that are identified important for dimerization from the structural analysis are included in this comparison, the alignment of the GpA GxxxG motif and the BNIP3 A176xxxG180 motif is the better even though this alignment still shows a mismatch for residue 183 of the BNIP3 TMD (Figure 6.7 C).



**Figure 6.6 Structure of the GpA transmembrane domain dimer**

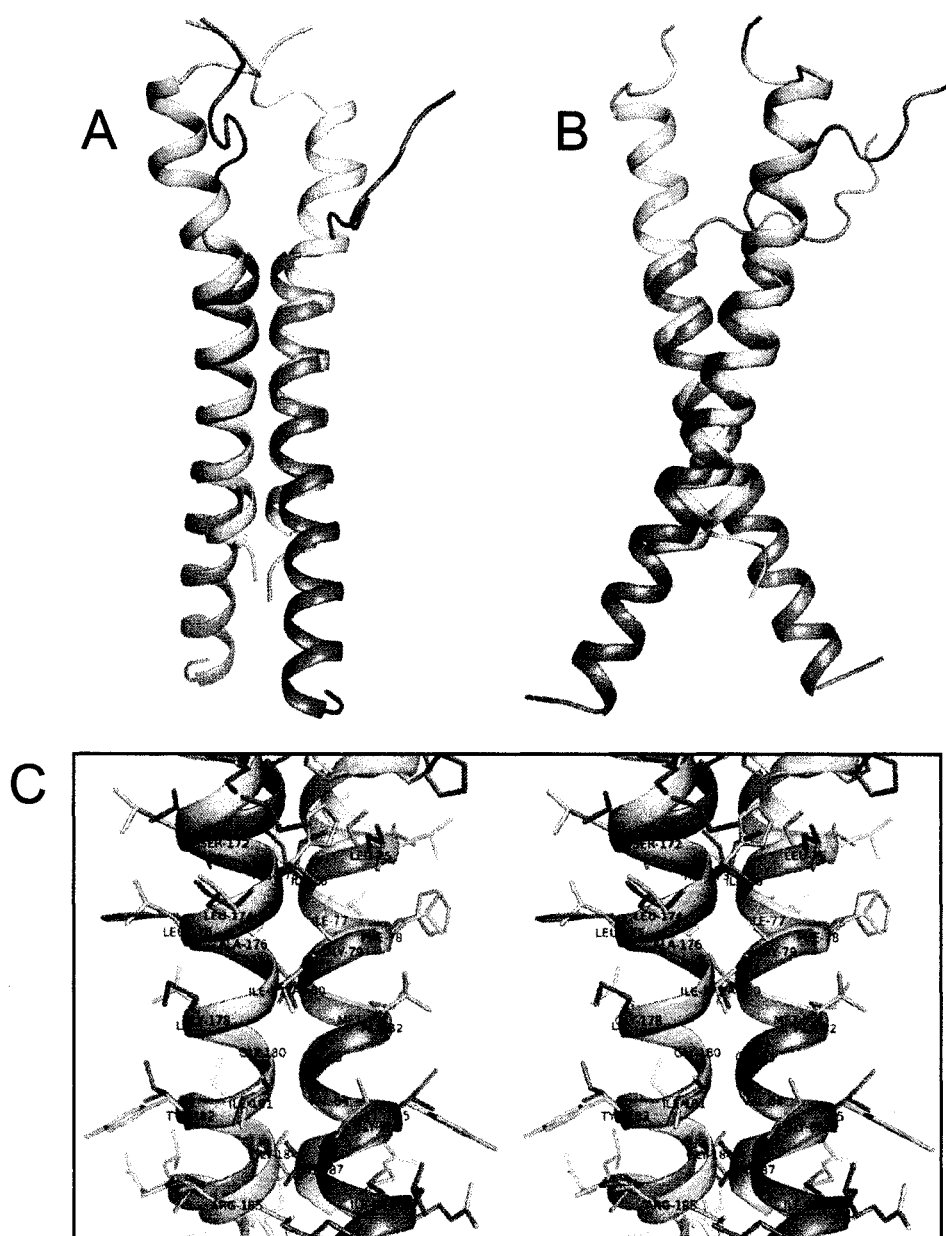
Presented are (A) the structure of the glycoporphin A transmembrane domain dimer (Mackenzie et al 1997, PDB accession number 1AFO) and (B) the stereo representation of the interacting region (res 75-88). Highlighted on the surface representation are atoms located within 4 Å of atoms on the opposite monomer, creating an interacting face for dimerization. The interfacial residues are labeled in panel B. The GpA transmembrane domain dimer structure resembles the BNIP3 transmembrane domain dimer structure shown in Figure 6.1.



**Figure 6.7 Alignment of the BNIP3 and GpA transmembrane domains interacting residues**

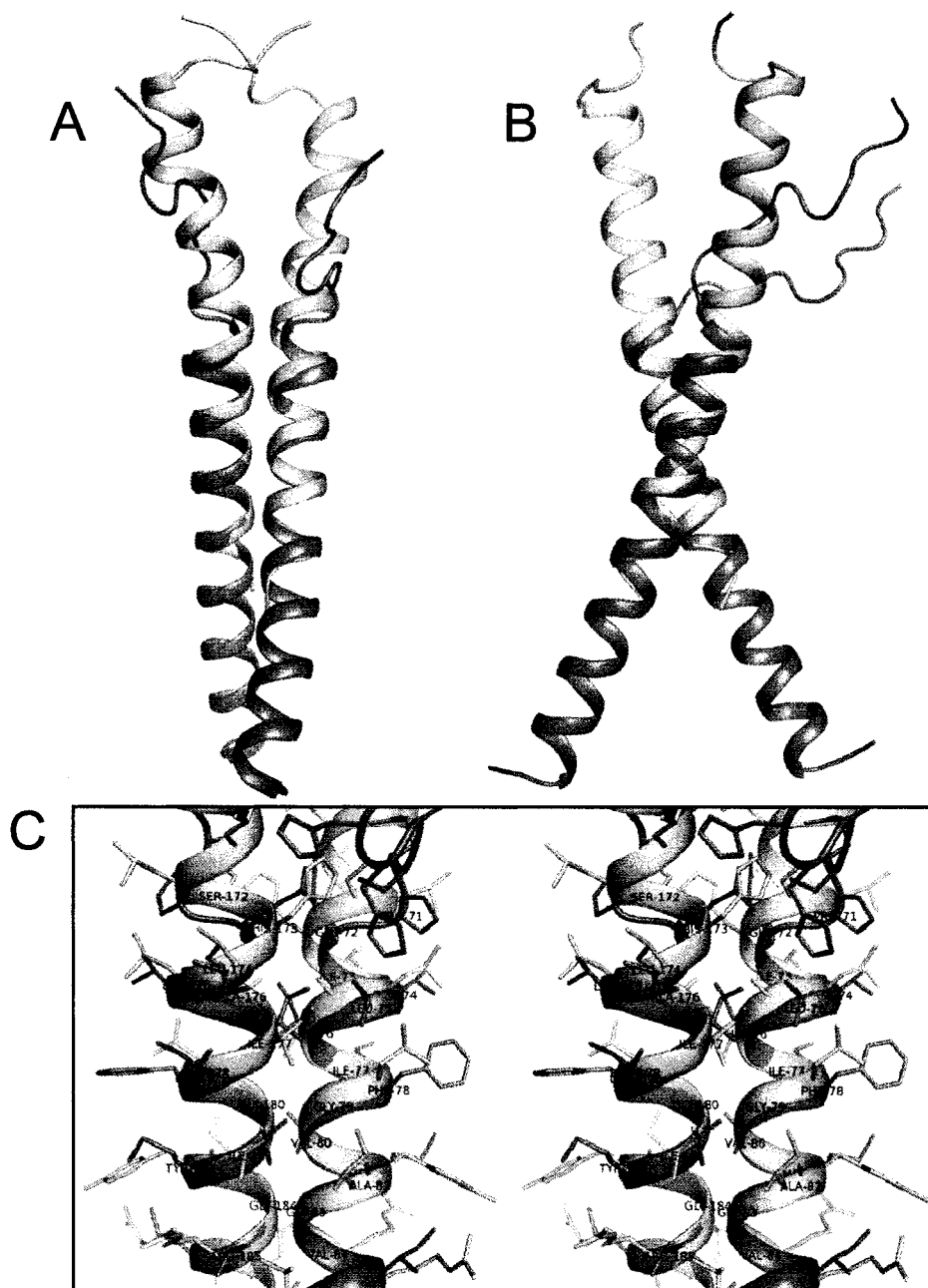
The sequences for BNIP3 and GpA transmembrane domains are aligned such that the GxxxG motif of GpA matches the BNIP3 A176xxxG180 (A and C) or G180xxxG184 (B). In panels A and B, positions that are found by mutagenesis to be important for dimerization in both sequences and that are within one rank when expressed on the GpA phenotype scale are boxed in both sequences with a dotted line (Lemmon et al. 1992, Sulistijo and MacKenzie 2006). Positions that are found to be important in one sequence but not in the other are boxed with a continuous line and emphasized with grey background when the phenotypes are not within one rank of the other phenotype scale. The mismatches on these two alignments suggest that BNIP3 and GpA transmembrane domain make rather different helix-helix contact.

Panel C shows that the GpA GxxxG to BNIP3 A176xxxG180 alignment matches better when structural data are included into consideration. However, the mismatch at BNIP3 Ile183 reaffirm the different contacts between the BNIP3 and GpA transmembrane domain dimers.



**Figure 6.8 Superposition of BNIP3 and GpA transmembrane domain structures when the BNIP3 A176xxxG180 motif is aligned with the GpA GxxxG motif**

Presented are overlay of the average structure of the BNIP3 TMD dimer (in cyan) and the structure of GpA TMD dimer (in purple), viewed from the side (A) and from the front (B). The amino terminus is on the top, and the carboxy terminus is at the bottom. Panel C shows stereo representation of the region covering residues 172-185 of the BNIP3 TMD. The residues of the BNIP3 TMD are labeled on the left helix, and the residues of the GpA TMD are labeled on the right helix. The superposition of the BNIP3 TMD (res 168-185) and the GpA TMD (res 71-88) has an RMS deviation value of  $1.2 \pm 0.2$  Å.



**Figure 6.9 Superposition of BNIP3 and GpA transmembrane domain structures when the BNIP3 G180xxxG184 motif is aligned with the GpA GxxxG motif**

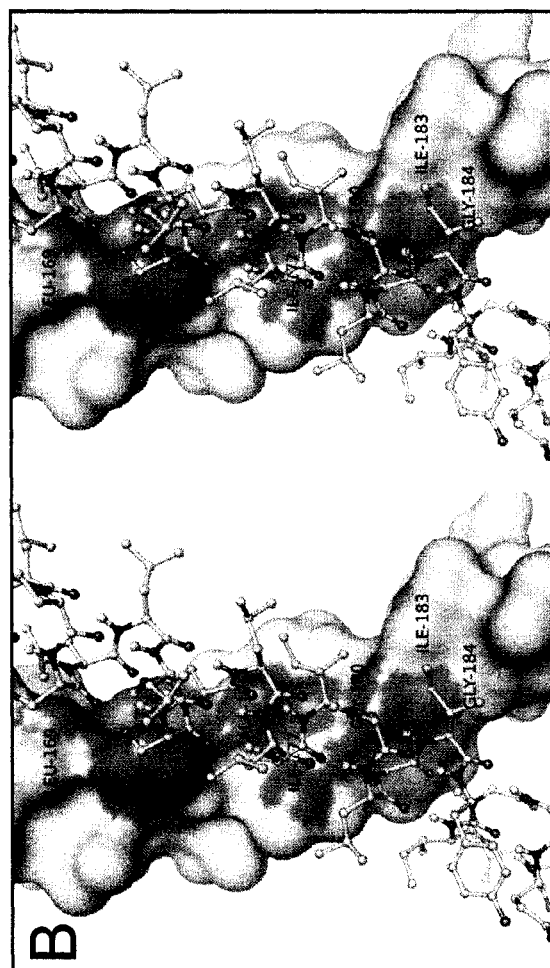
Presented are overlay of the average structure of the BNIP3 TMD dimer (in cyan) and the structure of GpA TMD dimer (in purple), viewed from the side (A) and from the front (B). The amino terminus is on the top, and the carboxy terminus is at the bottom. Panel C shows stereo representation of the region covering residues 172-185 of the BNIP3 TMD. The residues of the BNIP3 TMD are labeled on the left helix, and the residues of the GpA TMD are labeled on the right helix. The superposition of the BNIP3 TMD (res 172-185) and the GpA TMD (res 71-84) has an RMS deviation value of  $1.4 \pm 0.2$  Å.

A more detailed analysis of the BNIP3 TMD dimer and the GpA TMD dimer structures argues that the BNIP3 A176xxxG180 motif is more similar to the GpA GxxxG motif than the BNIP3 G180xxxG184 motif is. Superposition of the two TMD dimer structures for the first sequence alignment has an RMS deviation value of  $1.2 \pm 0.2$  Å (Figure 6.8), and superposition of the second sequence alignment has an RMS deviation of  $1.4 \pm 0.2$  Å (Figure 6.9). In addition, the structures indicate that helix-helix crossing is the closest between residues Ala176 and Gly180 in the BNIP3 TMD dimer and between residues Gly79 and Gly83 in the GpA TMD dimer, confirming the superiority of the BNIP3 A176xxxG180 motif and the GpA GxxxG motif sequence alignment.

The similarities between the BNIP3 TMD and the GpA TMD dimer structures, as well as the involvement of the GxxxG motif for dimer stability, imply that these helices form comparable associations despite helix-helix contact is closest at the carboxy terminus of the TMD for the first system and at the amino terminus of the TMD for the latter system. However, the lack of match for BNIP3 Ile183 in either sequence alignments and the requirement for inter-monomer hydrogen bonds between His173 and Ser172 suggest that BNIP3 TMD helix-helix association involves somewhat different contacts from the GpA TMD dimer. Therefore, even though mutagenesis and structural data indicate that the GxxxG motif is critical for dimerization in both systems, residues flanking this motif help define the specificity of the TMD dimer, thus preventing non-specific associations between various TMD helices containing the GxxxG motif.

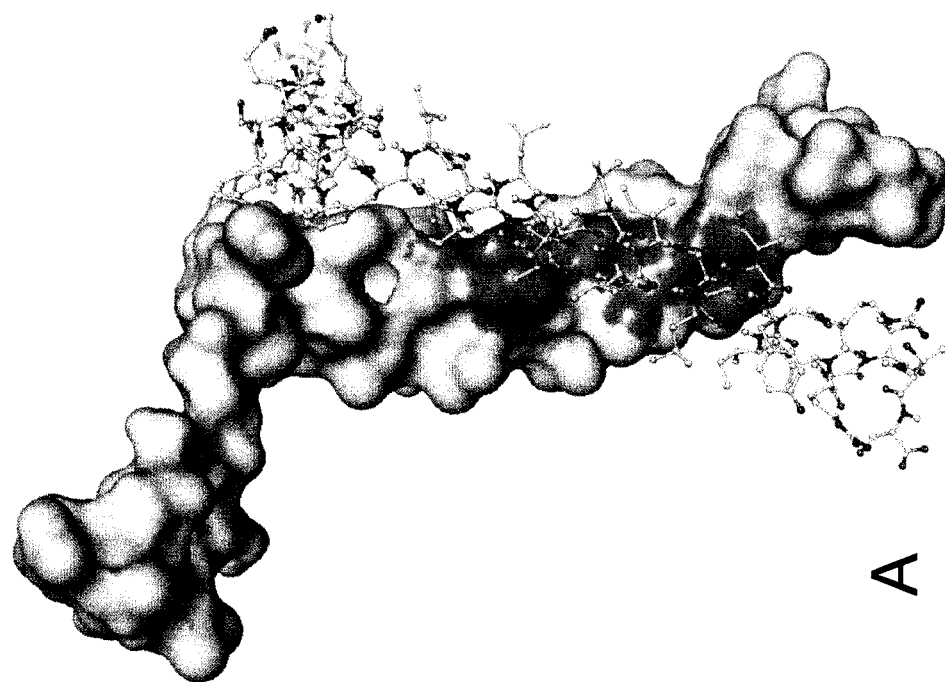
## 6.9 Comparison with a published BNIP3 transmembrane domain dimer structure

Recently, the Arseniev's group published a structure of the BNIP3 TMD dimer (Bocharov et al. 2007). The BNIP3 TMD peptide used for this structure determination is longer, covering residues 146-190, and the TMD peptide was prepared in the dimyristoylphosphatidyl-choline (DMPC)/dihexanoylphosphatidylcholine (DHPC) lipid bicelles instead of in detergent micelles. In general, the published structure looks very similar to the structure that I solved (compare Figure 6.10 to Figure 6.1), and the superposition of the two structures, between residues 168-185, which cover the well-defined region, has an RMS deviation value of  $1.0 \pm 0.2$  Å (Figure 6.11). The Arseniev's BNIP3 TMD dimer structure also shows that the TMD dimer has a right-handed crossing angle, and that the interaction face is composed of Phe165, Leu169, Ser172, His173, Ala176, Ile177, Gly180, Ile181, Ile183, Gly184, and Arg185, the same set of residues that I have identified from my structure. In addition, the Arseniev's group suggests that the BNIP3 TMD dimer is stabilized by backbone-backbone contacts, mediated by Ala176, Gly180, and Gly184, inter-monomer hydrogen bonds between His173 and Ser172, and hydrophobic side chain contacts, involving Ala176, Ile177, Ile181, and Ile183, similar to what I have suggested from the mutagenesis data and from the structure (Sulistijo and MacKenzie 2006; Bocharov et al. 2007). Detailed analysis of the Arseniev's BNIP3 TMD dimer structure indicates that the BNIP3 TMD helices pack with the ridges-to-grooves motif, and that the structure can also explain the dimerization phenotypes of the BNIP3 TMD mutants, as I have discussed in sections 6.2-6.6.

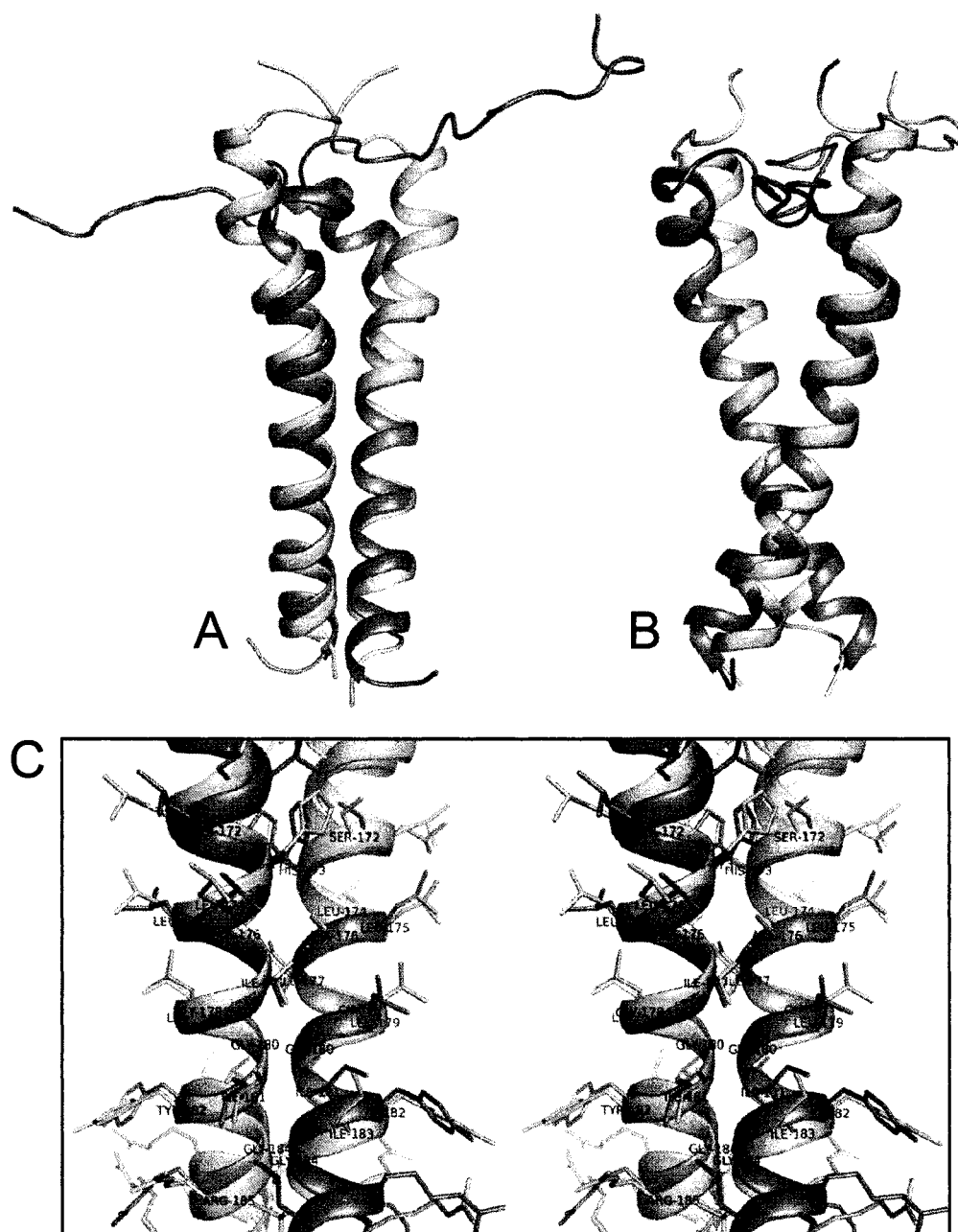


**Figure 6.10 The structure of the BNIP3 transmembrane domain dimer solved by the Arseniev's group**

Presented are (A) the structure of the BNIP3 transmembrane domain dimer solved by Arseniev and co-workers (Bocharov et al 2007, PDB accession number 2J5D) and (B) the stereo representation of the interacting region (res 169-184). Highlighted on the surface representation are atoms located within 4 Å of atoms on the opposite monomer, creating an interacting face for dimerization. The interfacial residues are labeled in panel B. Though the structure is similar to the structure that I calculated (see Figure 6.1), this BNIP3 transmembrane domain dimer structure shows a wider helix-helix crossing angle.







**Figure 6.11 Superposition of BNIP3 transmembrane domain dimer structure with the structure solved by the Arseniev's group**

Presented are overlay of the average structure of the BNIP3 transmembrane domain dimer (in cyan) that I solved using ARIA 1.2 and the structure solved by the Arseniev's group (in brown), viewed from the side (A) and from the front (B). The amino terminus is on the top, and the carboxy terminus is at the bottom. Panel C shows stereo representation of the region covering residues 172-185. The superposition of the two BNIP3 transmembrane domain dimer structures (res 168-185) has an RMS deviation value of  $1.0 \pm 0.2$  Å.

Detailed analysis of the Arseniev's structure cannot explain why substitutions of Ile183 to Met, Leu, and Val can disrupt the TMD dimer confirming that factors that cannot be identified from saturation mutagenesis and structural data are involved in the stability of BNIP3 TMD dimerization (see section 6.7).

A slight difference between the Arseniev's BNIP3 TMD dimer structure and my structure is in the angle between the TMD helices. The Arseniev's BNIP3 TMD dimer has wider crossing angle ( $\Omega = -45^\circ \pm 5^\circ$  in published literature (Bocharov et al. 2007),  $\Omega = -40^\circ \pm 5^\circ$  by my calculation for the region covering res 168-185), which means that the network of inter-monomer  $C^\alpha-H^\alpha \cdots O=C$  hydrogen bonds may involved a slightly different set of residues. Examination of the Arseniev's structure shows that Ala176, Ile177, Gly180, and Ile181 are implicated in the network of  $C^\alpha-H^\alpha \cdots O=C$  hydrogen bonds, but inter-monomer interaction between Ile181 and Gly184 may not be optimal ( $d_H > 4 \text{ \AA}$ ), and the distance between the  $H^\alpha$  of Arg185 and the carbonyl O of Gly184 ( $d_H > 6.5 \text{ \AA}$ ) does not satisfy the requirements for  $C^\alpha-H^\alpha \cdots O=C$  inter-monomer hydrogen bonds. Therefore, the Arseniev's BNIP3 TMD dimer structure indicates that each TMD monomer can provide three, or at most four,  $C^\alpha-H^\alpha \cdots O=C$  hydrogen bonds to the other monomer.

Despite some minor differences, having two independently solved BNIP3 TMD dimer structure is encouraging, particularly since both structures show similar helix-helix packing and implicate the same set of residues and the types of interactions to the stability of the TMD dimer. The fact that both the Arseniev's and my structures confirm structural inferences derived from saturation mutagenesis data suggests that the BNIP3

TMD dimer involves the same physical association in spite of being prepared in different analytical conditions (SDS micelles for mutagenesis analysis, DPC micelles and DMPC/DPPC bicelles for structure determination with solution NMR). These data also argue that mutagenesis analysis is a good approach to elucidate the basis of protein-protein interactions when structural data are not available.

## **7 Summary and future work**

### **7.1 Combining mutagenesis with structure determination reveals the basis for the specificity of membrane helix-helix associations**

This thesis has yielded a characterization of the sequence requirements for BNIP3 TMD dimerization and identification of interactions that participate in the stability of this association. Saturation mutagenesis offers a quick and simple approach for determining the sequence dependence of dimerization, and structure determination using solution NMR spectroscopy has provided a visual of the BNIP3 TMD dimer at atomic level, which allows interpretation of the mutagenesis data in the context of a three dimensional structure. Both saturation mutagenesis and structural data identify the same set of interacting residues and the types of interactions that are mediated by these residues, which include inter-monomer hydrogen bonds between His173 and Ser172 and tandem GxxxG motifs, mediated by Ala176, Gly180, and Gly184. Both approaches also agree that the tandem GxxxG motifs provide a context for the formation of close packing interactions between the two monomers and inter-monomer hydrogen bonds.

The good agreement between the mutagenesis and structural data may suggest that analysis using saturation mutagenesis would be sufficient to elucidate the sequence specificity for TMD helix-helix associations and to define the properties that drive and stabilize these interactions. Being simple and quick are also attractive features of this approach. However, there are interactions and physical characteristics that were only identified or explained when structural data became available, such as the inter-

monomer  $C^{\alpha}-H^{\alpha}\cdots O=C$  hydrogen bond between the TMD helix backbones.

Conversely, though the structure of the TMD dimer allows identification of interactions that are involved in the association, in the absence of mutagenesis data, the significance of certain interactions may be overlooked. For example, based on structural data alone inter-monomer hydrogen bonds between His173 and Ser172 may not be considered significant because replacement of these residues with other small amino acids can be accommodated sterically. Thus, one can argue that both saturation mutagenesis and structural data complement each other to better understand the chemical and physical requirements for membrane helix-helix associations.

## **7.2 Findings from saturation mutagenesis provide a framework for further studies of BNIP3 transmembrane domain dimerization**

### **7.2.1 Analysis of BNIP3 transmembrane domain dimerization in real membranes**

Saturation mutagenesis analysis and interpretation of the BNIP3 TMD dimer structure that are discussed in this thesis explain the chemical and physical properties that drive and stabilize the TMD self-association detergent micelles. Detergent micelles provide a convenient environment to study TMD helix-helix associations, but one can argue that detergent micelles may not accurately represent the conditions of real membranes and prompt a question whether the features and properties observed in detergent micelles would also occur in biological membranes. In the case of BNIP3 TMD dimerization, association in SDS micelles and in real membranes are likely

modulated by the same interactions because mutations in the TMD cause similar effects on dimerization in SDS micelles as observed with SDS-PAGE and in *E. coli* membranes as observed with TOXCAT assay (see section 3.2). This similarity also suggests that conclusions derived from saturation mutagenesis and SDS-PAGE analysis and information obtained from the TMD dimer structure in DPC micelles can be used to explain the basis for BNIP3 TMD dimerization in real membrane. However, verification of this hypothesis must wait for results from *in vivo* analysis using TOXCAT assay, which is currently ongoing in the MacKenzie lab.

### **7.2.2 Thermodynamic studies of BNIP3 transmembrane domain dimerization**

The reversibility of BNIP3 TMD dimerization in detergent micelles demonstrates that thermodynamic studies can be performed on this association. Serial two-fold dilutions of the wild type and His13Trp mutant of the SN-BNIP3TMD fusion protein enable calculation of apparent dissociation constants of these constructs with SDS-PAGE (section 3.4). However, because of uncertainty in determining the amount of protein in the dimeric and monomeric states from the protein band intensities, it is likely that the values obtained from this measurement are not highly accurate, though they should not deviate significantly from the correct values. Because of this limitation and because it is even more difficult to determine the amount of protein in the dimeric and monomeric states when the fusion protein constructs show smearing phenotypes (see Figure 3.7), SDS-PAGE is not the best approach to study the stoichiometry and thermodynamics of BNIP3 TMD dimerization. Therefore, other methods, such as analytical ultracentrifugation and fluorescence resonance energy transfer (FRET), are

needed to better analyze the thermodynamic of BNIP3 TMD helix-helix association (Fleming et al. 1997; Eisenhawer et al. 2001; Fleming and Engelman 2001; Gratkowski et al. 2001; Gratkowski et al. 2002; DeGrado et al. 2003; Lear et al. 2003; Lear et al. 2004), and TOXCAT assay can be employed to quantitatively determine the free energy for TMD helix-helix association in biological membranes (Duong et al. 2007).

### **7.2.3 Functional studies of BNIP3 transmembrane domain dimerization**

Even though overexpression of BNIP3 is associated with increased levels of apoptosis in yeast and cultured cells, little is known about the protein's function in regulating apoptosis. It has been reported that deletion of the carboxy terminal TMD cause a reduction of apoptosis levels, implying that the TMD is necessary for BNIP3 apoptosis function, and that interactions through the TMD may play a significant role in the regulation of this activity. However, this suggestion may have been obscured because the TMD is also necessary for subcellular localization to the mitochondrial membrane where the protein functions. Therefore, the use of BNIP3 mutants that have weakened TMD dimerization characteristics but do not have defects in mitochondrial targeting may provide better alternatives to elucidate the functional role for BNIP3 TMD association. Saturation mutagenesis and analysis using either SDS-PAGE or TOXCAT assay can identify such mutants, and thermodynamics studies enable finding of BNIP3 TMD mutants that show graded dimerization behavior, thus allowing more refined examination of the role of BNIP3 TMD association in apoptosis regulation.

## **7.3 Results from solution NMR spectroscopy call for additional analyses**

### **7.3.1 Determination of BNIP3 transmembrane domain peptide-phospholipid contacts**

Excellent peak resolution and signal-to-noise ratio in NMR spectra obtained with samples of BNIP3 TMD peptide that are prepared in the presence of phospholipids suggest that added phospholipids can modulate BNIP3 TMD dimer structure or conformational stability, perhaps through non-specific interactions (see section 5.3). With the data that I currently have, it is impossible to identify the best types of phospholipids for use in structural studies because these improvements appear to be independent of the types of the lipid head group or the length of the carbon tail (section 5.6). Addition of DPPC to NMR samples of BNIP3 TMD peptide to obtain good quality samples is reasonable because the mitochondrial membranes, which are the subcellular target of BNIP3, are enriched in phosphatidylcholine (PC). Yet, this does not rule out the possibility that other types of lipids may be able to enhance the quality of the BNIP3 TMD peptide samples better than PC.

My current data also prevent me from specifying which part of the lipids that interact with the BNIP3 TMD peptide and which part of the peptide that makes contacts with the lipids, because atoms in the lipid tail have similar chemical shift resonances to the methyl groups of aliphatic amino acids, and the NMR spectra that I have show poor resolution at the  $H^\alpha/C^\alpha$  region, which is also the region where resonances for atoms in the lipid head group can be detected. Information regarding peptide-lipid contacts can



be obtained by performing half-filtered NOESY-HSQC experiments on sample of [ $U$ - $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled BNIP3 TMD peptide prepared in the presence of unlabeled phospholipids. Though these data are not expected to affect the final structure of the BNIP3 TMD dimer, they could give insight into understanding how peptide-lipid interactions can influence the structure or conformational stability of the TMD association.

### **7.3.2 Study of BNIP3 transmembrane domain dimer dynamic**

The presence of two sets of chemical shifts for residues at the carboxy terminus of the BNIP3 TMD peptide (res 181-188) when DPPA is added to the sample suggests the possibility that this region is more mobile than other parts of the peptide in this condition, though the exchange rate between the two conformations is slow enough that they can be detected with NMR time scale (see section 5.6). To better understand the dynamics of BNIP3 TMD dimer, additional NMR experiments are needed, and analysis using molecular dynamics simulation will be necessary to complement the NMR data. Knowledge of the dynamics of BNIP3 TMD dimerization may also provide an explanation regarding the variations in the helix-helix crossing angle between my BNIP3 TMD dimer structure and that of the Arseniev's group.

## 7.4 Conclusion

The experimental evidence presented in this thesis demonstrates that the stability of BNIP3 TMD dimerization involves detailed packing interactions that can be identified by analysis using saturation mutagenesis and determining the structure of the TMD dimer using solution NMR spectroscopy. The good agreement between the mutagenesis and structural data suggest that either method may be sufficient to obtain valuable information about this type of association, but because the mutagenesis and structural data complement each other, both are needed to better elucidate the properties that modulate helix-helix interactions. The results obtain from this study provide a framework for further studies of BNIP3 TMD association to gain insight into understanding the role of this association in apoptosis regulation. In addition, this study also suggests that the methods used to characterize the BNIP3 TMD dimerization should prove applicable for the studies of other membrane helix-helix associations.

## Bibliography

- Adams, J. M. and S. Cory (1998). "The Bcl-2 protein family: arbiters of cell survival." Science **281**(5381): 1322-6.
- Antonsson, B. and J. C. Martinou (2000). "The Bcl-2 protein family." Exp Cell Res **256**(1): 50-7.
- Aouacheria, A., F. Brunet and M. Gouy (2005). "Phylogenomics of life-or-death switches in multicellular animals: Bcl-2, BH3-Only, and BNip families of apoptotic regulators." Mol Biol Evol **22**(12): 2395-416.
- Arkin, I. T., P. D. Adams, K. R. MacKenzie, M. A. Lemmon, A. T. Brunger and D. M. Engelman (1994). "Structural organization of the pentameric transmembrane alpha-helices of phospholamban, a cardiac ion channel." Embo J **13**(20): 4757-64.
- Blakey, D., A. Leech, G. H. Thomas, G. Coutts, K. Findlay and M. Merrick (2002). "Purification of the Escherichia coli ammonium transporter AmtB reveals a trimeric stoichiometry." Biochem J **364**(Pt 2): 527-35.
- Bocharov, E. V., Y. E. Pustovalova, K. V. Pavlov, P. E. Volynsky, M. V. Goncharuk, Y. S. Ermolyuk, D. V. Karpunin, A. A. Schulga, M. P. Kirpichnikov, R. G. Efremov, I. V. Maslennikov and A. S. Arseniev (2007). "Unique dimeric structure of BNip3 transmembrane domain suggests membrane permeabilization as a cell death trigger." J Biol Chem.
- Bowie, J. U. (1997). "Helix packing in membrane proteins." J Mol Biol **272**(5): 780-9.
- Boyd, J. M., S. Malstrom, T. Subramanian, L. K. Venkatesh, U. Schaeper, B. Elangovan, C. D'Sa-Eipper and G. Chinnadurai (1994). "Adenovirus E1B 19 kDa and Bcl-2 proteins interact with a common set of cellular proteins." Cell **79**(2): 341-51.
- Bruick, R. K. (2000). "Expression of the gene encoding the proapoptotic Nip3 protein is induced by hypoxia." Proc Natl Acad Sci U S A **97**(16): 9082-7.

- Brunger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54** ( Pt 5): 905-21.
- Chamberlain, A. K. and J. U. Bowie (2004). "Analysis of side-chain rotamers in transmembrane proteins." Biophys J **87**(5): 3460-9.
- Chen, G., J. Cizeau, C. Vande Velde, J. H. Park, G. Bozek, J. Bolton, L. Shi, D. Dubik and A. Greenberg (1999). "Nix and Nip3 form a subfamily of pro-apoptotic mitochondrial proteins." J Biol Chem **274**(1): 7-10.
- Chen, G., R. Ray, D. Dubik, L. Shi, J. Cizeau, R. C. Bleackley, S. Saxena, R. D. Gietz and A. H. Greenberg (1997). "The E1B 19K/Bcl-2-binding protein Nip3 is a dimeric mitochondrial protein that activates apoptosis." J Exp Med **186**(12): 1975-83.
- Choma, C., H. Gratkowski, J. D. Lear and W. F. DeGrado (2000). "Asparagine-mediated self-association of a model transmembrane helix." Nat Struct Biol **7**(2): 161-6.
- Chothia, C. (1984). "Principles that determine the structure of proteins." Annu Rev Biochem **53**: 537-72.
- Chothia, C., M. Levitt and D. Richardson (1981). "Helix to helix packing in proteins." J Mol Biol **145**(1): 215-50.
- Chou, J. J., H. Li, G. S. Salvesen, J. Yuan and G. Wagner (1999). "Solution structure of BID, an intracellular amplifier of apoptotic signaling." Cell **96**(5): 615-24.
- Cizeau, J., R. Ray, G. Chen, R. D. Gietz and A. H. Greenberg (2000). "The C. elegans orthologue ceBNIP3 interacts with CED-9 and CED-3 but kills through a BH3- and caspase-independent mechanism." Oncogene **19**(48): 5453-63.
- Cornea, R. L., L. R. Jones, J. M. Autry and D. D. Thomas (1997). "Mutation and phosphorylation change the oligomeric structure of phospholamban in lipid bilayers." Biochemistry **36**(10): 2960-7.
- Cory, S. and J. M. Adams (2002). "The Bcl2 family: regulators of the cellular life-or-death switch." Nat Rev Cancer **2**(9): 647-56.

- Cory, S., A. Strasser, T. Jacks, L. M. Corcoran, T. Metz, A. W. Harris and J. M. Adams (1994). "Enhanced cell survival and tumorigenesis." Cold Spring Harb Symp Quant Biol **59**: 365-75.
- Crick, F. H. C. (1953). "The packing of  $\alpha$ -helices: simple coiled-coils." Acta Crystallographica **6**: 689-697.
- Dawson, J. P., R. A. Melnyk, C. M. Deber and D. M. Engelman (2003). "Sequence context strongly modulates association of polar residues in transmembrane helices." J Mol Biol **331**(1): 255-62.
- DeGrado, W. F., H. Gratkowski and J. D. Lear (2003). "How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles." Protein Sci **12**(4): 647-65.
- Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax (1995). "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." J Biomol NMR **6**(3): 277-93.
- Derewenda, Z. S., L. Lee and U. Derewenda (1995). "The occurrence of C-H...O hydrogen bonds in proteins." J Mol Biol **252**(2): 248-62.
- Duong, M. T., T. M. Jaszewski, K. G. Fleming and K. R. Mackenzie (2007). "Changes in apparent free energy of helix-helix dimerization in a biological membrane due to point mutations." J Mol Biol **371**(2): 422-34.
- Eisenhawer, M., S. Cattarinussi, A. Kuhn and H. Vogel (2001). "Fluorescence resonance energy transfer shows a close helix-helix distance in the transmembrane M13 procoat protein." Biochemistry **40**(41): 12321-8.
- Engelman, D. M., T. A. Steitz and A. Goldman (1986). "Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins." Annu Rev Biophys Biophys Chem **15**: 321-53.
- Erkan, M., J. Kleeff, I. Esposito, T. Giese, K. Ketterer, M. W. Buchler, N. A. Giese and H. Friess (2005). "Loss of BNIP3 expression is a late event in pancreatic cancer contributing to chemoresistance and worsened prognosis." Oncogene **24**(27): 4421-32.

- Eskes, R., B. Antonsson, A. Osen-Sand, S. Montessuit, C. Richter, R. Sadoul, G. Mazzei, A. Nichols and J. C. Martinou (1998). "Bax-induced cytochrome C release from mitochondria is independent of the permeability transition pore but highly dependent on  $Mg^{2+}$  ions." J Cell Biol **143**(1): 217-24.
- Fesik, S. W. (2000). "Insights into programmed cell death through structural biology." Cell **103**(2): 273-82.
- Finucane, D. M., E. Bossy-Wetzel, N. J. Waterhouse, T. G. Cotter and D. R. Green (1999). "Bax-induced caspase activation and apoptosis via cytochrome c release from mitochondria is inhibitable by Bcl-xL." J Biol Chem **274**(4): 2225-33.
- Fleming, K. G., A. L. Ackerman and D. M. Engelman (1997). "The effect of point mutations on the free energy of transmembrane alpha-helix dimerization." J Mol Biol **272**(2): 266-75.
- Fleming, K. G. and D. M. Engelman (2001). "Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants." Proc Natl Acad Sci U S A **98**(25): 14340-4.
- Frazier, D. P., A. Wilson, R. M. Graham, J. W. Thompson, N. H. Bishopric and K. A. Webster (2006). "Acidosis regulates the stability, hydrophobicity, and activity of the BH3-only protein Bnip3." Antioxid Redox Signal **8**(9-10): 1625-34.
- Fujii, J., K. Maruyama, M. Tada and D. H. MacLennan (1989). "Expression and site-specific mutagenesis of phospholamban. Studies of residues involved in phosphorylation and pentamer formation." J Biol Chem **264**(22): 12950-5.
- Gratkowski, H., Q. H. Dai, A. J. Wand, W. F. DeGrado and J. D. Lear (2002). "Cooperativity and specificity of association of a designed transmembrane peptide." Biophys J **83**(3): 1613-9.
- Gratkowski, H., J. D. Lear and W. F. DeGrado (2001). "Polar side chains drive the association of model transmembrane peptides." Proc Natl Acad Sci U S A **98**(3): 880-5.
- Green, D. R. and G. I. Evan (2002). "A matter of life and death." Cancer Cell **1**(1): 19-30.
- Gross, A., J. M. McDonnell and S. J. Korsmeyer (1999). "BCL-2 family members and the mitochondria in apoptosis." Genes Dev **13**(15): 1899-911.

- Grzesiek, S., G. W. Vuister and A. Bax (1993). "A simple and sensitive experiment for measurement of JCC couplings between backbone carbonyl and methyl carbons in isotopically enriched proteins." J Biomol NMR **3**(4): 487-93.
- Guo, K., G. Searfoss, D. Krolikowski, M. Pagnoni, C. Franks, K. Clark, K. T. Yu, M. Jaye and Y. Ivashchenko (2001). "Hypoxia induces the expression of the pro-apoptotic gene BNIP3." Cell Death Differ **8**(4): 367-76.
- Guscetti, F., N. Nath and N. Denko (2005). "Functional characterization of human proapoptotic molecules in yeast *S. cerevisiae*." Faseb J **19**(3): 464-6.
- Hengartner, M. O. (2000). "The biochemistry of apoptosis." Nature **407**(6805): 770-6.
- Hessa, T., H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S. H. White and G. von Heijne (2005). "Recognition of transmembrane helices by the endoplasmic reticulum translocon." Nature **433**(7024): 377-81.
- Hinds, M. G., M. Lackmann, G. L. Skea, P. J. Harrison, D. C. Huang and C. L. Day (2003). "The structure of Bcl-w reveals a role for the C-terminal residues in modulating biological activity." Embo J **22**(7): 1497-507.
- Hirota, N., K. Mizuno and Y. Goto (1997). "Cooperative alpha-helix formation of beta-lactoglobulin and melittin induced by hexafluoroisopropanol." Protein Sci **6**(2): 416-21.
- Imagawa, T., T. Watanabe and T. Nakamura (1986). "Subunit structure and multiple phosphorylation sites of phospholamban." J Biochem (Tokyo) **99**(1): 41-53.
- Jacob, J., H. Duclouhier and D. S. Cafiso (1999). "The role of proline and glycine in determining the backbone flexibility of a channel-forming peptide." Biophys J **76**(3): 1367-76.
- Jones, T. A., J. Y. Zou, S. W. Cowan and M. Kjeldgaard (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta Crystallogr A **47** ( Pt 2): 110-9.
- Kay, L. E., P. Keifer and T. Saarinen (1992). "Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity " Journal of the American Chemical Society **114**(26): 10663-10665.

- Kelekar, A. and C. B. Thompson (1998). "Bcl-2-family proteins: the role of the BH3 domain in apoptosis." Trends Cell Biol **8**(8): 324-30.
- Kim, J. Y., J. J. Cho, J. Ha and J. H. Park (2002). "The carboxy terminal C-tail of BNip3 is crucial in induction of mitochondrial permeability transition in isolated mitochondria." Arch Biochem Biophys **398**(2): 147-52.
- Kim, S., A. K. Chamberlain and J. U. Bowie (2004). "Membrane channel structure of *Helicobacter pylori* vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels." Proc Natl Acad Sci U S A **101**(16): 5988-91.
- Kirkin, V., S. Joos and M. Zornig (2004). "The role of Bcl-2 family members in tumorigenesis." Biochim Biophys Acta **1644**(2-3): 229-49.
- Kleiger, G., R. Grothe, P. Mallick and D. Eisenberg (2002). "GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles." Biochemistry **41**(19): 5990-7.
- Korsmeyer, S. J., M. C. Wei, M. Saito, S. Weiler, K. J. Oh and P. H. Schlesinger (2000). "Pro-apoptotic cascade activates BID, which oligomerizes BAK or BAX into pores that result in the release of cytochrome c." Cell Death Differ **7**(12): 1166-73.
- Kubasiak, L. A., O. M. Hernandez, N. H. Bishopric and K. A. Webster (2002). "Hypoxia and acidosis activate cardiac myocyte death through the Bcl-2 family protein BNIP3." Proc Natl Acad Sci U S A **99**(20): 12825-30.
- Kubli, D. A., J. E. Ycaza and A. B. Gustafsson (2007). "Bnip3 mediates mitochondrial dysfunction and cell death through Bax and Bak." Biochem J **405**(3): 407-15.
- Kumar, Y., S. Tayyab and S. Muzammil (2004). "Molten-globule like partially folded states of human serum albumin induced by fluoro and alkyl alcohols at low pH." Arch Biochem Biophys **426**(1): 3-10.
- Kuwana, T., L. Bouchier-Hayes, J. E. Chipuk, C. Bonzon, B. A. Sullivan, D. R. Green and D. D. Newmeyer (2005). "BH3 domains of BH3-only proteins differentially regulate Bax-mediated mitochondrial membrane permeabilization both directly and indirectly." Mol Cell **17**(4): 525-35.
- Kuwana, T., M. R. Mackey, G. Perkins, M. H. Ellisman, M. Latterich, R. Schneider, D. R. Green and D. D. Newmeyer (2002). "Bid, Bax, and lipids cooperate to form



supramolecular openings in the outer mitochondrial membrane." Cell **111**(3): 331-42.

Kyte, J. and R. F. Doolittle (1982). "A simple method for displaying the hydropathic character of a protein." J Mol Biol **157**(1): 105-32.

Lamy, L., M. Ticchioni, A. K. Rouquette-Jazdanian, M. Samson, M. Deckert, A. H. Greenberg and A. Bernard (2003). "CD47 and the 19 kDa interacting protein-3 (BNIP3) in T cell apoptosis." J Biol Chem **278**(26): 23915-21.

Langosch, D. and J. Heringa (1998). "Interaction of transmembrane helices by a knobs-into-holes packing characteristic of soluble coiled coils." Proteins **31**(2): 150-9.

Laskowski, R., M. MacArthur, D. Moss and J. Thornton (1993). "PROCHECK: a program to check the stereochemical quality of protein structures." Journal of Applied Crystallography **26**: 283-291.

Lear, J. D., H. Gratkowski, L. Adamian, J. Liang and W. F. DeGrado (2003). "Position-dependence of stabilizing polar interactions of asparagine in transmembrane helical bundles." Biochemistry **42**(21): 6400-7.

Lear, J. D., A. L. Stouffer, H. Gratkowski, V. Nanda and W. F. Degrado (2004). "Association of a model transmembrane peptide containing gly in a heptad sequence motif." Biophys J **87**(5): 3421-9.

Lee, A. G. (2002). "Ca<sup>2+</sup> -ATPase structure in the E1 and E2 conformations: mechanism, helix-helix and helix-lipid interactions." Biochim Biophys Acta **1565**(2): 246-66.

Lee, H. and S. G. Paik (2006). "Regulation of BNIP3 in normal and cancer cells." Mol Cells **21**(1): 1-6.

Lemmon, M. A., J. M. Flanagan, J. F. Hunt, B. D. Adair, B. J. Bormann, C. E. Dempsey and D. M. Engelman (1992). "Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices." J Biol Chem **267**(11): 7683-9.

Lemmon, M. A., J. M. Flanagan, H. R. Treutlein, J. Zhang and D. M. Engelman (1992). "Sequence specificity in the dimerization of transmembrane alpha-helices." Biochemistry **31**(51): 12719-25.

- Leo, C., L. C. Horn and M. Hockel (2006). "Hypoxia and expression of the proapoptotic regulator BNIP3 in cervical cancer." Int J Gynecol Cancer **16**(3): 1314-20.
- Linge, J. P., M. Habeck, W. Rieping and M. Nilges (2003). "ARIA: automated NOE assignment and NMR structure calculation." Bioinformatics **19**(2): 315-6.
- Linge, J. P. and M. Nilges (1999). "Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation." J Biomol NMR **13**(1): 51-9.
- Linge, J. P., S. I. O'Donoghue and M. Nilges (2001). "Automated assignment of ambiguous nuclear overhauser effects with ARIA." Methods Enzymol **339**: 71-90.
- MacKenzie, K. R., J. H. Prestegard and D. M. Engelman (1996). "Leucine side-chain rotamers in a glycophorin A transmembrane peptide as revealed by three-bond carbon-carbon couplings and <sup>13</sup>C chemical shifts." J Biomol NMR **7**(3): 256-60.
- MacKenzie, K. R., J. H. Prestegard and D. M. Engelman (1997). "A transmembrane helix dimer: structure and implications." Science **276**(5309): 131-3.
- Manka, D., Z. Spicer and D. E. Millhorn (2005). "Bcl-2/adenovirus E1B 19 kDa interacting protein-3 knockdown enables growth of breast cancer metastases in the lung, liver, and bone." Cancer Res **65**(24): 11689-93.
- McClain, M. S., H. Iwamoto, P. Cao, A. D. Vinion-Dubiel, Y. Li, G. Szabo, Z. Shao and T. L. Cover (2003). "Essential role of a GXXXG motif for membrane channel formation by Helicobacter pylori vacuolating toxin." J Biol Chem **278**(14): 12101-8.
- McDonnell, J. M., D. Fushman, C. L. Milliman, S. J. Korsmeyer and D. Cowburn (1999). "Solution structure of the proapoptotic molecule BID: a structural basis for apoptotic agonists and antagonists." Cell **96**(5): 625-34.
- Meier, P., A. Finch and G. Evan (2000). "Apoptosis in development." Nature **407**(6805): 796-801.
- Morris, A. L., M. W. MacArthur, E. G. Hutchinson and J. M. Thornton (1992). "Stereochemical quality of protein structure coordinates." Proteins **12**(4): 345-64.

- Muchmore, S. W., M. Sattler, H. Liang, R. P. Meadows, J. E. Harlan, H. S. Yoon, D. Nettlesheim, B. S. Chang, C. B. Thompson, S. L. Wong, S. L. Ng and S. W. Fesik (1996). "X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death." Nature **381**(6580): 335-41.
- Nicholson, D. W. (2000). "From bench to clinic with apoptosis-based therapeutic agents." Nature **407**(6805): 810-6.
- Nilges, M. (1995). "Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities." J Mol Biol **245**(5): 645-60.
- Ohi, N., A. Tokunaga, H. Tsunoda, K. Nakano, K. Haraguchi, K. Oda, N. Motoyama and T. Nakajima (1999). "A novel adenovirus E1B19K-binding protein B5 inhibits apoptosis induced by Nip3 by forming a heterodimer through the C-terminal hydrophobic region." Cell Death Differ **6**(4): 314-25.
- Opferman, J. T. and S. J. Korsmeyer (2003). "Apoptosis in the development and maintenance of the immune system." Nat Immunol **4**(5): 410-5.
- Partridge, A. W., R. A. Melnyk and C. M. Deber (2002). "Polar residues in membrane domains of proteins: molecular basis for helix-helix association in a mutant CFTR transmembrane segment." Biochemistry **41**(11): 3647-53.
- Pastorino, J. G., S. T. Chen, M. Tafani, J. W. Snyder and J. L. Farber (1998). "The overexpression of Bax produces cell death upon induction of the mitochondrial permeability transition. PG - 7770-5." J Biol Chem **273**(13).
- Ponder, J. W. and F. M. Richards (1987). "Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes." J Mol Biol **193**(4): 775-91.
- Ray, R., G. Chen, C. Vande Velde, J. Cizeau, J. H. Park, J. C. Reed, R. D. Gietz and A. H. Greenberg (2000). "BNIP3 heterodimerizes with Bcl-2/Bcl-X(L) and induces cell death independent of a Bcl-2 homology 3 (BH3) domain at both mitochondrial and nonmitochondrial sites." J Biol Chem **275**(2): 1439-48.
- Reed, J. C. (1997). "Double identity for proteins of the Bcl-2 family." Nature **387**(6635): 773-6.

- Reed, J. C. (2001). "Apoptosis-regulating proteins as targets for drug discovery." Trends Mol Med **7**(7): 314-9.
- Reed, J. C. (2002). "Apoptosis-based therapies." Nat Rev Drug Discov **1**(2): 111-21.
- Reed, J. C., K. S. Doctor and A. Godzik (2004). "The domains of apoptosis: a genomics perspective." Sci STKE **2004**(239): re9.
- Regula, K. M., K. Ens and L. A. Kirshenbaum (2002). "Inducible expression of BNIP3 provokes mitochondrial defects and hypoxia-mediated cell death of ventricular myocytes." Circ Res **91**(3): 226-31.
- Richmond, T. J. and F. M. Richards (1978). "Packing of alpha-helices: geometrical constraints and contact areas." J Mol Biol **119**(4): 537-55.
- Russ, W. P. and D. M. Engelman (1999). "TOXCAT: a measure of transmembrane helix association in a biological membrane." Proc Natl Acad Sci U S A **96**(3): 863-8.
- Russ, W. P. and D. M. Engelman (2000). "The GxxxG motif: a framework for transmembrane helix-helix association." J Mol Biol **296**(3): 911-9.
- Sattler, M., H. Liang, D. Nettlesheim, R. P. Meadows, J. E. Harlan, M. Eberstadt, H. S. Yoon, S. B. Shuker, B. S. Chang, A. J. Minn, C. B. Thompson and S. W. Fesik (1997). "Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis." Science **275**(5302): 983-6.
- Scheiner, S., T. Kar and Y. Gu (2001). "Strength of the Calpha H..O hydrogen bond of amino acid residues." J Biol Chem **276**(13): 9832-7.
- Schendel, S. L., M. Montal and J. C. Reed (1998). "Bcl-2 family proteins as ion-channels." Cell Death Differ **5**(5): 372-80.
- Senes, A., D. E. Engel and W. F. DeGrado (2004). "Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs." Curr Opin Struct Biol **14**(4): 465-79.
- Senes, A., M. Gerstein and D. M. Engelman (2000). "Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in

association with beta-branched residues at neighboring positions." J Mol Biol **296**(3): 921-36.

Senes, A., I. Ubarretxena-Belandia and D. M. Engelman (2001). "The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions." Proc Natl Acad Sci U S A **98**(16): 9056-61.

Sowter, H. M., P. J. Ratcliffe, P. Watson, A. H. Greenberg and A. L. Harris (2001). "HIF-1-dependent regulation of hypoxic induction of the cell death factors BNIP3 and NIX in human tumors." Cancer Res **61**(18): 6669-73.

Strasser, A., D. C. Huang and D. L. Vaux (1997). "The role of the bcl-2/ced-9 gene family in cancer and general implications of defects in cell death control for tumourigenesis and resistance to chemotherapy.PG - F151-78." Biochim Biophys Acta **1333**(2).

Sulistijo, E. S., T. M. Jaszewski and K. R. MacKenzie (2003). "Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent." J Biol Chem **278**(51): 51950-6.

Sulistijo, E. S. and K. R. MacKenzie (2006). "Sequence dependence of BNIP3 transmembrane domain dimerization implicates side-chain hydrogen bonding and a tandem GxxxG motif in specific helix-helix interactions." J Mol Biol **364**(5): 974-90.

Suzuki, M., R. J. Youle and N. Tjandra (2000). "Structure of Bax: coregulation of dimer formation and intracellular localization." Cell **103**(4): 645-54.

Thompson, C. B. (1995). "Apoptosis in the pathogenesis and treatment of disease." Science **267**(5203): 1456-62.

van den Brink-van der Laan, E., V. Chupin, J. A. Killian and B. de Kruijff (2004). "Small alcohols destabilize the KcsA tetramer via their effect on the membrane lateral pressure." Biochemistry **43**(20): 5937-42.

Vande Velde, C., J. Cizeau, D. Dubik, J. Alimonti, T. Brown, S. Israels, R. Hakem and A. H. Greenberg (2000). "BNIP3 and genetic control of necrosis-like cell death through the mitochondrial permeability transition pore." Mol Cell Biol **20**(15): 5454-68.

- Vargas, R., J. Garza, D. A. Dixon and B. P. Hay (2000). "How Strong Is the C-H $\cdots$ O=C Hydrogen Bond?" Journal of the American Chemical Society **122**(19): 4750 - 4755.
- Vaux, D. L. and S. J. Korsmeyer (1999). "Cell death in development." Cell **96**(2): 245-54.
- Venyaminov, S. and N. N. Kalnin (1990). "Quantitative IR spectrophotometry of peptide compounds in water (H<sub>2</sub>O) solutions. I. Spectral parameters of amino acid residue absorption bands." Biopolymers **30**(13-14): 1243-57.
- Voulhoux, R., M. P. Bos, J. Geurtsen, M. Mols and J. Tommassen (2003). "Role of a highly conserved bacterial protein in outer membrane protein assembly." Science **299**(5604): 262-5.
- Vuister, G. W. and A. Bax (1993). "Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond J(HN $\alpha$ H) Coupling Constants in <sup>15</sup>N-Enriched Proteins." J. Am. Chem. Soc **115**(17): 7772-7777.
- Vuister, G. W., A. C. Wang and A. Bax (1993). "Measurement of Three-Bond Nitrogen-Carbon J Couplings in Proteins Uniformly Enriched in <sup>15</sup>N and <sup>13</sup>C." Journal of the American Chemical Society **115**(12): 5334-5335.
- Walters, R. F. and W. F. DeGrado (2006). "Helix-packing motifs in membrane proteins." Proc Natl Acad Sci U S A **103**(37): 13658-63.
- Walther, D., F. Eisenhaber and P. Argos (1996). "Principles of helix-helix packing in proteins: the helical lattice superposition model." J Mol Biol **255**(3): 536-53.
- Wan, J., D. Martinvalet, X. Ji, C. Lois, S. M. Kaech, U. H. Von Andrian, J. Lieberman, R. Ahmed and N. Manjunath (2003). "The Bcl-2 family pro-apoptotic molecule, BNIP3 regulates activation-induced cell death of effector cytotoxic T lymphocytes." Immunology **110**(1): 10-7.
- Watanabe, Y. (2002). "Effect of various mild surfactants on the reassembly of an oligomeric integral membrane protein OmpF porin." J Protein Chem **21**(3): 169-75.
- Webster, K. A., R. M. Graham and N. H. Bishopric (2005). "BNip3 and signal-specific programmed death in the heart." J Mol Cell Cardiol **38**(1): 35-45.

- Wimley, W. C., T. P. Creamer and S. H. White (1996). "Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides." Biochemistry **35**(16): 5109-24.
- Wimley, W. C. and S. H. White (1996). "Experimentally determined hydrophobicity scale for proteins at membrane interfaces." Nat Struct Biol **3**(10): 842-8.
- Wishart, D. S. and B. D. Sykes (1994). "The  $^{13}\text{C}$  chemical-shift index: a simple method for the identification of protein secondary structure using  $^{13}\text{C}$  chemical-shift data." J Biomol NMR **4**(2): 171-80.
- Wishart, D. S. and B. D. Sykes (1994). "Chemical shifts as a tool for structure determination." Methods Enzymol **239**: 363-92.
- Wishart, D. S., B. D. Sykes and F. M. Richards (1991). "Relationship between nuclear magnetic resonance chemical shift and protein secondary structure." J Mol Biol **222**(2): 311-33.
- Wishart, D. S., B. D. Sykes and F. M. Richards (1992). "The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy." Biochemistry **31**(6): 1647-51.
- Yamazaki, T., J. D. Forman-Kay and L. E. Kay (1993). "Two-dimensional NMR experiments for correlating  $^{13}\text{C}\beta$  and  $^1\text{H}\delta/\epsilon$  chemical shifts of aromatic residues in  $^{13}\text{C}$ -labeled proteins via scalar couplings." Journal of the American Chemical Society **115**(23): 11054-11055.
- Yasuda, M., P. Theodorakis, T. Subramanian and G. Chinnadurai (1998). "Adenovirus E1B-19K/BCL-2 interacting protein BNIP3 contains a BH3 domain and a mitochondrial targeting sequence." J Biol Chem **273**(20): 12415-21.
- Yuan, J. and B. A. Yankner (2000). "Apoptosis in the nervous system." Nature **407**(6805): 802-9.
- Zhou, F. X., M. J. Cocco, W. P. Russ, A. T. Brunger and D. M. Engelman (2000). "Interhelical hydrogen bonding drives strong interactions in membrane proteins." Nat Struct Biol **7**(2): 154-60.

Zhou, F. X., H. J. Merianos, A. T. Brunger and D. M. Engelman (2001). "Polar residues drive association of polyleucine transmembrane helices." Proc Natl Acad Sci U S A **98**(5): 2250-5.

Zwahlen, C., P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat and L. E. Kay (1997). "Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage lambda N-Peptide/boxB RNA Complex " Journal of the American Chemical Society **119**(29): 6711 -6721.



## **Appendix A: Oligonucleotide primers**

	Page
Sequence of mutagenic primers	185
Sequence of non-mutagenic primers	188

## Mutagenic primers

Primer names	Primer sequence (5' -> 3')	Target positions	Complementary primers
ES-NP-1A	CC CTG CTG TTA AGC <u>GNG</u> CTG CTG GCG ATC GGC	His173	ES-NP-1B
ES-NP-1B	GCC CAT CGC CAG CAG <u>CNC</u> GCT TAA CAG CAG GG	His173	ES-NP-1A
ES-NP-2A	CC CTG CTG TTA AGC <u>TKK</u> CTG CTG GCG ATC GGC	His173	ES-NP-2B
ES-NP-2B	GCC CAT CGC CAG CAG <u>MMA</u> GCT TAA CAG CAG GG	His173	ES-NP-2A
ES-NP-3	CC CTG CTG TTA AGC <u>MAD</u> CTG CTG GCG ATC GGC CTT	His173	ES-NP-C1
ES-NP-4	CC CTG CTG TTA <u>NBT</u> CAC CTG CTG GCG ATC GGC CTT	Ser172	ES-NP-C1
ES-NP-5	G CTG TTA AGC CAC CTG <u>NBT</u> GCG ATC GGC CTT GGA A	Leu175	ES-NP-C2
ES-NP-6	GTT TTC CTG CCG TCC <u>CTC CTC DBT</u> AGC CAC CTG CTG GCG ATC	Leu171	ES-NP-C1
ES-NP-7	GTT TTC CTG CCG TCC <u>DBT CTC CTC</u> AGC CAC CTG CTG GCG ATC	Leu169	ES-NP-C1
ES-NP-9	GCG ATC GGC CTT <u>NYT</u> ATC TAC ATC GGT CGC	Gly180	ES-NP-C3
ES-NP-10	CTG CTG TTA AGC CAC CTC CTC NBC ATC GGC CTT GGA ATC TAC	Ala176	ES-NP-C4
ES-NP-11	CAC CTG CTG GCG <u>NBT</u> GGC CTT GGA ATC TAC AT	Ile177	ES-NP-C4
ES-NP-12	GTT TTC CTG CCG <u>NBC CTC CTC</u> TTA AGC CAC CTG CTG	Ser168	ES-NP-C5
ES-NP-13	AGC CAC CTG CTG <u>NBC</u> ATC GGC CTT GGA ATC TAC	Ala176	ES-NP-C4
ES-NP-14	CTG TTA AGC GAG <u>CTC CTC</u> GCG ATC GGC CTT GGA ATC TAC		ES-NP-C4
ES-NP-15	CTG CCG TCC CTG <u>NBT</u> TTA AGC CAC CTG CTG	Leu170	ES-NP-C5
ES-NP-16	CAC CTG CTG GCG ATA <u>WKG</u> CTT GGA ATC TAC A	Gly178	ES-NP-C6
ES-NP-17	CAC CTG CTG GCG ATA <u>DYC</u> CTT GGA ATC TAC A	Gly178	ES-NP-C6
ES-NP-18	CTG TTA AGC CAC <u>NBT CTC</u> GCG ATC GGC CTT GGA	Leu174	ES-NP-C7
ES-NP-19	CTG CTG GCG ATC <u>GGA NBC</u> GGA ATC TAC ATC GGT	Leu179	ES-NP-C6
ES-NP-20	GAA TTC CTG AAA <u>GTG</u> TTC <u>NBC</u> CCG TCC CTG CTG TTA	Leu166	ES-NP-C8
ES-NP-21	GC CTT <u>GGG</u> ATC TAC ATC <u>GGG</u> CGC CGG CTG ACC ACG TC		ES-NP-C9
ES-NP-22	CTT GGG ATC TAC ATC <u>WKG</u> CGA CGG <u>ATG</u> ACC ACG	Gly184	ES-NP-C9
ES-NP-23	CTT GGG ATC TAC ATC <u>DYT</u> CGA CGG <u>ATG</u> ACC ACG	Gly184	ES-NP-C9
ES-NP-24	GAA TTC CTG AAA <u>WKK</u> TTT CTC CCG TCC CTG CTG TTA	Val164	ES-NP-C8
ES-NP-25	GAA TTC CTG AAA <u>VST</u> TTT CTC CCG TCC CTG CTG T	Val164	ES-NP-C8
ES-NP-26	GAA TTC CTG AAA <u>GTT ABK</u> CTC CCG TCC CTG CTG TTA	Phe165	ES-NP-C8
ES-NP-27	GAA TTC CTG AAA <u>GTT SBA</u> CTC CCG TCC CTG CTG TTA	Phe165	ES-NP-C8
ES-NP-28	CTG CTG GCG ATC <u>GGA WBG</u> GGA ATC TAC ATC GGT	Leu179	ES-NP-C6
ES-NP-29	CTG CTG GCG ATC <u>GGA GKG</u> GGA ATC TAC ATC GGT	Leu179	ES-NP-C6
ES-NP-30	GAA TTC CTG AAA <u>HTT</u> TTT CTC CCG TCC CTG CTG TA	Val164	ES-NP-C8
ES-NP-31	CTT GGA ATC TAC <u>CCC</u> GGG CGC CGG CTG ACC ACG	Ile183	ES-NP-C9
ES-NP-32	CTT GGA ATC TAC <u>WBK</u> GGA CGC CGG CTG ACC ACG	Ile183	ES-NP-C9
ES-NP-33	CTT GGA ATC TAC <u>GBT</u> GGA CGC CGG CTG ACC ACG	Ile183	ES-NP-C9
ES-NP-34	GC CTT <u>GGT DBT</u> TAC ATC GGG CGC CGG CTG ACC ACG TCG	Ile181	ES-NP-C9
ES-NP-35	GC CTT <u>GGT WBG</u> TAC ATC GGG CGC CGG CTG ACC ACG TCG	Ile181	ES-NP-C9

Primer names	Primer sequence (5' -> 3')	Target positions	Complementary primers
ES-NP-36	CTT GGT ATC <u>NBI</u> ATC GGG CGC CGG CTG ACC ACG TCG	Tyr182	ES-NP-C9
ES-NP-37	CTG AAA GTT TTC <u>CTC</u> <u>CTC</u> TCC CTG CTG T	Pro167	ES-NP-C10
ES-NP-38	CTG AAA GTT TTC CTC <u>WBK</u> TCC CTG CTG TT	Pro167	ES-NP-C10
ES-NP-39	CTG AAA GTT TTC CTC <u>GBT</u> TCC CTG CTG TT	Pro167	ES-NP-C10
ES-NP-40	C CGG CTG ACC <u>TAA</u> TCG ACC TCC ACC TTC TA	Thr189	ES-NP-C11
ES-NP-41	C CGG CTG ACC ACG TCG <u>TAA</u> TCC ACC TTC TA	Thr191	ES-NP-C11
ES-NP-42	TTC CTG CCG TCC <u>SCT</u> <u>CTC</u> <u>CTC</u> AGC CAC CTG CTG GCG AT	Leu169	ES-NP-C1
ES-NP-43	CTG CCG TCC <u>CTC</u> <u>CTC</u> <u>SCA</u> AGC CAC CTG CTG GCG AT	Leu171	ES-NP-C1
ES-NP-44	CTG TTA AGC CAC <u>SCA</u> <u>CTG</u> GCG ATC GGC CTT GGA AT	Leu174	ES-NP-C7
ES-NP-45	CAC CTG CTG GCG <u>SCT</u> GGC CTT GGA ATC TAC AT	Ile177	ES-NP-C4
ES-NP-46	GCG ATC GGA CTT <u>SCT</u> ATC TAC ATC GGT CGC	Gly180	ES-NP-C3
ES-NP-47A	AC ACT TCC GTT ATG <u>AAA</u> GGT GGT ATC TTC TC	Lys152	ES-NP-47B
ES-NP-47B	GA GAA GAT ACC ACC <u>TTT</u> CAT AAC GGA AGT GT	Lys152	ES-NP-47A
ES-NP-48	CC CTG CTG TTA AGC <u>RTR</u> CTG CTG GCG ATC GGC CTT GGA	His173	ES-NP-C1
ES-NP-49	AGC CAC CTG CTG <u>ATS</u> ATC GGC CTT GGA ATC TAC	Ala176	ES-NP-C4
ES-NP-50	CAC CTG CTG GCG <u>MTG</u> GGC CTT GGA ATC TAC A	Ile177	ES-NP-C4
ES-NP-51	CC CTG CTG TTA <u>MTR</u> CAC CTG CTG GCG ATC GGC CTT	Ser172	ES-NP-C1
ES-NP-52	GTT TTC CTG CCG <u>MTG</u> <u>CTC</u> <u>CTC</u> TTA AGC CAC CTG CTG	Ser168	ES-NP-C5
ES-NP-53	CC CTG CTG TTA AGT <u>HCT</u> CTG CTG GCG ATC GGC CTT GG	His173	ES-NP-C1
ES-NP-54	G CTG TTA AGT CAC CTG <u>ATR</u> GCG ATC GGC CTT GGA A	Leu175	ES-NP-C2
ES-NP-55	AGC CAC CTG CTG <u>MCA</u> <u>ATA</u> GGC CTT GGA ATC TAC AT	Ala176	ES-NP-C4
ES-NP-56	CAC CTG CTG GCG <u>ATA</u> <u>TKT</u> CTT GGA ATC TAC ATC GGT C	Gly178	ES-NP-C6
ES-NP-57	CAC CTG CTG GCG <u>ATA</u> <u>HCT</u> CTT GGA ATC TAC ATC GGT C	Gly178	ES-NP-C6
ES-NP-58	GC CTT GGT <u>SYA</u> TAC ATC GGG CGC CGG CTG ACC ACG TCG ACC TC	Ile181	ES-NP-C9
ES-NP-59	CTT GGT ATC <u>WTG</u> ATC GGG CGC CGG CTG ACC ACG TCG ACC TC	Tyr182	ES-NP-C9
ES-NP-60	CTT GGT ATC <u>SCA</u> ATC GGG CGC CGG CTG ACC ACG TCG ACC TC	Tyr182	ES-NP-C9
ES-NP-61	CTT GGA ATC TAC <u>WTG</u> <u>GGA</u> CGC CGG CTG ACC ACG TCG ACC TC	Ile183	ES-NP-C9
ES-NP-62	CTT GGG ATC TAC ATC <u>YCT</u> <u>CGA</u> CGG CTG ACC ACG TCG ACC TC	Gly184	ES-NP-C9
ES-NP-63	CTG AAA GTT TTC <u>CTG</u> <u>ATS</u> TCC CTG CTG TTA AGC CA	Pro167	ES-NP-C10
ES-NP-64	CTG CCG TCC <u>CTG</u> <u>RTG</u> TTA AGC CAC CTG C	Leu170	ES-NP-C5
ES-NP-65	GCG ATC GGA CTT <u>WTG</u> ATC TAC ATC GGT CGC	Gly180	ES-NP-C3
ES-NP-66	GTT TTC CTG CCG TCC <u>TTT</u> CTG TTA AGC CAC CTG C	Leu169	ES-NP-C5
ES-NP-67	CTG CCG TCC <u>CTG</u> <u>CCA</u> TTA AGC CAC CTG C	Leu170	ES-NP-C5
ES-NP-68	CC CTG CTG TTA <u>CCA</u> CAC CTG CTG GCG ATC GGC CTT	Ser172	ES-NP-C1
ES-NP-69	CC CTG CTG TTA AGT <u>CCT</u> CTG CTG GCG ATC GGC CTT GG	His173	ES-NP-C1
ES-NP-70	CAC CTG CTG GCG <u>ATA</u> <u>CCT</u> CTT GGA ATC TAC ATC GGT C	Gly178	ES-NP-C6

Primer names	Primer sequence (5' -> 3')	Target positions	Complementary primers
ES-NP-71	ATC TAC ATC GGT CGC <u>TAA</u> CTG ACC ACG TCG ACC TC	Arg186	ES-NP-C9
ES-NP-72	CTT GGG ATC TAC ATC <u>CCT</u> CGA CGG CTG ACC ACG TCG ACC TC	Gly184	ES-NP-C9
ES-NP-73	ATC TAC ATC GGT <u>VYT</u> CGG CTG ACC ACG TCG ACC TCC A	Arg185	ES-NP-C9
ES-NP-74	ATC TAC ATC GGT <u>VAW</u> CGG CTG ACC ACG TCG ACC TCC A	Arg185	ES-NP-C9
ES-NP-75	ATC TAC ATC GGT CGC <u>VYT</u> CTG ACC ACG TCG ACC TCC A	Arg186	ES-NP-C9
ES-NP-76	ATC TAC ATC GGT CGC <u>VAW</u> CTG ACC ACG TCG ACC TCC A	Arg186	ES-NP-C9
ES-NP-77	ATC GGT CGC CGG <u>VYT</u> ACC ACG TCG ACC TCC ACC TTC T	Leu187	ES-NP-C12
ES-NP-78	ATC GGT CGC CGG CTG <u>VYT</u> ACG TCG ACC TCC ACC TTC T	Thr188	ES-NP-C12
ES-NP-79	CC CTG CTG TTA <u>CAT</u> CAC CTG CTG GCG ATC GGC CTT	Ser172	ES-NP-C1
ES-NP-80	ATC GGT CGC CGG <u>CCT</u> ACC ACG TCG ACC TCC ACC TTC T	Leu187	ES-NP-C12

## Non-mutagenic primers

Primer names	Primer sequence (5' -> 3')
ES-NP-C1	GAT TCC AAG GCC GAT CGC CAG
ES-NP-C2	GTA GAT TCC AAG GCC GAT CGC
ES-NP-C3	CGA CGT GGT CAG CCG GCG CCA GAT G
ES-NP-C4	AG CCG GCG ACC GAT GTA GAT TCC AAG
ES-NP-C5	GA TCG CCA GCA GGT GGC TTA
ES-NP-C6	GT CAG CCG GCG ACC GAT GTA GAT TC
ES-NP-C7	ACC GAT GTA GAT TCC AAG GCC GAT
ES-NP-C8	CAG GTG GCT TTA CAG CAG GGA CG
ES-NP-C9	GAA GGT GGA GGT CGA CGT GGT CAG
ES-NP-C10	CAG CAG GTG GCT TAA CAG CAG GG
ES-NP-C11	CA GCC GGA TCC TCA TTA GAA GGT G
ES-NP-C12	TCC TCA TTA GAA GGT GGA GGT CG

## **Appendix B: Glycerol- and urea-SDS-PAGE**

	Page
SDS-PAGE	190
Urea-SDS-PAGE	191
Glycerol-SDS-PAGE	192

**5 ml 12% separating gel:**

	<u>no urea gel</u>
ddH <sub>2</sub> O	1.6 mL
30% acrylamide:bis (37.5:1)	2 mL
1.5 M Tris-HCl pH 8.8	1.3 mL
10% (w/v) SDS	50 µL
10% (w/v) ammonium persulfate	50 µL
TEMED	2 µL

**2 ml 5% stacking gel (for stacking two gels):**

ddH <sub>2</sub> O	1.4 mL
30% acrylamide:bis (37.5:1)	330 µL
1.0 M Tris-HCl pH 6.8	250 µL
10% (w/v) SDS	20 µL
10% (w/v) ammonium persulfate	20 µL
TEMED	2 µL

**5 ml 12% separating gel:**

	<u>urea gel, 8 M</u>
ddH <sub>2</sub> O	to vol
30% acrylamide:bis (37.5:1)	2 mL
1.5 M Tris-HCl pH 8.8	1.3 mL
10% (w/v) SDS	50 µL
urea	2.4 g
10% (w/v) ammonium persulfate	50 µL
TEMED	2 µL

**2 ml 5% stacking gel (for stacking two gels):**

ddH <sub>2</sub> O	1.4 mL
30% acrylamide:bis (37.5:1)	330 µL
1.0 M Tris-HCl pH 6.8	250 µL
10% (w/v) SDS	20 µL
10% (w/v) ammonium persulfate	20 µL
TEMED	2 µL

**5 ml 1x urea loading buffer (8 M urea):**

urea	2.4 g (8 M)
10% (w/v) SDS	1 ml (2%)
1 M Tris-HCl pH 6.8	250 µL (50 mM)
bromophenol blue	5 mg (0.1%)
ddH <sub>2</sub> O	to 5 ml



**5 ml 12% separating gel:**

	<u>glycerol, 13%</u>
ddH <sub>2</sub> O	0.9 mL
30% acrylamide:bis (37.5:1)	2 mL
1.5 M Tris-HCl pH 8.8	1.3 mL
10% (w/v) SDS	50 µL
75% glycerol	0.7 mL
10% (w/v) ammonium persulfate	50 µL
TEMED	2 µL

**2 ml 5% stacking gel (for stacking two gels):**

ddH <sub>2</sub> O	1.4 mL
30% acrylamide:bis (37.5:1)	330 µL
1.0 M Tris-HCl pH 6.8	250 µL
10% (w/v) SDS	20 µL
10% (w/v) ammonium persulfate	20 µL
TEMED	2 µL

**5 ml 1x glycerol loading buffer (13% glycerol):**

75% glycerol	0.7 mL (13%)
10% (w/v) SDS	1 ml (2%)
1 M Tris-HCl pH 6.8	250 µL (50 mM)
bromophenol blue	5 mg (0.1%)
ddH <sub>2</sub> O	to 5 ml

## Appendix C: List of assigned chemical shifts

Listed in the following table are assigned chemical shifts for each residue in the BNIP3 transmembrane domain peptide.

The chirality for Val  $\gamma$ -methyls and Leu  $\delta$ -methyls in this list are resolved, with the exceptions for Leu162 and Leu187, though a different list of chemical shifts was used for BNIP3 transmembrane domain dimer structure calculation.

The chirality for methylene protons are not resolved, except for His173  $H^\beta$  and  $H^\beta$  of Leu166, Leu169, Leu170, Leu171, Leu174, Leu175, and Leu179.

Residue #	Residue	Atom	Chemical shift
154	GLY	HA2	3.88
154	GLY	C	169.94
154	GLY	CA	42.77
155	GLY	H	8.57
155	GLY	HA2	3.95
155	GLY	HA3	4.07
155	GLY	C	173.69
155	GLY	CA	44.56
155	GLY	N	108.94
156	ILE	H	8.34
156	ILE	HA	3.64
156	ILE	HB	1.41
156	ILE	HG12	0.80
156	ILE	HG13	0.97
156	ILE	HG2	0.30
156	ILE	HD1	0.53
156	ILE	C	174.47
156	ILE	CA	62.06
156	ILE	CB	37.68
156	ILE	CG1	26.74
156	ILE	CG2	16.59
156	ILE	CD1	12.99
156	ILE	N	119.87
157	PHE	H	7.85
157	PHE	HA	4.57
157	PHE	HB2	3.14
157	PHE	HB3	2.97
157	PHE	HD1	7.13
157	PHE	HE1	7.09
157	PHE	C	174.32
157	PHE	CA	56.31
157	PHE	CB	37.84
157	PHE	CD1	131.23
157	PHE	CE1	130.24
157	PHE	N	117.79
158	SER	H	7.97
158	SER	HA	4.56
158	SER	HB2	3.96
158	SER	HB3	4.18
158	SER	C	174.16
158	SER	CA	57.01
158	SER	CB	64.87
158	SER	N	115.67
159	ALA	H	9.00
159	ALA	HA	4.08

Residue #	Residue	Atom	Chemical shift
159	ALA	HB	1.52
159	ALA	C	78.29
159	ALA	CA	54.85
159	ALA	CB	18.07
159	ALA	N	125.39
160	GLU	H	8.59
160	GLU	HA	3.99
160	GLU	HB2	2.08
160	GLU	HB3	2.01
160	GLU	HG2	2.43
160	GLU	HG3	2.35
160	GLU	C	177.44
160	GLU	CA	59.01
160	GLU	CB	28.22
160	GLU	CG	35.18
160	GLU	N	116.18
161	PHE	H	7.86
161	PHE	HA	4.33
161	PHE	HB2	3.27
161	PHE	HB3	3.25
161	PHE	HD1	7.14
161	PHE	HE1	6.84
161	PHE	C	177.13
161	PHE	CA	60.33
161	PHE	CB	39.16
161	PHE	CD1	131.53
161	PHE	CE1	130.00
161	PHE	N	118.05
162	LEU	H	8.24
162	LEU	HA	3.99
162	LEU	HB2	1.55
162	LEU	HB3	2.05
162	LEU	HG	2.05
162	LEU	HD1	1.00
162	LEU	HD2	1.01
162	LEU	C	177.13
162	LEU	CA	57.99
162	LEU	CB	41.51
162	LEU	CG	26.43
162	LEU	CD1	25.26
162	LEU	CD2	22.84
162	LEU	N	119.11
163	LYS	H	7.95
163	LYS	HA	3.92
163	LYS	HB2	1.94
163	LYS	HB3	1.90
163	LYS	HG2	1.64

Residue #	Residue	Atom	Chemical shift
163	LYS	HG3	1.42
163	LYS	HD2	1.73
163	LYS	HD3	1.70
163	LYS	HE2	2.89
163	LYS	HE3	2.85
163	LYS	C	174.01
163	LYS	CA	58.85
163	LYS	CB	32.39
163	LYS	CG	25.65
163	LYS	CD	29.32
163	LYS	CE	41.59
163	LYS	N	115.32
164	VAL	H	6.94
164	VAL	HA	4.18
164	VAL	HB	1.96
164	VAL	HG1	0.83
164	VAL	HG2	0.97
164	VAL	C	177.29
164	VAL	CA	62.68
164	VAL	CB	32.68
164	VAL	CG1	21.19
164	VAL	CG2	21.12
164	VAL	N	110.25
165	PHE	H	8.61
165	PHE	HA	4.18
165	PHE	HB2	2.60
165	PHE	HB3	2.56
165	PHE	HD1	7.04
165	PHE	HE1	7.24
165	PHE	C	175.72
165	PHE	CA	60.41
165	PHE	CB	40.33
165	PHE	CD1	131.06
165	PHE	CE1	130.18
165	PHE	N	121.29
166	LEU	H	8.42
166	LEU	HA	4.16
166	LEU	HB2	1.95
166	LEU	HB3	1.50
166	LEU	HG	1.85
166	LEU	HD1	0.97
166	LEU	HD2	0.90
166	LEU	CA	59.09
166	LEU	CB	38.38
166	LEU	CG	26.98
166	LEU	CD1	24.87
166	LEU	CD2	22.84

Residue #	Residue	Atom	Chemical shift
166	LEU	N	115.42
167	PRO	HA	4.13
167	PRO	HB2	2.09
167	PRO	HB3	1.81
167	PRO	HG2	2.46
167	PRO	HG3	1.84
167	PRO	HD2	3.55
167	PRO	HD3	3.50
167	PRO	C	177.15
167	PRO	CA	65.80
167	PRO	CB	30.34
167	PRO	CG	28.15
167	PRO	CD	49.71
168	SER	H	7.15
168	SER	HA	4.04
168	SER	HB2	4.02
168	SER	HB3	3.68
168	SER	C	174.16
168	SER	CA	63.33
168	SER	CB	62.06
168	SER	N	112.33
169	LEU	H	7.77
169	LEU	HA	4.80
169	LEU	HB2	1.71
169	LEU	HB3	1.32
169	LEU	HG	1.57
169	LEU	HD1	0.74
169	LEU	HD2	0.83
169	LEU	C	180.10
169	LEU	CA	55.34
169	LEU	CB	41.98
169	LEU	CG	26.04
169	LEU	CD1	25.10
169	LEU	CD2	23.62
169	LEU	N	120.48
170	LEU	H	8.05
170	LEU	HA	4.07
170	LEU	HB2	1.70
170	LEU	HB3	1.73
170	LEU	HG	1.67
170	LEU	HD1	0.82
170	LEU	HD2	0.78
170	LEU	C	176.97
170	LEU	CA	57.91
170	LEU	CB	40.80
170	LEU	CG	26.51
170	LEU	CD1	23.77

Residue #	Residue	Atom	Chemical shift
170	LEU	CD2	23.85
170	LEU	N	122.15
171	LEU	H	7.94
171	LEU	HA	3.99
171	LEU	HB2	1.77
171	LEU	HB3	1.75
171	LEU	HG	1.70
171	LEU	HD1	0.84
171	LEU	HD2	0.88
171	LEU	C	177.91
171	LEU	CA	57.76
171	LEU	CB	41.12
171	LEU	CG	26.59
171	LEU	CD1	23.96
171	LEU	CD2	23.56
171	LEU	N	117.49
172	SER	H	8.97
172	SER	HA	3.84
172	SER	HB2	3.73
172	SER	HB3	3.89
172	SER	C	176.82
172	SER	CA	63.19
172	SER	CB	59.73
172	SER	N	115.97
173	HIS	H	7.60
173	HIS	HA	4.08
173	HIS	HB2	2.91
173	HIS	HB3	2.73
173	HIS	HD2	6.15
173	HIS	HE1	6.82
173	HIS	C	176.97
173	HIS	CA	62.22
173	HIS	CB	27.44
173	HIS	CD2	96.31
173	HIS	CE1	135.45
173	HIS	N	119.72
174	LEU	H	8.20
174	LEU	HA	4.07
174	LEU	HB2	1.98
174	LEU	HB3	1.51
174	LEU	HG	1.87
174	LEU	HD1	0.79
174	LEU	HD2	0.78
174	LEU	C	178.22
174	LEU	CA	57.91
174	LEU	CB	40.96
174	LEU	CG	26.27

Residue #	Residue	Atom	Chemical shift
174	LEU	CD1	24.48
174	LEU	CD2	22.91
174	LEU	N	120.78
175	LEU	H	8.51
175	LEU	HA	3.93
175	LEU	HB2	1.93
175	LEU	HB3	1.45
175	LEU	HG	1.84
175	LEU	HD1	0.79
175	LEU	HD2	0.77
175	LEU	C	177.44
175	LEU	CA	57.76
175	LEU	CB	40.80
175	LEU	CG	26.19
175	LEU	CD1	24.63
175	LEU	CD2	22.84
175	LEU	N	118.60
176	ALA	H	7.83
176	ALA	HA	3.97
176	ALA	HB	1.75
176	ALA	C	178.22
176	ALA	CA	55.26
176	ALA	CB	18.07
176	ALA	N	121.03
177	ILE	H	8.21
177	ILE	HA	3.90
177	ILE	HB	2.06
177	ILE	HG12	0.91
177	ILE	HG13	0.95
177	ILE	HG2	0.85
177	ILE	HD1	0.82
177	ILE	C	178.38
177	ILE	CA	64.87
177	ILE	CB	37.37
177	ILE	CG1	28.46
177	ILE	CG2	16.43
177	ILE	CD1	13.07
177	ILE	N	118.33
178	GLY	H	8.59
178	GLY	HA2	3.60
178	GLY	HA3	3.63
178	GLY	C	173.85
178	GLY	CA	47.53
178	GLY	N	107.32
179	LEU	H	8.83
179	LEU	HA	4.11
179	LEU	HB2	2.03



Residue #	Residue	Atom	Chemical shift
179	LEU	HB3	1.64
179	LEU	HG	1.71
179	LEU	HD1	0.87
179	LEU	HD2	0.94
179	LEU	C	178.22
179	LEU	CA	57.60
179	LEU	CB	41.04
179	LEU	CG	26.43
179	LEU	CD1	24.79
179	LEU	CD2	23.07
179	LEU	N	122.35
180	GLY	H	8.52
180	GLY	HA2	3.76
180	GLY	HA3	4.52
180	GLY	C	174.16
180	GLY	CA	47.92
180	GLY	N	106.26
181	ILE	H	8.78
181	ILE	HA	3.59
181	ILE	HB	2.15
181	ILE	HG12	0.95
181	ILE	HG13	1.01
181	ILE	HG2	0.82
181	ILE	HD1	0.84
181	ILE	C	177.13
181	ILE	CA	65.34
181	ILE	CB	36.90
181	ILE	CG1	28.93
181	ILE	CG2	16.90
181	ILE	CD1	13.93
181	ILE	N	123.97
182	TYR	H	8.11
182	TYR	HA	3.88
182	TYR	HB2	3.49
182	TYR	HB3	2.99
182	TYR	HD1	6.92
182	TYR	HE1	6.70
182	TYR	C	176.97
182	TYR	CA	61.90
182	TYR	CB	37.91
182	TYR	CD1	132.05
182	TYR	CE1	117.41
182	TYR	N	119.57
183	ILE	H	8.49
183	ILE	HA	3.40
183	ILE	HB	2.04
183	ILE	HG12	1.01

Residue #	Residue	Atom	Chemical shift
183	ILE	HG13	1.06
183	ILE	HG2	0.86
183	ILE	HD1	0.85
183	ILE	C	177.76
183	ILE	CA	65.02
183	ILE	CB	37.60
183	ILE	CG1	28.46
183	ILE	CG2	16.74
183	ILE	CD1	13.15
183	ILE	N	118.00
184	GLY	H	9.36
184	GLY	HA2	3.65
184	GLY	HA3	3.94
184	GLY	C	174.32
184	GLY	CA	47.38
184	GLY	N	106.26
185	ARG	H	8.09
185	ARG	HA	4.03
185	ARG	HB2	1.86
185	ARG	HB3	1.78
185	ARG	HG2	1.76
185	ARG	HG3	1.74
185	ARG	HD2	3.18
185	ARG	HD3	3.08
185	ARG	HE	7.61
185	ARG	HH11	6.86
185	ARG	C	177.29
185	ARG	CA	57.05
185	ARG	CB	29.63
185	ARG	CG	26.98
185	ARG	CD	42.60
185	ARG	CZ	179.47
185	ARG	N	118.05
185	ARG	NE	83.75
185	ARG	NH1	72.20
186	ARG	H	7.76
186	ARG	HA	4.26
186	ARG	HB2	1.87
186	ARG	HB3	1.63
186	ARG	HG2	1.42
186	ARG	HG3	1.40
186	ARG	HD2	2.92
186	ARG	HD3	2.89
186	ARG	HE	7.40
186	ARG	HH11	6.83
186	ARG	C	176.19
186	ARG	CA	55.67

Residue #	Residue	Atom	Chemical shift
186	ARG	CB	29.71
186	ARG	CG	26.51
186	ARG	CD	42.05
186	ARG	CZ	179.16
186	ARG	N	117.39
186	ARG	NE	84.50
186	ARG	NH1	72.20
187	LEU	H	7.53
187	LEU	HA	4.39
187	LEU	HB2	1.66
187	LEU	HB3	1.77
187	LEU	HG	1.90
187	LEU	HD1	0.87
187	LEU	HD2	0.84
187	LEU	C	175.41
187	LEU	CA	54.44
187	LEU	CB	41.59
187	LEU	CG	25.65
187	LEU	CD1	25.49
187	LEU	CD2	22.60
187	LEU	N	119.36
188	THR	H	7.43
188	THR	HA	4.13
188	THR	HB	4.24
188	THR	HG2	1.19
188	THR	CA	62.78
188	THR	CB	70.18
188	THR	CG2	21.63
188	THR	N	116.99

## **Appendix D: Lists of experimentally derived restraints**

	Page
Helical hydrogen bond restraint table	204
Torsion angle restraint table	207
Karplus restraint table	213
List of inter-monomer NOE to the amide proton	214
List of inter-monomer NOE to the carbon	215

```
! bnip3_hbond_two_ver8.tbl
```

```
! Hydrogen bond restraints for alpha-helix are applied for residues 168-185, which is an  
! alpha-helical region based on evaluation with CSI
```

```
! O::HN restraint for I -> i-4 H bonds
```

```
! helix "AAAA"
```

```
assign ( segid "AAAA" and residue 168 and name O )  
      ( segid "AAAA" and residue 172 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 169 and name O )  
      ( segid "AAAA" and residue 173 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 170 and name O )  
      ( segid "AAAA" and residue 174 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 171 and name O )  
      ( segid "AAAA" and residue 175 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 172 and name O )  
      ( segid "AAAA" and residue 176 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 173 and name O )  
      ( segid "AAAA" and residue 177 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 174 and name O )  
      ( segid "AAAA" and residue 178 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 175 and name O )  
      ( segid "AAAA" and residue 179 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 176 and name O )  
      ( segid "AAAA" and residue 180 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 177 and name O )  
      ( segid "AAAA" and residue 181 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 178 and name O )  
      ( segid "AAAA" and residue 182 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 179 and name O )  
      ( segid "AAAA" and residue 183 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 180 and name O )  
      ( segid "AAAA" and residue 184 and name HN ) 1.9 0.3 0.3  
assign ( segid "AAAA" and residue 181 and name O )  
      ( segid "AAAA" and residue 185 and name HN ) 1.9 0.3 0.3
```

```
! O::HN restraint for I -> i-4 H bonds
```

```
! helix "AAAA"
```

```
assign ( segid "AAAA" and residue 168 and name O )  
      ( segid "AAAA" and residue 172 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 169 and name O )  
      ( segid "AAAA" and residue 173 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 170 and name O )  
      ( segid "AAAA" and residue 174 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 171 and name O )  
      ( segid "AAAA" and residue 175 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 172 and name O )  
      ( segid "AAAA" and residue 176 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 173 and name O )  
      ( segid "AAAA" and residue 177 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 174 and name O )  
      ( segid "AAAA" and residue 178 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 175 and name O )  
      ( segid "AAAA" and residue 179 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 176 and name O )  
      ( segid "AAAA" and residue 180 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 177 and name O )  
      ( segid "AAAA" and residue 181 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 178 and name O )  
      ( segid "AAAA" and residue 182 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 179 and name O )  
      ( segid "AAAA" and residue 183 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 180 and name O )  
      ( segid "AAAA" and residue 184 and name N ) 2.9 0.3 0.3  
assign ( segid "AAAA" and residue 181 and name O )  
      ( segid "AAAA" and residue 185 and name N ) 2.9 0.3 0.3
```



```

! Inter-molecular hydrogen bond restraints between His173 NE2 (HE2) and Ser172 OG
! helix "AAAA"
assign ( segid "AAAA" and residue 172 and name OG )
      ( segid "BBBB" and residue 173 and name NE2 ) 2.9  0.3  0.3
assign ( segid "AAAA" and residue 172 and name OG )
      ( segid "BBBB" and residue 173 and name HE2 ) 1.9  0.3  0.3

! helix "BBBB"
assign ( segid "BBBB" and residue 172 and name OG )
      ( segid "AAAA" and residue 173 and name NE2 ) 2.9  0.3  0.3
assign ( segid "BBBB" and residue 172 and name OG )
      ( segid "AAAA" and residue 173 and name HE2 ) 1.9  0.3  0.3

! Intra-molecular hydrogen bond restraints between Ser172 HG and Ser168 O
! helix "AAAA"
assign ( segid "AAAA" and residue 168 and name O )
      ( segid "AAAA" and residue 172 and name HG ) 1.9  0.3  0.3

! helix "BBBB"
assign ( segid "BBBB" and residue 168 and name O )
      ( segid "BBBB" and residue 172 and name HG ) 1.9  0.3  0.3

```

```
! bnip3_dihedrals_two_ver15.tbl
```

```
! Backbone torsion angle restraints (phi and psi) for alpha-helix are applied for
! residues 168-185, which is an alpha-helical region based on evaluation with CSI
```

```
! helix "AAAA", phi angle
```

```
assign (segid "AAAA" and resid 168 and name C) (segid "AAAA" and resid 169 and name N)
      (segid "AAAA" and resid 169 and name CA) (segid "AAAA" and resid 169 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 169 and name C) (segid "AAAA" and resid 170 and name N)
      (segid "AAAA" and resid 170 and name CA) (segid "AAAA" and resid 170 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 170 and name C) (segid "AAAA" and resid 171 and name N)
      (segid "AAAA" and resid 171 and name CA) (segid "AAAA" and resid 171 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 171 and name C) (segid "AAAA" and resid 172 and name N)
      (segid "AAAA" and resid 172 and name CA) (segid "AAAA" and resid 172 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 172 and name C) (segid "AAAA" and resid 173 and name N)
      (segid "AAAA" and resid 173 and name CA) (segid "AAAA" and resid 173 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 173 and name C) (segid "AAAA" and resid 174 and name N)
      (segid "AAAA" and resid 174 and name CA) (segid "AAAA" and resid 174 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 174 and name C) (segid "AAAA" and resid 175 and name N)
      (segid "AAAA" and resid 175 and name CA) (segid "AAAA" and resid 175 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 175 and name C) (segid "AAAA" and resid 176 and name N)
      (segid "AAAA" and resid 176 and name CA) (segid "AAAA" and resid 176 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 176 and name C) (segid "AAAA" and resid 177 and name N)
      (segid "AAAA" and resid 177 and name CA) (segid "AAAA" and resid 177 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 177 and name C) (segid "AAAA" and resid 178 and name N)
      (segid "AAAA" and resid 178 and name CA) (segid "AAAA" and resid 178 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 178 and name C) (segid "AAAA" and resid 179 and name N)
      (segid "AAAA" and resid 179 and name CA) (segid "AAAA" and resid 179 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 179 and name C) (segid "AAAA" and resid 180 and name N)
      (segid "AAAA" and resid 180 and name CA) (segid "AAAA" and resid 180 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 180 and name C) (segid "AAAA" and resid 181 and name N)
      (segid "AAAA" and resid 181 and name CA) (segid "AAAA" and resid 181 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 181 and name C) (segid "AAAA" and resid 182 and name N)
      (segid "AAAA" and resid 182 and name CA) (segid "AAAA" and resid 182 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 182 and name C) (segid "AAAA" and resid 183 and name N)
      (segid "AAAA" and resid 183 and name CA) (segid "AAAA" and resid 183 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 183 and name C) (segid "AAAA" and resid 184 and name N)
      (segid "AAAA" and resid 184 and name CA) (segid "AAAA" and resid 184 and name C)
1 -57 5 2
assign (segid "AAAA" and resid 184 and name C) (segid "AAAA" and resid 185 and name N)
      (segid "AAAA" and resid 185 and name CA) (segid "AAAA" and resid 185 and name C)
1 -57 5 2
```

```
! helix "AAAA", psi angle
```

```
assign (segid "AAAA" and resid 168 and name N) (segid "AAAA" and resid 168 and name CA)
      (segid "AAAA" and resid 168 and name C) (segid "AAAA" and resid 169 and name N)
1 -47 5 2
assign (segid "AAAA" and resid 169 and name N) (segid "AAAA" and resid 169 and name CA)
      (segid "AAAA" and resid 169 and name C) (segid "AAAA" and resid 170 and name N)
1 -47 5 2
assign (segid "AAAA" and resid 170 and name N) (segid "AAAA" and resid 170 and name CA)
      (segid "AAAA" and resid 170 and name C) (segid "AAAA" and resid 171 and name N)
1 -47 5 2
```



[illegible]



```

assign (segid "BBBB" and resid 181 and name N) (segid "BBBB" and resid 181 and name CA)
      (segid "BBBB" and resid 181 and name C) (segid "BBBB" and resid 182 and name N)
1 -47 5 2
assign (segid "BBBB" and resid 182 and name N) (segid "BBBB" and resid 182 and name CA)
      (segid "BBBB" and resid 182 and name C) (segid "BBBB" and resid 183 and name N)
1 -47 5 2
assign (segid "BBBB" and resid 183 and name N) (segid "BBBB" and resid 183 and name CA)
      (segid "BBBB" and resid 183 and name C) (segid "BBBB" and resid 184 and name N)
1 -47 5 2
assign (segid "BBBB" and resid 184 and name N) (segid "BBBB" and resid 184 and name CA)
      (segid "BBBB" and resid 184 and name C) (segid "BBBB" and resid 185 and name N)
1 -47 5 2

! Side chain torsion angle restraints for isoleucines obtained from J-coupling
! experiments Jcco and Jcn
! helix "AAAA", chi-1 restraints
assign (segid "AAAA" and resid 177 and name N) (segid "AAAA" and resid 177 and name CA)
      (segid "AAAA" and resid 177 and name CB) (segid "AAAA" and resid 177 and name CG2)
1 180 10 2
assign (segid "AAAA" and resid 181 and name N) (segid "AAAA" and resid 181 and name CA)
      (segid "AAAA" and resid 181 and name CB) (segid "AAAA" and resid 181 and name CG2)
1 180 10 2

! helix "BBBB", chi-1 restraints
assign (segid "BBBB" and resid 177 and name N) (segid "BBBB" and resid 177 and name CA)
      (segid "BBBB" and resid 177 and name CB) (segid "BBBB" and resid 177 and name CG2)
1 180 10 2
assign (segid "BBBB" and resid 181 and name N) (segid "BBBB" and resid 181 and name CA)
      (segid "BBBB" and resid 181 and name CB) (segid "BBBB" and resid 181 and name CG2)
1 180 10 2

! Side chain torsion angle restraints chi-1 for His173, Leu166, Leu169, Leu170, Leu171,
! Leu174, Leu175, Leu179, and side chain torsion angle restraints chi-2 for Leu166,
! Leu169, Leu170, Leu171, Leu174, Leu175, Leu179
! The chi-1 angle restraints were determined by examining the intensities of NOE cross
! peaks
! The chi-2 angle restraints were determined based on data from J-coupling Jcca
! experiment and by examining the intensities of NOE cross peaks
! helix "AAAA", chi-1 angle
assign (segid "AAAA" and resid 173 and name N) (segid "AAAA" and resid 173 and name CA)
      (segid "AAAA" and resid 173 and name CB) (segid "AAAA" and resid 173 and name HB2)
1 60 10 2

assign (segid "AAAA" and resid 166 and name N) (segid "AAAA" and resid 166 and name CA)
      (segid "AAAA" and resid 166 and name CB) (segid "AAAA" and resid 166 and name HB2)
1 60 10 2
assign (segid "AAAA" and resid 169 and name N) (segid "AAAA" and resid 169 and name CA)
      (segid "AAAA" and resid 169 and name CB) (segid "AAAA" and resid 169 and name HB2)
1 60 10 2
assign (segid "AAAA" and resid 170 and name N) (segid "AAAA" and resid 170 and name CA)
      (segid "AAAA" and resid 170 and name CB) (segid "AAAA" and resid 170 and name HB2)
1 60 10 2
assign (segid "AAAA" and resid 171 and name N) (segid "AAAA" and resid 171 and name CA)
      (segid "AAAA" and resid 171 and name CB) (segid "AAAA" and resid 171 and name HB2)
1 60 10 2
assign (segid "AAAA" and resid 174 and name N) (segid "AAAA" and resid 174 and name CA)
      (segid "AAAA" and resid 174 and name CB) (segid "AAAA" and resid 174 and name HB2)
1 60 10 2
assign (segid "AAAA" and resid 175 and name N) (segid "AAAA" and resid 175 and name CA)
      (segid "AAAA" and resid 175 and name CB) (segid "AAAA" and resid 175 and name HB2)
1 60 10 2
assign (segid "AAAA" and resid 179 and name N) (segid "AAAA" and resid 179 and name CA)
      (segid "AAAA" and resid 179 and name CB) (segid "AAAA" and resid 179 and name HB2)
1 60 10 2

```

```

! helix "BBBB", chi-1 angle
assign (segid "BBBB" and resid 173 and name N) (segid "BBBB" and resid 173 and name CA)
      (segid "BBBB" and resid 173 and name CB) (segid "BBBB" and resid 173 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 166 and name N) (segid "BBBB" and resid 166 and name CA)
      (segid "BBBB" and resid 166 and name CB) (segid "BBBB" and resid 166 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 169 and name N) (segid "BBBB" and resid 169 and name CA)
      (segid "BBBB" and resid 169 and name CB) (segid "BBBB" and resid 169 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 170 and name N) (segid "BBBB" and resid 170 and name CA)
      (segid "BBBB" and resid 170 and name CB) (segid "BBBB" and resid 170 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 171 and name N) (segid "BBBB" and resid 171 and name CA)
      (segid "BBBB" and resid 171 and name CB) (segid "BBBB" and resid 171 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 174 and name N) (segid "BBBB" and resid 174 and name CA)
      (segid "BBBB" and resid 174 and name CB) (segid "BBBB" and resid 174 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 175 and name N) (segid "BBBB" and resid 175 and name CA)
      (segid "BBBB" and resid 175 and name CB) (segid "BBBB" and resid 175 and name HB2)
1 60 10 2

assign (segid "BBBB" and resid 179 and name N) (segid "BBBB" and resid 179 and name CA)
      (segid "BBBB" and resid 179 and name CB) (segid "BBBB" and resid 179 and name HB2)
1 60 10 2

! helix "AAAA", chi-2 angle
assign (segid "AAAA" and resid 171 and name CA) (segid "AAAA" and resid 171 and name CB)
      (segid "AAAA" and resid 171 and name CG) (segid "AAAA" and resid 171 and name CD1)
1 180 10 2

assign (segid "AAAA" and resid 174 and name CA) (segid "AAAA" and resid 174 and name CB)
      (segid "AAAA" and resid 174 and name CG) (segid "AAAA" and resid 174 and name CD1)
1 180 10 2

assign (segid "AAAA" and resid 175 and name CA) (segid "AAAA" and resid 175 and name CB)
      (segid "AAAA" and resid 175 and name CG) (segid "AAAA" and resid 175 and name CD1)
1 180 10 2

assign (segid "AAAA" and resid 179 and name CA) (segid "AAAA" and resid 179 and name CB)
      (segid "AAAA" and resid 179 and name CG) (segid "AAAA" and resid 179 and name CD1)
1 180 10 2

! helix "BBBB", chi-2 angle
assign (segid "BBBB" and resid 171 and name CA) (segid "BBBB" and resid 171 and name CB)
      (segid "BBBB" and resid 171 and name CG) (segid "BBBB" and resid 171 and name CD1)
1 180 10 2

assign (segid "BBBB" and resid 174 and name CA) (segid "BBBB" and resid 174 and name CB)
      (segid "BBBB" and resid 174 and name CG) (segid "BBBB" and resid 174 and name CD1)
1 180 10 2

assign (segid "BBBB" and resid 175 and name CA) (segid "BBBB" and resid 175 and name CB)
      (segid "BBBB" and resid 175 and name CG) (segid "BBBB" and resid 175 and name CD1)
1 180 10 2

assign (segid "BBBB" and resid 179 and name CA) (segid "BBBB" and resid 179 and name CB)
      (segid "BBBB" and resid 179 and name CG) (segid "BBBB" and resid 179 and name CD1)
1 180 10 2

! Side chain torsion angle restraints (chi-1) for Ser172 OG and Ser168 OG
! helix "AAAA", chi-1 angle
assign (segid "AAAA" and resid 172 and name N) (segid "AAAA" and resid 172 and name CA)
      (segid "AAAA" and resid 172 and name CB) (segid "AAAA" and resid 172 and name OG)
1 -60 10 2

assign (segid "AAAA" and resid 168 and name N) (segid "AAAA" and resid 168 and name CA)
      (segid "AAAA" and resid 168 and name CB) (segid "AAAA" and resid 168 and name OG)
1 -60 10 2

! helix "BBBB", chi-1 angle
assign (segid "BBBB" and resid 172 and name N) (segid "BBBB" and resid 172 and name CA)
      (segid "BBBB" and resid 172 and name CB) (segid "BBBB" and resid 172 and name OG)
1 -60 10 2

```

```
assign (segid "BBBB" and resid 168 and name N) (segid "BBBB" and resid 168 and name CA)
      (segid "BBBB" and resid 168 and name CB) (segid "BBBB" and resid 168 and name OG)
1 -60 10 2
```

```

! bnip3_karplus_two.tbl
!
! Karplus restraints for phi angle were obtained from 3J(HN,HA)
! Restraints for F161 is not included because its amide overlap with F157 amide
! i.e. the peak intensity recorded for F161 amide may not be correct

coefficients 6.98 -1.38 1.72 -60.00

! Helix "AAAA"
assign (segid "AAAA" and resid 155 and name C) (segid "AAAA" and resid 156 and name N)
      (segid "AAAA" and resid 156 and name CA) (segid "AAAA" and resid 156 and name C)
      5.61      0.50
assign (segid "AAAA" and resid 156 and name C) (segid "AAAA" and resid 157 and name N)
      (segid "AAAA" and resid 157 and name CA) (segid "AAAA" and resid 157 and name C)
      6.39      0.50
assign (segid "AAAA" and resid 157 and name C) (segid "AAAA" and resid 158 and name N)
      (segid "AAAA" and resid 158 and name CA) (segid "AAAA" and resid 158 and name C)
      6.02      0.50
assign (segid "AAAA" and resid 158 and name C) (segid "AAAA" and resid 159 and name N)
      (segid "AAAA" and resid 159 and name CA) (segid "AAAA" and resid 159 and name C)
      3.79      0.50
assign (segid "AAAA" and resid 159 and name C) (segid "AAAA" and resid 160 and name N)
      (segid "AAAA" and resid 160 and name CA) (segid "AAAA" and resid 160 and name C)
      3.78      0.50
assign (segid "AAAA" and resid 161 and name C) (segid "AAAA" and resid 162 and name N)
      (segid "AAAA" and resid 162 and name CA) (segid "AAAA" and resid 162 and name C)
      4.48      0.50
assign (segid "AAAA" and resid 162 and name C) (segid "AAAA" and resid 163 and name N)
      (segid "AAAA" and resid 163 and name CA) (segid "AAAA" and resid 163 and name C)
      3.78      0.50

! Helix "BBBB"
assign (segid "BBBB" and resid 155 and name C) (segid "BBBB" and resid 156 and name N)
      (segid "BBBB" and resid 156 and name CA) (segid "BBBB" and resid 156 and name C)
      5.61      0.50
assign (segid "BBBB" and resid 156 and name C) (segid "BBBB" and resid 157 and name N)
      (segid "BBBB" and resid 157 and name CA) (segid "BBBB" and resid 157 and name C)
      6.39      0.50
assign (segid "BBBB" and resid 157 and name C) (segid "BBBB" and resid 158 and name N)
      (segid "BBBB" and resid 158 and name CA) (segid "BBBB" and resid 158 and name C)
      6.02      0.50
assign (segid "BBBB" and resid 158 and name C) (segid "BBBB" and resid 159 and name N)
      (segid "BBBB" and resid 159 and name CA) (segid "BBBB" and resid 159 and name C)
      3.79      0.50
assign (segid "BBBB" and resid 159 and name C) (segid "BBBB" and resid 160 and name N)
      (segid "BBBB" and resid 160 and name CA) (segid "BBBB" and resid 160 and name C)
      3.78      0.50
assign (segid "BBBB" and resid 161 and name C) (segid "BBBB" and resid 162 and name N)
      (segid "BBBB" and resid 162 and name CA) (segid "BBBB" and resid 162 and name C)
      4.48      0.50
assign (segid "BBBB" and resid 162 and name C) (segid "BBBB" and resid 163 and name N)
      (segid "BBBB" and resid 163 and name CA) (segid "BBBB" and resid 163 and name C)
      3.78      0.50

```

```

! 15Ninter.tbl
! Inter-monomer NOE for the amide region; ARIA uncalibrated data
! The peak intensities listed are half the original peak intensities

! Helix "AAAA"
ASSI ( resid 180 and name HN and segid "AAAA" )
      ( ( attr store1 < 0.885 and attr store1 > 0.785 ) and segid "BBBB" )
      6.0 0.1 0.1 peak 1 spectrum 5 volume -170705 hpm1 106.290 ppm1 8.480 ppm2 0.835
ASSI ( resid 180 and name HN and segid "AAAA" )
      ( resid 177 and name HA and segid "BBBB" )
      6.0 0.1 0.1 peak 2 spectrum 5 volume -154211 hpm1 106.388 ppm1 8.467 ppm2 3.910
ASSI ( resid 184 and name HN and segid "AAAA" )
      ( ( attr store1 < 0.880 and attr store1 > 0.780 ) and segid "BBBB" )
      6.0 0.1 0.1 peak 3 spectrum 5 volume -152472 hpm1 106.334 ppm1 9.302 ppm2 0.830
ASSI ( resid 184 and name HN and segid "AAAA" )
      ( resid 181 and name HA and segid "BBBB" )
      6.0 0.1 0.1 peak 4 spectrum 5 volume -175893 hpm1 106.013 ppm1 9.299 ppm2 3.617
!
! Helix "BBBB"
ASSI ( resid 180 and name HN and segid "BBBB" )
      ( ( attr store1 < 0.885 and attr store1 > 0.785 ) and segid "AAAA" )
      6.0 0.1 0.1 peak 1 spectrum 5 volume -170705 hpm1 106.290 ppm1 8.480 ppm2 0.835
ASSI ( resid 180 and name HN and segid "BBBB" )
      ( resid 177 and name HA and segid "AAAA" )
      6.0 0.1 0.1 peak 2 spectrum 5 volume -154211 hpm1 106.388 ppm1 8.467 ppm2 3.910
ASSI ( resid 184 and name HN and segid "BBBB" )
      ( ( attr store1 < 0.880 and attr store1 > 0.780 ) and segid "AAAA" )
      6.0 0.1 0.1 peak 3 spectrum 5 volume -152472 hpm1 106.334 ppm1 9.302 ppm2 0.830
ASSI ( resid 184 and name HN and segid "BBBB" )
      ( resid 181 and name HA and segid "AAAA" )
      6.0 0.1 0.1 peak 4 spectrum 5 volume -175893 hpm1 106.013 ppm1 9.299 ppm2 3.617

```

! 13Cinter.tbl  
 ! Inter-monomer NOE for the carbon; ARIA uncalibrated data  
 ! The peak intensities listed are half the original peak intensities

! Helix "AAAA"

```
ASSI ( resid 169 and name HD1# and segid "AAAA" )
      ( resid 165 and name HE# and segid "BBBB" )
      6.0 0.1 0.1 peak 3 spectrum 4 volume -271651 hpm1 25.675 ppm1 0.734 ppm2 7.235
ASSI ( resid 169 and name HD2# and segid "AAAA" )
      ( resid 165 and name HE# and segid "BBBB" )
      6.0 0.1 0.1 peak 5 spectrum 4 volume -229362 hpm1 24.231 ppm1 0.831 ppm2 7.238
ASSI ( resid 169 and name HD2# and segid "AAAA" )
      ( resid 173 and name HE1 and segid "BBBB" )
      6.0 0.1 0.1 peak 6 spectrum 4 volume -358411 hpm1 24.257 ppm1 0.830 ppm2 6.837
ASSI ( resid 173 and name HA and segid "AAAA" )
      ( resid 173 and name HD2 and segid "BBBB" )
      6.0 0.1 0.1 peak 7 spectrum 4 volume -195116 hpm1 62.635 ppm1 4.077 ppm2 6.161
ASSI ( resid 173 and name HA and segid "AAAA" )
      ( resid 176 and name HB# and segid "BBBB" )
      6.0 0.1 0.1 peak 8 spectrum 4 volume -157827 hpm1 62.665 ppm1 4.082 ppm2 1.747
ASSI ( resid 176 and name HB# and segid "AAAA" )
      ( resid 177 and name HD1# and segid "BBBB" )
      6.0 0.1 0.1 peak 10 spectrum 4 volume -468088 hpm1 18.598 ppm1 1.747 ppm2 0.823
ASSI ( resid 176 and name HB# and segid "AAAA" )
      ( resid 173 and name HA and segid "BBBB" )
      6.0 0.1 0.1 peak 11 spectrum 4 volume -325457 hpm1 18.607 ppm1 1.743 ppm2 4.086
ASSI ( resid 176 and name HB# and segid "AAAA" )
      ( resid 173 and name HD2 and segid "BBBB" )
      6.0 0.1 0.1 peak 12 spectrum 4 volume -729488 hpm1 18.610 ppm1 1.745 ppm2 6.159
ASSI ( resid 176 and name HB# and segid "AAAA" )
      ( resid 176 and name HB# and segid "BBBB" )
      6.0 0.1 0.1 peak 13 spectrum 4 volume -747975 hpm1 18.608 ppm1 1.746 ppm2 1.752
ASSI ( resid 176 and name HB# and segid "AAAA" )
      ( resid 177 and name HN and segid "BBBB" )
      6.0 0.1 0.1 peak 14 spectrum 4 volume -217663 hpm1 18.580 ppm1 1.741 ppm2 8.205
ASSI ( resid 176 and name HB# and segid "AAAA" )
      ( resid 177 and name HA and segid "BBBB" )
      6.0 0.1 0.1 peak 16 spectrum 4 volume -203406 hpm1 18.537 ppm1 1.742 ppm2 3.928
ASSI ( resid 177 and name HD1# and segid "AAAA" )
      ( resid 176 and name HA and segid "BBBB" )
      6.0 0.1 0.1 peak 17 spectrum 4 volume -202372 hpm1 13.764 ppm1 0.815 ppm2 3.982
ASSI ( resid 177 and name HD1# and segid "AAAA" )
      ( resid 176 and name HB# and segid "BBBB" )
      6.0 0.1 0.1 peak 18 spectrum 4 volume -479145 hpm1 13.696 ppm1 0.811 ppm2 1.748
ASSI ( resid 177 and name HG2# and segid "AAAA" )
      ( resid 179 and name HG and segid "BBBB" )
      6.0 0.1 0.1 peak 21 spectrum 4 volume -258283 hpm1 17.025 ppm1 0.844 ppm2 1.727
ASSI ( resid 177 and name HG2# and segid "AAAA" )
      ( resid 180 and name HA2 and segid "BBBB" )
      6.0 0.1 0.1 peak 22 spectrum 4 volume -135642 hpm1 17.020 ppm1 0.849 ppm2 3.771
ASSI ( resid 177 and name HG2# and segid "AAAA" )
      ( resid 177 and name HA and segid "BBBB" )
      6.0 0.1 0.1 peak 23 spectrum 4 volume -157538 hpm1 17.056 ppm1 0.849 ppm2 3.895
ASSI ( resid 177 and name HG2# and segid "AAAA" )
      ( ( attr store1 < 2.104 and attr store1 > 2.004 ) and segid "BBBB" )
      6.0 0.1 0.1 peak 24 spectrum 4 volume -160594 hpm1 17.017 ppm1 0.847 ppm2 2.054
ASSI ( resid 177 and name HG2# and segid "AAAA" )
      ( resid 180 and name HN and segid "BBBB" )
      6.0 0.1 0.1 peak 25 spectrum 4 volume -230894 hpm1 17.074 ppm1 0.846 ppm2 8.543
ASSI ( resid 177 and name HG2# and segid "AAAA" )
      ( resid 180 and name HA3 and segid "BBBB" )
      6.0 0.1 0.1 peak 26 spectrum 4 volume -158725 hpm1 17.032 ppm1 0.841 ppm2 4.555
ASSI ( resid 179 and name HG and segid "AAAA" )
      ( resid 177 and name HG2# and segid "BBBB" )
      6.0 0.1 0.1 peak 28 spectrum 4 volume -224386 hpm1 27.089 ppm1 1.700 ppm2 0.850
```



```

ASSI ( resid 181 and name HG2# and segid "AAAA" )
    ( resid 184 and name HA2 and segid "BBBB" )
6.0 0.1 0.1 peak 33 spectrum 4 volume -208236 hpm1 17.530 ppm1 0.816 ppm2 3.678
ASSI ( resid 181 and name HG2# and segid "AAAA" )
    ( resid 184 and name HA3 and segid "BBBB" )
6.0 0.1 0.1 peak 34 spectrum 4 volume -181653 hpm1 17.530 ppm1 0.816 ppm2 3.934
ASSI ( resid 181 and name HG2# and segid "AAAA" )
    ( resid 184 and name HN and segid "BBBB" )
6.0 0.1 0.1 peak 39 spectrum 4 volume -136403 hpm1 17.682 ppm1 0.814 ppm2 9.366
ASSI ( resid 181 and name HA and segid "AAAA" )
    ( ( attr store1 < 0.900 and attr store1 > 0.800 ) and segid "BBBB" )
6.0 0.1 0.1 peak 40 spectrum 4 volume -110995 hpm1 65.830 ppm1 3.603 ppm2 0.850
ASSI ( resid 181 and name HA and segid "AAAA" )
    ( resid 184 and name HA2 and segid "BBBB" )
6.0 0.1 0.1 peak 41 spectrum 4 volume -125655 hpm1 65.830 ppm1 3.620 ppm2 3.630
ASSI ( resid 181 and name HA and segid "AAAA" )
    ( resid 184 and name HA3 and segid "BBBB" )
6.0 0.1 0.1 peak 43 spectrum 4 volume -89129 hpm1 65.830 ppm1 3.623 ppm2 3.961
ASSI ( resid 177 and name HA and segid "AAAA" )
    ( resid 177 and name HA and segid "BBBB" )
6.0 0.1 0.1 peak 44 spectrum 4 volume -139679 hpm1 65.445 ppm1 3.907 ppm2 3.919
ASSI ( resid 177 and name HA and segid "AAAA" )
    ( resid 176 and name HB# and segid "BBBB" )
6.0 0.1 0.1 peak 45 spectrum 4 volume -106586 hpm1 65.451 ppm1 3.891 ppm2 1.747
ASSI ( resid 177 and name HA and segid "AAAA" )
    ( resid 180 and name HN and segid "BBBB" )
6.0 0.1 0.1 peak 46 spectrum 4 volume -111480 hpm1 65.392 ppm1 3.902 ppm2 8.540
ASSI ( resid 184 and name HA2 and segid "AAAA" )
    ( resid 181 and name HA and segid "BBBB" )
6.0 0.1 0.1 peak 49 spectrum 4 volume -112052 hpm1 48.000 ppm1 3.660 ppm2 3.622
ASSI ( resid 184 and name HA3 and segid "AAAA" )
    ( resid 181 and name HA and segid "BBBB" )
6.0 0.1 0.1 peak 50 spectrum 4 volume -103693 hpm1 48.000 ppm1 3.947 ppm2 3.612
ASSI ( resid 184 and name HA2 and segid "AAAA" )
    ( resid 181 and name HG2# and segid "BBBB" )
6.0 0.1 0.1 peak 51 spectrum 4 volume -99755 hpm1 48.000 ppm1 3.668 ppm2 0.812
ASSI ( resid 184 and name HA3 and segid "AAAA" )
    ( resid 181 and name HG2# and segid "BBBB" )
6.0 0.1 0.1 peak 52 spectrum 4 volume -91323 hpm1 48.000 ppm1 3.931 ppm2 0.812
ASSI ( resid 180 and name HA2 and segid "AAAA" )
    ( resid 177 and name HG2# and segid "BBBB" )
6.0 0.1 0.1 peak 53 spectrum 4 volume -163440 hpm1 48.530 ppm1 3.780 ppm2 0.850
ASSI ( resid 180 and name HA3 and segid "AAAA" )
    ( ( attr store1 < 0.894 and attr store1 > 0.794 ) and segid "BBBB" )
6.0 0.1 0.1 peak 54 spectrum 4 volume -148560 hpm1 48.498 ppm1 4.536 ppm2 0.844
ASSI ( resid 183 and name HB and segid "AAAA" )
    ( ( attr store1 < 0.897 and attr store1 > 0.797 ) and segid "BBBB" )
6.0 0.1 0.1 peak 64 spectrum 4 volume -198282 hpm1 38.135 ppm1 2.029 ppm2 0.847
ASSI ( resid 179 and name HB2 and segid "AAAA" )
    ( ( attr store1 < 0.902 and attr store1 > 0.802 ) and segid "BBBB" )
6.0 0.1 0.1 peak 72 spectrum 4 volume -66915 hpm1 41.593 ppm1 1.639 ppm2 0.852

! Helix "BBBB"

ASSI ( resid 169 and name HD1# and segid "BBBB" )
    ( resid 165 and name HE# and segid "AAAA" )
6.0 0.1 0.1 peak 3 spectrum 4 volume -271651 hpm1 25.675 ppm1 0.734 ppm2 7.235
ASSI ( resid 169 and name HD2# and segid "BBBB" )
    ( resid 165 and name HE# and segid "AAAA" )
6.0 0.1 0.1 peak 5 spectrum 4 volume -229362 hpm1 24.231 ppm1 0.831 ppm2 7.238
ASSI ( resid 169 and name HD2# and segid "BBBB" )
    ( resid 173 and name HE1 and segid "AAAA" )
6.0 0.1 0.1 peak 6 spectrum 4 volume -358411 hpm1 24.257 ppm1 0.830 ppm2 6.837
ASSI ( resid 173 and name HA and segid "BBBB" )
    ( resid 173 and name HD2 and segid "AAAA" )
6.0 0.1 0.1 peak 7 spectrum 4 volume -195116 hpm1 62.635 ppm1 4.077 ppm2 6.161
ASSI ( resid 173 and name HA and segid "BBBB" )
    ( resid 176 and name HB# and segid "AAAA" )
6.0 0.1 0.1 peak 8 spectrum 4 volume -157827 hpm1 62.665 ppm1 4.082 ppm2 1.747

```

```

ASSI ( resid 176 and name HB# and segid "BBBB" )
    ( resid 177 and name HD1# and segid "BBBB" )
6.0 0.1 0.1 peak 10 spectrum 4 volume -468088 hpm1 18.598 ppm1 1.747 ppm2 0.823
ASSI ( resid 176 and name HB# and segid "BBBB" )
    ( resid 173 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 11 spectrum 4 volume -325457 hpm1 18.607 ppm1 1.743 ppm2 4.086
ASSI ( resid 176 and name HB# and segid "BBBB" )
    ( resid 173 and name HD2 and segid "AAAA" )
6.0 0.1 0.1 peak 12 spectrum 4 volume -729488 hpm1 18.610 ppm1 1.745 ppm2 6.159
ASSI ( resid 176 and name HB# and segid "BBBB" )
    ( resid 176 and name HB# and segid "AAAA" )
6.0 0.1 0.1 peak 13 spectrum 4 volume -747975 hpm1 18.608 ppm1 1.746 ppm2 1.752
ASSI ( resid 176 and name HB# and segid "BBBB" )
    ( resid 177 and name HN and segid "AAAA" )
6.0 0.1 0.1 peak 14 spectrum 4 volume -217663 hpm1 18.580 ppm1 1.741 ppm2 8.205
ASSI ( resid 176 and name HB# and segid "BBBB" )
    ( resid 177 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 16 spectrum 4 volume -203406 hpm1 18.537 ppm1 1.742 ppm2 3.928
ASSI ( resid 177 and name HD1# and segid "BBBB" )
    ( resid 176 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 17 spectrum 4 volume -202372 hpm1 13.764 ppm1 0.815 ppm2 3.982
ASSI ( resid 177 and name HD1# and segid "BBBB" )
    ( resid 176 and name HB# and segid "AAAA" )
6.0 0.1 0.1 peak 18 spectrum 4 volume -479145 hpm1 13.696 ppm1 0.811 ppm2 1.748
ASSI ( resid 177 and name HG2# and segid "BBBB" )
    ( resid 179 and name HG and segid "AAAA" )
6.0 0.1 0.1 peak 21 spectrum 4 volume -258283 hpm1 17.025 ppm1 0.844 ppm2 1.727
ASSI ( resid 177 and name HG2# and segid "BBBB" )
    ( resid 180 and name HA2 and segid "AAAA" )
6.0 0.1 0.1 peak 22 spectrum 4 volume -135642 hpm1 17.020 ppm1 0.849 ppm2 3.771
ASSI ( resid 177 and name HG2# and segid "BBBB" )
    ( resid 177 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 23 spectrum 4 volume -157538 hpm1 17.056 ppm1 0.849 ppm2 3.895
ASSI ( resid 177 and name HG2# and segid "BBBB" )
    ( ( attr store1 < 2.104 and attr store1 > 2.004 ) and segid "AAAA" )
6.0 0.1 0.1 peak 24 spectrum 4 volume -160594 hpm1 17.017 ppm1 0.847 ppm2 2.054
ASSI ( resid 177 and name HG2# and segid "BBBB" )
    ( resid 180 and name HN and segid "AAAA" )
6.0 0.1 0.1 peak 25 spectrum 4 volume -230894 hpm1 17.074 ppm1 0.846 ppm2 8.543
ASSI ( resid 177 and name HG2# and segid "BBBB" )
    ( resid 180 and name HA3 and segid "AAAA" )
6.0 0.1 0.1 peak 26 spectrum 4 volume -158725 hpm1 17.032 ppm1 0.841 ppm2 4.555
ASSI ( resid 179 and name HG and segid "BBBB" )
    ( resid 177 and name HG2# and segid "AAAA" )
6.0 0.1 0.1 peak 28 spectrum 4 volume -224386 hpm1 27.089 ppm1 1.700 ppm2 0.850
ASSI ( resid 181 and name HG2# and segid "BBBB" )
    ( resid 184 and name HA2 and segid "AAAA" )
6.0 0.1 0.1 peak 33 spectrum 4 volume -208236 hpm1 17.530 ppm1 0.816 ppm2 3.678
ASSI ( resid 181 and name HG2# and segid "BBBB" )
    ( resid 184 and name HA3 and segid "AAAA" )
6.0 0.1 0.1 peak 34 spectrum 4 volume -181653 hpm1 17.530 ppm1 0.816 ppm2 3.934
ASSI ( resid 181 and name HG2# and segid "BBBB" )
    ( resid 184 and name HN and segid "AAAA" )
6.0 0.1 0.1 peak 39 spectrum 4 volume -136403 hpm1 17.682 ppm1 0.814 ppm2 9.366
ASSI ( resid 181 and name HA and segid "BBBB" )
    ( ( attr store1 < 0.900 and attr store1 > 0.800 ) and segid "AAAA" )
6.0 0.1 0.1 peak 40 spectrum 4 volume -110995 hpm1 65.830 ppm1 3.603 ppm2 0.850
ASSI ( resid 181 and name HA and segid "BBBB" )
    ( resid 184 and name HA2 and segid "AAAA" )
6.0 0.1 0.1 peak 41 spectrum 4 volume -125655 hpm1 65.830 ppm1 3.620 ppm2 3.630
ASSI ( resid 181 and name HA and segid "BBBB" )
    ( resid 184 and name HA3 and segid "AAAA" )
6.0 0.1 0.1 peak 43 spectrum 4 volume -89129 hpm1 65.830 ppm1 3.623 ppm2 3.961
ASSI ( resid 177 and name HA and segid "BBBB" )
    ( resid 177 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 44 spectrum 4 volume -139679 hpm1 65.445 ppm1 3.907 ppm2 3.919
ASSI ( resid 177 and name HA and segid "BBBB" )
    ( resid 176 and name HB# and segid "AAAA" )
6.0 0.1 0.1 peak 45 spectrum 4 volume -106586 hpm1 65.451 ppm1 3.891 ppm2 1.747

```

```

ASSI ( resid 177 and name HA and segid "BBBB" )
      ( resid 180 and name HN and segid "AAAA" )
6.0 0.1 0.1 peak 46 spectrum 4 volume -111480 hpm1 65.392 ppm1 3.902 ppm2 8.540
ASSI ( resid 184 and name HA2 and segid "BBBB" )
      ( resid 181 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 49 spectrum 4 volume -112052 hpm1 48.000 ppm1 3.660 ppm2 3.622
ASSI ( resid 184 and name HA3 and segid "BBBB" )
      ( resid 181 and name HA and segid "AAAA" )
6.0 0.1 0.1 peak 50 spectrum 4 volume -103693 hpm1 48.000 ppm1 3.947 ppm2 3.612
ASSI ( resid 184 and name HA2 and segid "BBBB" )
      ( resid 181 and name HG2# and segid "AAAA" )
6.0 0.1 0.1 peak 51 spectrum 4 volume -99755 hpm1 48.000 ppm1 3.668 ppm2 0.812
ASSI ( resid 184 and name HA3 and segid "BBBB" )
      ( resid 181 and name HG2# and segid "AAAA" )
6.0 0.1 0.1 peak 52 spectrum 4 volume -91323 hpm1 48.000 ppm1 3.931 ppm2 0.812
ASSI ( resid 180 and name HA2 and segid "BBBB" )
      ( resid 177 and name HG2# and segid "AAAA" )
6.0 0.1 0.1 peak 53 spectrum 4 volume -163440 hpm1 48.530 ppm1 3.780 ppm2 0.850
ASSI ( resid 180 and name HA3 and segid "BBBB" )
      ( ( attr store1 < 0.894 and attr store1 > 0.794 ) and segid "AAAA" )
6.0 0.1 0.1 peak 54 spectrum 4 volume -148560 hpm1 48.498 ppm1 4.536 ppm2 0.844
ASSI ( resid 183 and name HB and segid "BBBB" )
      ( ( attr store1 < 0.897 and attr store1 > 0.797 ) and segid "AAAA" )
6.0 0.1 0.1 peak 64 spectrum 4 volume -198282 hpm1 38.135 ppm1 2.029 ppm2 0.847
ASSI ( resid 179 and name HB2 and segid "BBBB" )
      ( ( attr store1 < 0.902 and attr store1 > 0.802 ) and segid "AAAA" )
6.0 0.1 0.1 peak 72 spectrum 4 volume -66915 hpm1 41.593 ppm1 1.639 ppm2 0.852

```

## Appendix E: Modified ARIA 1.2 code

	Page
run.cns	220
Modified code in the directory runX/protocols/	
generate.inp	239
newsymmetry.cns (for CA atoms of residues 168-185)	242
Modified code in the directory runX/data/sequence/ directory	
stereoassign.cns	243
secondarystructure.cns	245
Modified code in the directory runX/begin/	
setup_swap_list.tbl	246

```

!$Revision: 2.10 $
!$Date: 2002/07/23 16:19:27 $
!$RCSfile: run.cns,v $

module(
spectrum;
iteration;
filenames;
spectra;
data;
iterations;
saprotocol;
refine;
relax;
toppar;
analysis;
)

{+ file: run.cns +}
{+ description: The file run.cns contains all necessary information to run ARIA.
ARIA automatically sets the default values.
Please change the values for the mixing time, rotation
correlation time and spectrometer frequency for the spin diffusion correction.
For the large CNS log files, you should use a temporary directory. version 1.2 +}
{+ authors: Jens Linge, Michael Nilges +}

set message off echo off end

! Please cite the following references when using this protocol:
{+ reference: M. Nilges (1995) Calculation of protein structures with ambiguous distance
restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide
connectivities. J. Mol. Biol. 245, 645-660 +}
{+ reference: M. Nilges and Sean O'Donoghue (1998) Ambiguous NOEs and automated NOE
assignment. Prog. NMR Spect. 32, 107-139 +}
{+ reference: J.P. Linge and M. Nilges (1999) Influence of non-bonded parameters on the
quality of NMR structures: a new force-field for NMR structure calculation. J. Biomol.
NMR 13, 51-59 +}
{+ reference: J.P. Linge (2001) New methods for automated NOE assignment and NMR
structure calculation. Book on Demand Verlag ISBN: 383111482X +}
{+ reference: J.P. Linge, S.I. O'Donoghue and M. Nilges (2001). Assigning
Ambiguous NOEs with ARIA.
Methods in Enzymology, 339, 71-90. +}

{- Guidelines for using this file:
- all strings must be quoted by double-quotes
- logical variables (true/false) are not quoted
- do not remove any evaluate statements from the file
- pathnames should not exceed 80 characters -}
{- begin block parameter definition -} define(

{===== filenames =====}
{* the name of your current project *}
{==>} fileroot="dimer-070730";

{* RUN directory *}
{* the absolute path of your current run, e.g. /home/linge/werner/run3 *}
{==>} run_dir="/data/endah/ARIA1.2/dimer-070730/run4b";

{* molecular topology files (mtf) *}
{* will be generated by ARIA from the sequence / add other mtf files here *}
{==>} structure="dimer.psf";
{==>} structure_2="";
{==>} structure_3="";
{==>} structure_4="";
{==>} structure_5="";

{* PDB or sequence (3-letter code) file*}
{+ choice: "PDB" "sequence" +}

```

```

{==>} pdb_or_sequence="sequence";
{==>} prot_coor_1="bnip3.seq";
{==>} pdb_or_sequence_2="sequence";          ! ESS 02/01/2007
{==>} prot_coor_2="bnip3.seq";              ! ESS 02/01/2007

(* new segment id (segid) *)
(* will appear in the pdb files, must have 4 digits, leave it blank if you don't want one *)
{==>} prot_segid_1="AAAA";
{==>} prot_segid_2="BBBB";                  ! ESS 02/01/2007

(* Atomname nomenclature *)
(* set true if you have IUPAC (e.g. LEU HB2 and HB3 and not HB2 and HB1) data (e.g. from XEASY) *)
(+ choice: true false +)
{==>} xploriodiana=true;

(* ARIA directory *)
(* the absolute path of the ARIA program files *)
{==>} aria_dir="/divsite/opt.BIOC/aria1.2/redhat9";

(* Logfile directory *)
(* specify a directory for the large CNS log files *)
{==>} temprash_dir="/data/endah/ARIA1.2/dimer-070730/run4b/log";

{===== disulphide bonds =====}

(* number of disulphide bridges *)
(* set to zero if none are present *)
{==>} ss_bridge=0;

(* unambiguous or ambiguous disulphide bridges *)
(* default: unambiguous *)
(+ choice: "unambiguous" "ambiguous" +)
{==>} ss_ambigunambig="unambiguous";

(* repeat the four lines for each disulfide bond *)
(* first pair of entries are the resid and segid of the first cysteine *)
(* second pair of entries are the resid id and segid of the second cysteine *)
{==>}
ss_i_resid_1=11;
ss_i_segid_1="";
ss_j_resid_1=22;
ss_j_segid_1="";

ss_i_resid_2=33;
ss_i_segid_2="";
ss_j_resid_2=44;
ss_j_segid_2="";

ss_i_resid_3=55;
ss_i_segid_3="";
ss_j_resid_3=66;
ss_j_segid_3="";

ss_i_resid_4=77;
ss_i_segid_4="";
ss_j_resid_4=88;
ss_j_segid_4="";

ss_i_resid_5=99;
ss_i_segid_5="";
ss_j_resid_5=01;
ss_j_segid_5="";

{<==}

```

```

{===== histidine patches =====}
{* Patch to change doubly protonated HIS to singly protonated histidine (HD1) *}
{* just give the residue number of the histidines for the HISD patch, set them to zero if
you don't want them *}
{+ table: rows=2 "residue" "segid" cols=10 "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" +}
{==>} hisd_resid_1=0;
{==>} hisd_resid_2=0;
{==>} hisd_resid_3=0;
{==>} hisd_resid_4=0;
{==>} hisd_resid_5=0;
{==>} hisd_resid_6=0;
{==>} hisd_resid_7=0;
{==>} hisd_resid_8=0;
{==>} hisd_resid_9=0;
{==>} hisd_resid_10=0;
{==>} hisd_segid_1="";
{==>} hisd_segid_2="";
{==>} hisd_segid_3="";
{==>} hisd_segid_4="";
{==>} hisd_segid_5="";
{==>} hisd_segid_6="";
{==>} hisd_segid_7="";
{==>} hisd_segid_8="";
{==>} hisd_segid_9="";
{==>} hisd_segid_10="";

{* Patch to change doubly protonated HIS to singly protonated histidine (HE2) *}
{* just give the residue number of the histidines for the HISE patch, set them to zero if
you don't want them *}
{+ table: rows=2 "residue" "segid" cols=10 "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" +}
{==>} hise_resid_1=0;
{==>} hise_resid_2=0;
{==>} hise_resid_3=0;
{==>} hise_resid_4=0;
{==>} hise_resid_5=0;
{==>} hise_resid_6=0;
{==>} hise_resid_7=0;
{==>} hise_resid_8=0;
{==>} hise_resid_9=0;
{==>} hise_resid_10=0;
{==>} hise_segid_1="";
{==>} hise_segid_2="";
{==>} hise_segid_3="";
{==>} hise_segid_4="";
{==>} hise_segid_5="";
{==>} hise_segid_6="";
{==>} hise_segid_7="";
{==>} hise_segid_8="";
{==>} hise_segid_9="";
{==>} hise_segid_10="";

{===== Distance restraints =====}
{* Do you want to use hbond restraints? *}
{+ choice: true false +}
{==>} hbonds_on=true;

{* energy constants *}
{+ table: rows=3 "unambig" "ambig" "hbonds" cols=5 "firstIteration" "hot" "cool1_ini"
"cool1_fin" "cool2" +}

{==>} unamb_firstit=0;
{==>} unamb_hot=10;
{==>} unamb_cool1_ini=10;
{==>} unamb_cool1_fin=50;
{==>} unamb_cool2=50;
{==>} amb_firstit=0;
{==>} amb_hot=10;
{==>} amb_cool1_ini=10;

```

```

{====>} amb_cool1_fin=50;
{====>} amb_cool2=50;
{====>} hbond_firstit=0;
{====>} hbond_hot=10;
{====>} hbond_cool1_ini=10;
{====>} hbond_cool1_fin=50;
{====>} hbond_cool2=50;

(* potential shape *)
{+ table: rows=4 "mRswitch" "rswitch" "mAsymptote" "asymptote" cols=3 "hot" "cool1"
"cool2" +}

{====>} mrswi_hot=0.5;
{====>} mrswi_cool1=0.5;
{====>} mrswi_cool2=0.5;
{====>} rswi_hot=0.5;
{====>} rswi_cool1=0.5;
{====>} rswi_cool2=0.5;
{====>} masy_hot=-1.0;
{====>} masy_cool1=-1.0;
{====>} masy_cool2=-0.1;
{====>} asy_hot=1.0;
{====>} asy_cool1=1.0;
{====>} asy_cool2=0.1;

{===== dihedrals =====}
(* energy constants *)
{+ table: rows=1 "dihedrals" cols=4 "use?" "hot" "cool1" "cool2" +}

{+ choice: true false +}
{====>} dihedrals_on=true;
{====>} dihedrals_hot=5;
{====>} dihedrals_cool1=25;
{====>} dihedrals_cool2=200;

{===== Karplus coupling restraints =====}

(* Karplus coefficients *)
{+ table: rows=5 "class1" "class2" "class3" "class4" "class5"
cols=8 "use?" "A" "B" "C" "delta" "E(hot)" "E(cool1)" "E(cool2)" +}
{+ choice: true false +}
{====>} c1_on=true;
{====>} c1_karplusa=6.98;
{====>} c1_karplusb=-1.38;
{====>} c1_karplusc=1.72;
{====>} c1_karplusd=-60.0;
{====>} c1_hot=0.0;
{====>} c1_cool1=0.2;
{====>} c1_cool2=1.0;
{+ choice: true false +}
{====>} c2_on=false;
{====>} c2_karplusa=6.98;
{====>} c2_karplusb=-1.38;
{====>} c2_karplusc=1.72;
{====>} c2_karplusd=-120.0;
{====>} c2_hot=0.0;
{====>} c2_cool1=0.2;
{====>} c2_cool2=1.0;
{+ choice: true false +}
{====>} c3_on=false;
{====>} c3_karplusa=6.98;
{====>} c3_karplusb=-1.38;
{====>} c3_karplusc=1.72;
{====>} c3_karplusd=-120.0;
{====>} c3_hot=0.0;
{====>} c3_cool1=0.2;

```



```

{====> c3_cool2=1.0;
{+ choice: true false +}
{====> c4_on=false;
{====> c4_karplusa=6.98;
{====> c4_karplusb=-1.38;
{====> c4_karplusc=1.72;
{====> c4_karplusd=-120.0;
{====> c4_hot=0.0;
{====> c4_cool1=0.2;
{====> c4_cool2=1.0;
{+ choice: true false +}
{====> c5_on=false;
{====> c5_karplusa=6.98;
{====> c5_karplusb=-1.38;
{====> c5_karplusc=1.72;
{====> c5_karplusd=-120.0;
{====> c5_hot=0.0;
{====> c5_cool1=0.2;
{====> c5_cool2=1.0;

(===== residual dipolar couplings =====)

{* Parameters *}
{+ table: rows=5 "class1" "class2" "class3" "class4" "class5"
      cols=19 "type" "firstIt" "E(hot)" "E(cool1)" "E(cool2)" "R" "D" "ini_bor_hot"
"fin_bor_hot"
"ini_bor_cool1" "fin_bor_cool1" "ini_bor_cool2" "fin_bor_cool2" "ini_cen_hot"
"fin_cen_hot"
"ini_cen_cool1" "fin_cen_cool1" "ini_cen_cool2" "fin_cen_cool2" +}
{+ choice: "NO" "SANI" "VANGLE" +}
{====> rdc1_choice="NO";
{====> rdc1_firstIt=0;
{====> rdc1_hot=0.0;
{====> rdc1_cool1=0.2;
{====> rdc1_cool2=1.0;
{====> rdc1_r=0.4;
{====> rdc1_d=8.0;
{====> ini_bor_hot_1=0.1;
{====> fin_bor_hot_1=40.0;
{====> ini_bor_cool1_1=40.0;
{====> fin_bor_cool1_1=40.0;
{====> ini_bor_cool2_1=40.0;
{====> fin_bor_cool2_1=40.0;
{====> ini_cen_hot_1=0.1;
{====> fin_cen_hot_1=0.1;
{====> ini_cen_cool1_1=10.0;
{====> fin_cen_cool1_1=10.0;
{====> ini_cen_cool2_1=10.0;
{====> fin_cen_cool2_1=10.0;

{+ choice: "NO" "SANI" "VANGLE" +}
{====> rdc2_choice="NO";
{====> rdc2_firstIt=0;
{====> rdc2_hot=0.0;
{====> rdc2_cool1=0.2;
{====> rdc2_cool2=1.0;
{====> rdc2_r=0.4;
{====> rdc2_d=8.0;
{====> ini_bor_hot_2=0.1;
{====> fin_bor_hot_2=40.0;
{====> ini_bor_cool1_2=40.0;
{====> fin_bor_cool1_2=40.0;
{====> ini_bor_cool2_2=40.0;
{====> fin_bor_cool2_2=40.0;
{====> ini_cen_hot_2=0.1;
{====> fin_cen_hot_2=0.1;
{====> ini_cen_cool1_2=10.0;
{====> fin_cen_cool1_2=10.0;
{====> ini_cen_cool2_2=10.0;

```

```

{==>} fin_cen_cool2_2=10.0;

{+ choice: "NO" "SANI" "VANGLE" +}
{==>} rdc3_choice="NO";
{==>} rdc3_firstIt=0;
{==>} rdc3_hot=0.0;
{==>} rdc3_cool1=0.2;
{==>} rdc3_cool2=1.0;
{==>} rdc3_r=0.4;
{==>} rdc3_d=8.0;
{==>} ini_bor_hot_3=0.1;
{==>} fin_bor_hot_3=40.0;
{==>} ini_bor_cool1_3=40.0;
{==>} fin_bor_cool1_3=40.0;
{==>} ini_bor_cool2_3=40.0;
{==>} fin_bor_cool2_3=40.0;
{==>} ini_cen_hot_3=0.1;
{==>} fin_cen_hot_3=0.1;
{==>} ini_cen_cool1_3=10.0;
{==>} fin_cen_cool1_3=10.0;
{==>} ini_cen_cool2_3=10.0;
{==>} fin_cen_cool2_3=10.0;

{+ choice: "NO" "SANI" "VANGLE" +}
{==>} rdc4_choice="NO";
{==>} rdc4_firstIt=0;
{==>} rdc4_hot=0.0;
{==>} rdc4_cool1=0.2;
{==>} rdc4_cool2=1.0;
{==>} rdc4_r=0.4;
{==>} rdc4_d=8.0;
{==>} ini_bor_hot_4=0.1;
{==>} fin_bor_hot_4=40.0;
{==>} ini_bor_cool1_4=40.0;
{==>} fin_bor_cool1_4=40.0;
{==>} ini_bor_cool2_4=40.0;
{==>} fin_bor_cool2_4=40.0;
{==>} ini_cen_hot_4=0.1;
{==>} fin_cen_hot_4=0.1;
{==>} ini_cen_cool1_4=10.0;
{==>} fin_cen_cool1_4=10.0;
{==>} ini_cen_cool2_4=10.0;
{==>} fin_cen_cool2_4=10.0;

{+ choice: "NO" "SANI" "VANGLE" +}
{==>} rdc5_choice="NO";
{==>} rdc5_firstIt=0;
{==>} rdc5_hot=0.0;
{==>} rdc5_cool1=0.2;
{==>} rdc5_cool2=1.0;
{==>} rdc5_r=0.4;
{==>} rdc5_d=8.0;
{==>} ini_bor_hot_5=0.1;
{==>} fin_bor_hot_5=40.0;
{==>} ini_bor_cool1_5=40.0;
{==>} fin_bor_cool1_5=40.0;
{==>} ini_bor_cool2_5=40.0;
{==>} fin_bor_cool2_5=40.0;
{==>} ini_cen_hot_5=0.1;
{==>} fin_cen_hot_5=0.1;
{==>} ini_cen_cool1_5=10.0;
{==>} fin_cen_cool1_5=10.0;
{==>} ini_cen_cool2_5=10.0;
{==>} fin_cen_cool2_5=10.0;

{===== CSI restraints =====}
{* Do you want to use CSI derived hbond restraints? *}
{* uses the files generated by the program CSI *}
{+ choice: true false +}

```

```

{==>} hbondscsi_on=false;
{* Do you want to use CSI derived dihedral restraints? *}
{+ choice: true false +}
{==>} dihedralscsi_on=false;

{===== TALOS restraints =====}
{* Do you want to use TALOS derived dihedral restraints? *}
{+ choice: true false +}
{==>} dihedraltalos_on=false;

{===== topology and parameter files =====}

{* topology file *}
{==>} prot_top="topallhdg5.3.pro";

{* linkage file *}
{==>} prot_link="topallhdg5.3.pep";

{* energy parameter file *}
{==>} prot_par_1="parallhdg5.3.pro";
{==>} prot_par_2="dna-rna-allatom.param";
{==>} prot_par_3="";
{==>} prot_par_4="";
{==>} prot_par_5="";

{* type of non-bonded parameters *}
{* specify the type of non-bonded interaction *}
{+ choice: "PROLSQ" "PARMALLH6" "PARALLHDG" "OPLSX" +}
{==>} par_nonbonded="PARALLHDG";

{* Do you want to include dihedral angle energy terms? *}
{+ choice: true false +}
{==>} dihedflag=true;

{* Do you want to use floating chirality assignment (swapping)? *}
{+ choice: true false +}
{==>} swapflag=true;

{===== spectra parameters =====}

{* parameters *}
{+ table: rows=19
  "spectra names" "qshifts" "qcalib"
  "qrelax" "Tmix" "Tcorrel" "frequency"
  "qerrset" "errmod" "err0" "err1" "err2" "err3" "qexclude" "qmove" "het1window"
  "pro1window" "het2window" "pro2window"
  cols = 5 "1" "2" "3" "4" "5" +}

{==>} aspectrum_1="13C";
{==>} aspectrum_2="13Calpha";
{==>} aspectrum_3="15N";
{==>} aspectrum_4="13Cinter";
{==>} aspectrum_5="15Ninter";

{+ choice: true false +}
{==>} qshifts_1=true;
{+ choice: true false +}
{==>} qshifts_2=true;
{+ choice: true false +}
{==>} qshifts_3=true;
{+ choice: true false +}
{==>} qshifts_4=true;
{+ choice: true false +}

```

```

{==>} qshifts_5=true;

{+ choice: true false +}
{==>} qcalib_1=true;
{+ choice: true false +}
{==>} qcalib_2=true;
{+ choice: true false +}
{==>} qcalib_3=true;
{+ choice: true false +}
{==>} qcalib_4=true;
{+ choice: true false +}
{==>} qcalib_5=true;

{+ choice: true false +}
{==>} qrelax_1=true;
{+ choice: true false +}
{==>} qrelax_2=true;
{+ choice: true false +}
{==>} qrelax_3=true;
{+ choice: true false +}
{==>} qrelax_4=true;
{+ choice: true false +}
{==>} qrelax_5=true;

{==>} tmix_1=0.050;
{==>} tmix_2=0.050;
{==>} tmix_3=0.050;
{==>} tmix_4=0.060;
{==>} tmix_5=0.060;

{==>} tcorrel_1=1.5E-8;
{==>} tcorrel_2=1.5E-8;
{==>} tcorrel_3=1.5E-8;
{==>} tcorrel_4=1.5E-8;
{==>} tcorrel_5=1.5E-8;

{==>} frequency_1=800E6;
{==>} frequency_2=800E6;
{==>} frequency_3=800E6;
{==>} frequency_4=800E6;
{==>} frequency_5=800E6;

{+ choice: true false +}
{==>} qerrset_1=true;
{+ choice: true false +}
{==>} qerrset_2=true;
{+ choice: true false +}
{==>} qerrset_3=true;
{+ choice: true false +}
{==>} qerrset_4=true;
{+ choice: true false +}
{==>} qerrset_5=true;

{==>} errmod_1="DIST";
{==>} errmod_2="DIST";
{==>} errmod_3="DIST";
{==>} errmod_4="DIST";
{==>} errmod_5="DIST";

{==>} err0_1=0.0;
{==>} err0_2=0.0;
{==>} err0_3=0.0;
{==>} err0_4=0.0;
{==>} err0_5=0.0;

{==>} err1_1=0.0;
{==>} err1_2=0.0;
{==>} err1_3=0.0;
{==>} err1_4=0.0;

```

```

{====>} err1_5=0.0;

{====>} err2_1=0.125;
{====>} err2_2=0.125;
{====>} err2_3=0.125;
{====>} err2_4=0.125;
{====>} err2_5=0.125;

{====>} err3_1=0.0;
{====>} err3_2=0.0;
{====>} err3_3=0.0;
{====>} err3_4=0.0;
{====>} err3_5=0.0;

{+ choice: true false +}
{====>} qexclude_1=true;
{+ choice: true false +}
{====>} qexclude_2=true;
{+ choice: true false +}
{====>} qexclude_3=true;
{+ choice: true false +}
{====>} qexclude_4=false;          ! set to false for inter-monomer NOE
{+ choice: true false +}
{====>} qexclude_5=false;          ! set to false for inter-monomer NOE

{+ choice: true false +}
{====>} qmove_1=true;
{+ choice: true false +}
{====>} qmove_2=true;
{+ choice: true false +}
{====>} qmove_3=true;
{+ choice: true false +}
{====>} qmove_4=false;          ! set to false for inter-monomer NOE
{+ choice: true false +}
{====>} qmove_5=false;          ! set to false for inter-monomer NOE

{====>} het1window_1=0.5;
{====>} het1window_2=0.5;
{====>} het1window_3=0.5;
{====>} het1window_4=0.5;
{====>} het1window_5=0.5;

{====>} pro1window_1=0.04;
{====>} pro1window_2=0.04;
{====>} pro1window_3=0.04;
{====>} pro1window_4=0.04;
{====>} pro1window_5=0.04;

{====>} het2window_1=0.5;
{====>} het2window_2=0.5;
{====>} het2window_3=0.5;
{====>} het2window_4=0.5;
{====>} het2window_5=0.5;

{====>} pro2window_1=0.04;
{====>} pro2window_2=0.04;
{====>} pro2window_3=0.04;
{====>} pro2window_4=0.04;
{====>} pro2window_5=0.04;

{===== relaxation matrix parameters =====}
{* distance cutoff *}
{====>} relax_cutoff=5.0;
{* number of cutoff layers *}
{====>} relax_nlayers=3;
{* number of matrix doublings *}

```

```

{====>} relax_mdoube=10;

{===== restraint combination =====}
{* Last iteration during which to use restraint combination as proposed by Peter
Guentert, for unambiguous long-range NOEs, set to -1 in order to switch it off *}
{+ choice: "-1" "0" "1" "2" "3" "4" "5" "6" "7" "8" +}
{====>} combination_last_it="-1";

{===== iterations =====}
{* parameters for the 8 iterations *}

{+ table: rows=8 "structures" "keep_structures" "assign_structures"
          "ambiguous_cutoff" "max_n" "violation_tolerance" "violation_ratio"
  "window_weight"
  cols=9 "It0" "It1" "It2" "It3" "It4" "It5" "It6" "It7" "It8" +}

{====>} structures_0=20;
{====>} structures_1=20;
{====>} structures_2=20;
{====>} structures_3=20;
{====>} structures_4=20;
{====>} structures_5=20;
{====>} structures_6=20;
{====>} structures_7=20;
{====>} structures_8=50;

{====>} keepstruct_0=0;
{====>} keepstruct_1=10;
{====>} keepstruct_2=10;
{====>} keepstruct_3=10;
{====>} keepstruct_4=10;
{====>} keepstruct_5=10;
{====>} keepstruct_6=10;
{====>} keepstruct_7=10;
{====>} keepstruct_8=25;

{====>} assignstruct_0=7;
{====>} assignstruct_1=7;
{====>} assignstruct_2=7;
{====>} assignstruct_3=7;
{====>} assignstruct_4=7;
{====>} assignstruct_5=7;
{====>} assignstruct_6=7;
{====>} assignstruct_7=7;
{====>} assignstruct_8=20;

{====>} ambigcutoff_0=1.01;
{====>} ambigcutoff_1=0.9999;
{====>} ambigcutoff_2=0.999;
{====>} ambigcutoff_3=0.99;
{====>} ambigcutoff_4=0.98;
{====>} ambigcutoff_5=0.96;
{====>} ambigcutoff_6=0.93;
{====>} ambigcutoff_7=0.90;
{====>} ambigcutoff_8=0.80;

{====>} maxn_0=20;
{====>} maxn_1=20;
{====>} maxn_2=20;
{====>} maxn_3=20;
{====>} maxn_4=20;
{====>} maxn_5=20;
{====>} maxn_6=20;

```

```

{==>} maxn_7=20;
{==>} maxn_8=20;

{==>} violtoler_0=1000.0;
{==>} violtoler_1=1000.0;
{==>} violtoler_2=1.0;
{==>} violtoler_3=0.5;
{==>} violtoler_4=0.1;
{==>} violtoler_5=1.0;
{==>} violtoler_6=0.1;
{==>} violtoler_7=0.1;
{==>} violtoler_8=0.1;

{==>} violratio_0=0.5;
{==>} violratio_1=0.5;
{==>} violratio_2=0.5;
{==>} violratio_3=0.5;
{==>} violratio_4=0.5;
{==>} violratio_5=0.5;
{==>} violratio_6=0.5;
{==>} violratio_7=0.5;
{==>} violratio_8=0.5;

{==>} window_weight_0=1.0;
{==>} window_weight_1=1.0;
{==>} window_weight_2=1.0;
{==>} window_weight_3=1.0;
{==>} window_weight_4=1.0;
{==>} window_weight_5=1.0;
{==>} window_weight_6=1.0;
{==>} window_weight_7=1.0;
{==>} window_weight_8=1.0;

{===== parallel jobs =====}
{* How many nodes do you want to use in parallel? *}
{* leave unused fields blank, make sure that the queues are actually running *}
{+ table: rows=10 "1" "2" "3" "4" "5" "6" "7" "8" "9" "10"
  cols=3 "queue command" "cns executable" "number of jobs" +}

{==>} queue_1="csh";
{==>} cns_exe_1="/usr/site/cns_solve_aria/intel-i686-linux/bin/cns";
{==>} cpunumber_1=1;

{==>} queue_2="";
{==>} cns_exe_2="";
{==>} cpunumber_2=0;

{==>} queue_3="";
{==>} cns_exe_3="";
{==>} cpunumber_3=0;

{==>} queue_4="";
{==>} cns_exe_4="";
{==>} cpunumber_4=0;

{==>} queue_5="";
{==>} cns_exe_5="";
{==>} cpunumber_5=0;

{==>} queue_6="";
{==>} cns_exe_6="";
{==>} cpunumber_6=0;

{==>} queue_7="";
{==>} cns_exe_7="";
{==>} cpunumber_7=0;

```

```

{==>} queue_8="";
{==>} cns_exe_8="";
{==>} cpunumber_8=0;

{==>} queue_9="";
{==>} cns_exe_9="";
{==>} cpunumber_9=0;

{==>} queue_10="";
{==>} cns_exe_10="";
{==>} cpunumber_10=0;

{===== SA protocol =====}
{* type of molecular dynamics *}
{+ choice: "torsion" "Cartesian" +}
{==>} md_type="torsion";

{* initial seed for random number generator *}
{* change to get different initial velocities *}
{==>} iniseed=89764443;

{* initial temperature for TAD *}
{==>} tadinit_t=10000;
{* initial temperature for TAD and Cartesian dynamics *}
{==>} carinit_t=2000;

{* final temperature after first cooling step *}
{==>} finall_t=1000;

{* finale temperature after second cooling step *}
{==>} final2_t=50;

{* Cartesian time step *}
{==>} timestep=0.003;
{* factor for timestep and number of steps in TAD *}
{==>} tadfactor=9;

{* initial number of MD steps *}
{==>} initiosteps=10000;

{* number of MD steps for refinement *}
{==>} refinesteps=4000;

{* number of MD steps during first cooling stage *}
{==>} cool1_steps=5000;

{* number of MD steps during second cooling stage *}
{==>} cool2_steps=4000;

{* first iteration for Cartesian refinement (after TAD) *}
{==>} cart_firstit=0;

{* Sort structures accordingly to total energy or sum of restraints energies? *}
{+ choice: "totener" "restener" +}
{==>} filesort="totener";

{===== final water refinement =====}
{* Do you want water refinement for the last iteration? *}
{+ choice: "yes" "no" +}
{==>} firstwater="yes";

{* Which solvent do you want to use? *}
{+ choice: "water" "dmso" +}
{==>} solvent="water";

```



```

{* number of structures for the water refinement *}
{* refine the n best structures regarding energy *}
{==>} waterrefine=10;

{===== proccheck, whatcheck and prosa analysis of the last iteration
=====}
{* Make sure that your PRODIR and PROSA_BASE system variables are set correctly. Leave
fields empty if you don't want to perform these checks 1. proccheck executable *}
{==>} proccheck_exe="/divsite/opt.BIOC/proccheck-3.5.4/irix6/proccheck/proccheck.scr";
{* 2. proccheck_comp.scr script *}
{==>} proccheckcomp_exe="/divsite/opt.BIOC/proccheck-
3.5.4/irix6/proccheck/proccheck_comp.scr";
{* 3. whatif executable *}
{==>} whatif_exe="";
{* 4. prosaII executable *}
{==>} prosa_exe="";
{* number of pdb files for analysis *}
{==>} how_many_pdb="20";

{=====}
{      things below this line do not normally need to be changed      }
{=====}

) {- end block parameter definition -}

!set countspec (important for the correct numbering of the spectra):
if ($spectrum =&aspectrum_1) then
  evaluate ($countspec=1)
elseif ($spectrum =&aspectrum_2) then
  evaluate ($countspec=2)
elseif ($spectrum =&aspectrum_3) then
  evaluate ($countspec=3)
elseif ($spectrum =&aspectrum_4) then
  evaluate ($countspec=4)
elseif ($spectrum =&aspectrum_5) then
  evaluate ($countspec=5)
end if

!extensions:
evaluate (&spectra.restraintfile = $spectrum + ".tbl")
evaluate (&spectra.listfile = $spectrum + ".list")

!for global parameters (local variables (suffix ) => global variables):
evaluate (&saprotocol.iniseed=&iniseed)
evaluate (&saprotocol.carinit_t=&carinit_t)
evaluate (&saprotocol.tadinit_t=&tadinit_t)
evaluate (&saprotocol.finall_t=&finall_t)
evaluate (&saprotocol.final2_t=&final2_t)
evaluate (&saprotocol.tempstep=50)
evaluate (&saprotocol.timestep=&timestep)
evaluate (&saprotocol.tadfactor=&tadfactor)
evaluate (&saprotocol.initiosteps=&initiosteps)
evaluate (&saprotocol.refinesteps=&refinesteps)
evaluate (&saprotocol.cool1_steps=&cool1_steps)
evaluate (&saprotocol.cool2_steps=&cool2_steps)
evaluate (&saprotocol.cart_firstit=&cart_firstit)
evaluate (&saprotocol.fbeta=20)
evaluate (&saprotocol.mass=100)
evaluate (&filenames.fileroot=&fileroot)
evaluate (&filenames.reference=&reference)
evaluate (&filenames.structure =&structure)
evaluate (&filenames.structure_1=&structure)
evaluate (&filenames.structure_2=&structure_2)
evaluate (&filenames.structure_3=&structure_3)
evaluate (&filenames.structure_4=&structure_4)
evaluate (&filenames.structure_5=&structure_5)

evaluate (&filenames.template=&fileroot + "_template.pdb")

```

```

evaluate (&iterations.ini_count      =1)
evaluate (&iterations.structures     =&structures_$iteration)
evaluate (&iterations.keepstruct     =&keepstruct_$iteration)
evaluate (&iterations.ambigcutoff    =&ambigcutoff_$iteration)
evaluate (&iterations.maxn           =&maxn_$iteration)
evaluate (&iterations.violtoler      =&violtoler_$iteration)
evaluate (&iterations.violratio      =&violratio_$iteration)
evaluate (&iterations.assignstruct  =&assignstruct_$iteration)
evaluate (&iterations.window_weight=&window_weight_$iteration)

!set nspectra (important for merging protocol):
evaluate (&spectra.nspectra=0)
if (&aspectrum_1 ne "") then
  evaluate (&spectra.nspectra=1)
end if
if (&aspectrum_2 ne "") then
  evaluate (&spectra.nspectra=2)
end if
if (&aspectrum_3 ne "") then
  evaluate (&spectra.nspectra=3)
end if
if (&aspectrum_4 ne "") then
  evaluate (&spectra.nspectra=4)
end if
if (&aspectrum_5 ne "") then
  evaluate (&spectra.nspectra=5)
end if

evaluate (&spectra.aspectrum=&aspectrum_$countspec)

evaluate (&spectra.namelist_1=&aspectrum_1)
evaluate (&spectra.namelist_2=&aspectrum_2)
evaluate (&spectra.namelist_3=&aspectrum_3)
evaluate (&spectra.namelist_4=&aspectrum_4)
evaluate (&spectra.namelist_5=&aspectrum_5)

evaluate (&spectra.peaklist=&aspectrum_$countspec + ".tbl")
evaluate (&spectra.qshifts=&qshifts_$countspec)
evaluate (&spectra.shifts=&aspectrum_$countspec + ".ppm")

evaluate (&spectra.qcalib=&qcalib_$countspec)
evaluate (&spectra.qrelax=&qrelax_$countspec)
evaluate (&spectra.tmix=&tmix_$countspec)
evaluate (&spectra.tcorrel=&tcorrel_$countspec)
evaluate (&spectra.frequency=&frequency_$countspec)
evaluate (&spectra.solvent=&solvent_$countspec)
evaluate (&spectra.qerrset=&qerrset_$countspec)
evaluate (&spectra.errmod=&errmod_$countspec)
evaluate (&spectra.err0=&err0_$countspec)
evaluate (&spectra.err1=&err1_$countspec)
evaluate (&spectra.err2=&err2_$countspec)
evaluate (&spectra.err3=&err3_$countspec)
evaluate (&spectra.qexclude=&qexclude_$countspec)
evaluate (&spectra.qmove=&qmove_$countspec)
evaluate (&spectra.het1window=&het1window_$countspec)
evaluate (&spectra.pro1window=&pro1window_$countspec)
evaluate (&spectra.het2window=&het2window_$countspec)
evaluate (&spectra.pro2window=&pro2window_$countspec)

!topology and parameters, sequence file, template file, ss-bridges, HISD, HISE:
evaluate (&toppar.prot_top=&prot_top )
evaluate (&toppar.prot_link=&prot_link )
evaluate (&toppar.prot_par  =&prot_par_1 )
evaluate (&toppar.prot_par_1=&prot_par_1 )
evaluate (&toppar.prot_par_2=&prot_par_2 )
evaluate (&toppar.prot_par_3=&prot_par_3 )
evaluate (&toppar.prot_par_4=&prot_par_4 )
evaluate (&toppar.prot_par_5=&prot_par_5 )

```

```

evaluate (&toppar.par_nonbonded=&par_nonbonded)
evaluate (&toppar.pdb_or_sequence=&pdb_or_sequence )
evaluate (&toppar.prot_coor_1=&prot_coor_1)
evaluate (&toppar.prot_segid_1=&prot_segid_1)
evaluate (&toppar.pdb_or_sequence_2=&pdb_or_sequence_2 ) ! ESS 02/01/2007
evaluate (&toppar.prot_coor_2=&prot_coor_2) ! ESS 02/01/2007
evaluate (&toppar.prot_segid_2=&prot_segid_2) ! ESS 02/01/2007

```

```

evaluate (&toppar.ss_bridge=&ss_bridge)
evaluate (&toppar.ss_i_resid_1=&ss_i_resid_1)
evaluate (&toppar.ss_i_segid_1=&ss_i_segid_1)
evaluate (&toppar.ss_j_resid_1=&ss_j_resid_1)
evaluate (&toppar.ss_j_segid_1=&ss_j_segid_1)
evaluate (&toppar.ss_i_resid_2=&ss_i_resid_2)
evaluate (&toppar.ss_i_segid_2=&ss_i_segid_2)
evaluate (&toppar.ss_j_resid_2=&ss_j_resid_2)
evaluate (&toppar.ss_j_segid_2=&ss_j_segid_2)
evaluate (&toppar.ss_i_resid_3=&ss_i_resid_3)
evaluate (&toppar.ss_i_segid_3=&ss_i_segid_3)
evaluate (&toppar.ss_j_resid_3=&ss_j_resid_3)
evaluate (&toppar.ss_j_segid_3=&ss_j_segid_3)
evaluate (&toppar.ss_i_resid_4=&ss_i_resid_4)
evaluate (&toppar.ss_i_segid_4=&ss_i_segid_4)
evaluate (&toppar.ss_j_resid_4=&ss_j_resid_4)
evaluate (&toppar.ss_j_segid_4=&ss_j_segid_4)
evaluate (&toppar.ss_i_resid_5=&ss_i_resid_5)
evaluate (&toppar.ss_i_segid_5=&ss_i_segid_5)
evaluate (&toppar.ss_j_resid_5=&ss_j_resid_5)
evaluate (&toppar.ss_j_segid_5=&ss_j_segid_5)
evaluate (&toppar.ss_ambigunambig=&ss_ambigunambig)

```

```

evaluate (&toppar.hisd_resid_1=&hisd_resid_1)
evaluate (&toppar.hisd_segid_1=&hisd_segid_1)
evaluate (&toppar.hisd_resid_2=&hisd_resid_2)
evaluate (&toppar.hisd_segid_2=&hisd_segid_2)
evaluate (&toppar.hisd_resid_3=&hisd_resid_3)
evaluate (&toppar.hisd_segid_3=&hisd_segid_3)
evaluate (&toppar.hisd_resid_4=&hisd_resid_4)
evaluate (&toppar.hisd_segid_4=&hisd_segid_4)
evaluate (&toppar.hisd_resid_5=&hisd_resid_5)
evaluate (&toppar.hisd_segid_5=&hisd_segid_5)
evaluate (&toppar.hisd_resid_6=&hisd_resid_6)
evaluate (&toppar.hisd_segid_6=&hisd_segid_6)
evaluate (&toppar.hisd_resid_7=&hisd_resid_7)
evaluate (&toppar.hisd_segid_7=&hisd_segid_7)
evaluate (&toppar.hisd_resid_8=&hisd_resid_8)
evaluate (&toppar.hisd_segid_8=&hisd_segid_8)
evaluate (&toppar.hisd_resid_9=&hisd_resid_9)
evaluate (&toppar.hisd_segid_9=&hisd_segid_9)
evaluate (&toppar.hisd_resid_10=&hisd_resid_10)
evaluate (&toppar.hisd_segid_10=&hisd_segid_10)
evaluate (&toppar.hise_resid_1=&hise_resid_1)
evaluate (&toppar.hise_segid_1=&hise_segid_1)
evaluate (&toppar.hise_resid_2=&hise_resid_2)
evaluate (&toppar.hise_segid_2=&hise_segid_2)
evaluate (&toppar.hise_resid_3=&hise_resid_3)
evaluate (&toppar.hise_segid_3=&hise_segid_3)
evaluate (&toppar.hise_resid_4=&hise_resid_4)
evaluate (&toppar.hise_segid_4=&hise_segid_4)
evaluate (&toppar.hise_resid_5=&hise_resid_5)
evaluate (&toppar.hise_segid_5=&hise_segid_5)
evaluate (&toppar.hise_resid_6=&hise_resid_6)
evaluate (&toppar.hise_segid_6=&hise_segid_6)
evaluate (&toppar.hise_resid_7=&hise_resid_7)
evaluate (&toppar.hise_segid_7=&hise_segid_7)
evaluate (&toppar.hise_resid_8=&hise_resid_8)
evaluate (&toppar.hise_segid_8=&hise_segid_8)
evaluate (&toppar.hise_resid_9=&hise_resid_9)
evaluate (&toppar.hise_segid_9=&hise_segid_9)

```

```

evaluate (&toppar.hise_resid_10=&hise_resid_10)
evaluate (&toppar.hise_segid_10=&hise_segid_10)

evaluate (&toppar.xplortodiana=&xplortodiana)
evaluate (&toppar.float=&swapflag)

!Dihedrals, Jcouplings, Residual dipolar couplings, Hbonds, Analysis:
evaluate (&Data.flags.bonds = "TRUE")
evaluate (&Data.flags.angle = "TRUE")
evaluate (&Data.flags.impro = "TRUE")
evaluate (&Data.flags.dihed = &dihedflag)
evaluate (&Data.flags.vdw = "TRUE")
evaluate (&Data.flags.elec = "FALSE")

evaluate (&Data.flags.noe = "TRUE")
evaluate (&Data.flags.cdih = &dihedrals_on)
evaluate (&Data.cdih.on = &dihedrals_on)

evaluate (&Data.flags.coup = "FALSE")
evaluate (&Data.flags.vean = "FALSE")
evaluate (&Data.flags.sani = "FALSE")
evaluate (&Data.flags.dani = "FALSE")

evaluate (&Data.flags.plan = "FALSE")
evaluate (&Data.flags.ncs = "FALSE")

evaluate (&data.unamb_firstit=&unamb_firstit)
evaluate (&data.unamb_hot=&unamb_hot)
evaluate (&data.unamb_cool1_ini=&unamb_cool1_ini)
evaluate (&data.unamb_cool1_fin=&unamb_cool1_fin)
evaluate (&data.unamb_cool2=&unamb_cool2)

evaluate (&data.amb_firstit=&amb_firstit)
evaluate (&data.amb_hot=&amb_hot)
evaluate (&data.amb_cool1_ini=&amb_cool1_ini)
evaluate (&data.amb_cool1_fin=&amb_cool1_fin)
evaluate (&data.amb_cool2=&amb_cool2)

evaluate (&data.hbond_firstit=&hbond_firstit)
evaluate (&data.hbond_hot=&hbond_hot)
evaluate (&data.hbond_cool1_ini=&hbond_cool1_ini)
evaluate (&data.hbond_cool1_fin=&hbond_cool1_fin)
evaluate (&data.hbond_cool2=&hbond_cool2)

evaluate (&data.mrswi_hot=&mrswi_hot)
evaluate (&data.mrswi_cool1=&mrswi_cool1)
evaluate (&data.mrswi_cool2=&mrswi_cool2)

evaluate (&data.rswi_hot=&rswi_hot)
evaluate (&data.rswi_cool1=&rswi_cool1)
evaluate (&data.rswi_cool2=&rswi_cool2)

evaluate (&data.masy_hot=&masy_hot)
evaluate (&data.masy_cool1=&masy_cool1)
evaluate (&data.masy_cool2=&masy_cool2)

evaluate (&data.asy_hot=&asy_hot)
evaluate (&data.asy_cool1=&asy_cool1)
evaluate (&data.asy_cool2=&asy_cool2)

evaluate (&data.dihedrals.on=&dihedrals_on)
evaluate (&data.dihedrals_hot=&dihedrals_hot)
evaluate (&data.dihedrals_cool1=&dihedrals_cool1)
evaluate (&data.dihedrals_cool2=&dihedrals_cool2)
evaluate (&data.hbonds_on=&hbonds_on)
evaluate (&data.hbondscsi_on=&hbondscsi_on)
evaluate (&data.dihedralscsi.on=&dihedralscsi_on)
evaluate (&data.cdihcsi.on=&dihedralscsi_on)
evaluate (&data.dihedralstalos.on=&dihedralstalos_on)

```

```
evaluate (&data.cdihtalos.on=&dihedralstalos_on)
```

```
evaluate (&data.c1_on=&c1_on)
evaluate (&data.c1_karplusa=&c1_karplusa)
evaluate (&data.c1_karplusb=&c1_karplusb)
evaluate (&data.c1_karplusc=&c1_karplusc)
evaluate (&data.c1_karplused=&c1_karplused)
evaluate (&data.c1_hot=&c1_hot)
evaluate (&data.c1_cool1=&c1_cool1)
evaluate (&data.c1_cool2=&c1_cool2)
evaluate (&data.c2_on=&c2_on)
evaluate (&data.c2_karplusa=&c2_karplusa)
evaluate (&data.c2_karplusb=&c2_karplusb)
evaluate (&data.c2_karplusc=&c2_karplusc)
evaluate (&data.c2_karplused=&c2_karplused)
evaluate (&data.c2_hot=&c2_hot)
evaluate (&data.c2_cool1=&c2_cool1)
evaluate (&data.c2_cool2=&c2_cool2)
evaluate (&data.c3_on=&c3_on)
evaluate (&data.c3_karplusa=&c3_karplusa)
evaluate (&data.c3_karplusb=&c3_karplusb)
evaluate (&data.c3_karplusc=&c3_karplusc)
evaluate (&data.c3_karplused=&c3_karplused)
evaluate (&data.c3_hot=&c3_hot)
evaluate (&data.c3_cool1=&c3_cool1)
evaluate (&data.c3_cool2=&c3_cool2)
evaluate (&data.c4_on=&c4_on)
evaluate (&data.c4_karplusa=&c4_karplusa)
evaluate (&data.c4_karplusb=&c4_karplusb)
evaluate (&data.c4_karplusc=&c4_karplusc)
evaluate (&data.c4_karplused=&c4_karplused)
evaluate (&data.c4_hot=&c4_hot)
evaluate (&data.c4_cool1=&c4_cool1)
evaluate (&data.c4_cool2=&c4_cool2)
evaluate (&data.c5_on=&c5_on)
evaluate (&data.c5_karplusa=&c5_karplusa)
evaluate (&data.c5_karplusb=&c5_karplusb)
evaluate (&data.c5_karplusc=&c5_karplusc)
evaluate (&data.c5_karplused=&c5_karplused)
evaluate (&data.c5_hot=&c5_hot)
evaluate (&data.c5_cool1=&c5_cool1)
evaluate (&data.c5_cool2=&c5_cool2)
```

```
evaluate (&data.rdc1_choice=&rdc1_choice)
evaluate (&data.rdc1_firstIt=&rdc1_firstIt)
evaluate (&data.rdc1_hot=&rdc1_hot)
evaluate (&data.rdc1_cool1=&rdc1_cool1)
evaluate (&data.rdc1_cool2=&rdc1_cool2)
evaluate (&data.rdc1_r=&rdc1_r)
evaluate (&data.rdc1_d=&rdc1_d)
```

```
evaluate (&data.rdc2_choice=&rdc2_choice)
evaluate (&data.rdc2_firstIt=&rdc2_firstIt)
evaluate (&data.rdc2_hot=&rdc2_hot)
evaluate (&data.rdc2_cool1=&rdc2_cool1)
evaluate (&data.rdc2_cool2=&rdc2_cool2)
evaluate (&data.rdc2_r=&rdc2_r)
evaluate (&data.rdc2_d=&rdc2_d)
```

```
evaluate (&data.rdc3_choice=&rdc3_choice)
evaluate (&data.rdc3_firstIt=&rdc3_firstIt)
evaluate (&data.rdc3_hot=&rdc3_hot)
evaluate (&data.rdc3_cool1=&rdc3_cool1)
evaluate (&data.rdc3_cool2=&rdc3_cool2)
evaluate (&data.rdc3_r=&rdc3_r)
evaluate (&data.rdc3_d=&rdc3_d)
```

```
evaluate (&data.rdc4_choice=&rdc4_choice)
evaluate (&data.rdc4_firstIt=&rdc4_firstIt)
```

```

evaluate (&data.rdc4_hot=&rdc4_hot)
evaluate (&data.rdc4_cool1=&rdc4_cool1)
evaluate (&data.rdc4_cool2=&rdc4_cool2)
evaluate (&data.rdc4_r=&rdc4_r)
evaluate (&data.rdc4_d=&rdc4_d)

evaluate (&data.rdc5_choice=&rdc5_choice)
evaluate (&data.rdc5_firstIt=&rdc5_firstIt)
evaluate (&data.rdc5_hot=&rdc5_hot)
evaluate (&data.rdc5_cool1=&rdc5_cool1)
evaluate (&data.rdc5_cool2=&rdc5_cool2)
evaluate (&data.rdc5_r=&rdc5_r)
evaluate (&data.rdc5_d=&rdc5_d)

!VEAN statement:
evaluate (&data.ini_bor_hot_1=&ini_bor_hot_1)
evaluate (&data.ini_bor_cool1_1=&ini_bor_cool1_1)
evaluate (&data.ini_bor_cool2_1=&ini_bor_cool2_1)
evaluate (&data.ini_cen_hot_1=&ini_cen_hot_1)
evaluate (&data.ini_cen_cool1_1=&ini_cen_cool1_1)
evaluate (&data.ini_cen_cool2_1=&ini_cen_cool2_1)
evaluate (&data.fin_bor_hot_1=&fin_bor_hot_1)
evaluate (&data.fin_bor_cool1_1=&fin_bor_cool1_1)
evaluate (&data.fin_bor_cool2_1=&fin_bor_cool2_1)
evaluate (&data.fin_cen_hot_1=&fin_cen_hot_1)
evaluate (&data.fin_cen_cool1_1=&fin_cen_cool1_1)
evaluate (&data.fin_cen_cool2_1=&fin_cen_cool2_1)

evaluate (&data.ini_bor_hot_2=&ini_bor_hot_2)
evaluate (&data.ini_bor_cool1_2=&ini_bor_cool1_2)
evaluate (&data.ini_bor_cool2_2=&ini_bor_cool2_2)
evaluate (&data.ini_cen_hot_2=&ini_cen_hot_2)
evaluate (&data.ini_cen_cool1_2=&ini_cen_cool1_2)
evaluate (&data.ini_cen_cool2_2=&ini_cen_cool2_2)
evaluate (&data.fin_bor_hot_2=&fin_bor_hot_2)
evaluate (&data.fin_bor_cool1_2=&fin_bor_cool1_2)
evaluate (&data.fin_bor_cool2_2=&fin_bor_cool2_2)
evaluate (&data.fin_cen_hot_2=&fin_cen_hot_2)
evaluate (&data.fin_cen_cool1_2=&fin_cen_cool1_2)
evaluate (&data.fin_cen_cool2_2=&fin_cen_cool2_2)

evaluate (&data.ini_bor_hot_3=&ini_bor_hot_3)
evaluate (&data.ini_bor_cool1_3=&ini_bor_cool1_3)
evaluate (&data.ini_bor_cool2_3=&ini_bor_cool2_3)
evaluate (&data.ini_cen_hot_3=&ini_cen_hot_3)
evaluate (&data.ini_cen_cool1_3=&ini_cen_cool1_3)
evaluate (&data.ini_cen_cool2_3=&ini_cen_cool2_3)
evaluate (&data.fin_bor_hot_3=&fin_bor_hot_3)
evaluate (&data.fin_bor_cool1_3=&fin_bor_cool1_3)
evaluate (&data.fin_bor_cool2_3=&fin_bor_cool2_3)
evaluate (&data.fin_cen_hot_3=&fin_cen_hot_3)
evaluate (&data.fin_cen_cool1_3=&fin_cen_cool1_3)
evaluate (&data.fin_cen_cool2_3=&fin_cen_cool2_3)

evaluate (&data.ini_bor_hot_4=&ini_bor_hot_4)
evaluate (&data.ini_bor_cool1_4=&ini_bor_cool1_4)
evaluate (&data.ini_bor_cool2_4=&ini_bor_cool2_4)
evaluate (&data.ini_cen_hot_4=&ini_cen_hot_4)
evaluate (&data.ini_cen_cool1_4=&ini_cen_cool1_4)
evaluate (&data.ini_cen_cool2_4=&ini_cen_cool2_4)
evaluate (&data.fin_bor_hot_4=&fin_bor_hot_4)
evaluate (&data.fin_bor_cool1_4=&fin_bor_cool1_4)
evaluate (&data.fin_bor_cool2_4=&fin_bor_cool2_4)
evaluate (&data.fin_cen_hot_4=&fin_cen_hot_4)
evaluate (&data.fin_cen_cool1_4=&fin_cen_cool1_4)
evaluate (&data.fin_cen_cool2_4=&fin_cen_cool2_4)

evaluate (&data.ini_bor_hot_5=&ini_bor_hot_5)
evaluate (&data.ini_bor_cool1_5=&ini_bor_cool1_5)

```

```

evaluate (&data.ini_bor_cool2_5=&ini_bor_cool2_5)
evaluate (&data.ini_cen_hot_5=&ini_cen_hot_5)
evaluate (&data.ini_cen_cool1_5=&ini_cen_cool1_5)
evaluate (&data.ini_cen_cool2_5=&ini_cen_cool2_5)
evaluate (&data.fin_bor_hot_5=&fin_bor_hot_5)
evaluate (&data.fin_bor_cool1_5=&fin_bor_cool1_5)
evaluate (&data.fin_bor_cool2_5=&fin_bor_cool2_5)
evaluate (&data.fin_cen_hot_5=&fin_cen_hot_5)
evaluate (&data.fin_cen_cool1_5=&fin_cen_cool1_5)
evaluate (&data.fin_cen_cool2_5=&fin_cen_cool2_5)

!relaxation matrix refinement:
evaluate (&relax.nlayers=&relax_nlayers)
evaluate (&relax.cutoff=&relax_cutoff)
evaluate (&relax.mdoube=&relax_mdoube)

!water refinement
evaluate (&refine.firstwater=&firstwater)
evaluate (&refine.waterrefine=&waterrefine)
evaluate (&refine.solvent=&solvent)
evaluate (&refine.pmrefine_on=&pmrefine_on)

!free R-factor:
evaluate (&analysis.on=&rfree_on)
evaluate (&analysis.oneset=&rfree_oneset)
evaluate (&analysis.method=&rfree_method)
evaluate (&analysis.percent=&rfree_percent)
evaluate (&analysis.signal1=&rfree_signal1)
evaluate (&analysis.signal2=&rfree_signal2)
evaluate (&analysis.putative1=&rfree_putative1)
evaluate (&analysis.putative2=&rfree_putative2)
evaluate (&analysis.noise1=&rfree_noise1)
evaluate (&analysis.noise2=&rfree_noise2)

! evaluate (&analysis.fitselect_ordered=&fitselect_ordered)
! evaluate (&analysis.fitselect_all=&fitselect_all)

!for the non-bonded parameters (the section was taken out of
!parallhdg5.0.pro and parallhdg5.1.pro, so be careful!):
if (&toppar.par_nonbonded eq "PROLSQ") then
  evaluate (&toppar.repel_radius = 1.0)
  evaluate (&toppar.repel_rcons = 20)
  evaluate (&toppar.repel_rexpo = 4)
  evaluate (&toppar.repel_irexp = 1)
elseif (&toppar.par_nonbonded eq "PARMALLH6") then
  evaluate (&toppar.repel_radius = 0.8)
  evaluate (&toppar.repel_rcons = 5.0)
  evaluate (&toppar.repel_rexpo = 2)
  evaluate (&toppar.repel_irexp = 2)
elseif (&toppar.par_nonbonded eq "OPLSX") then
  evaluate (&toppar.repel_radius = 0.0)
else
  {...now the standard PARALLHDG parameters}
  evaluate (&toppar.repel_radius = 0.78)
  evaluate (&toppar.repel_rcons = 5.0)
  evaluate (&toppar.repel_rexpo = 2)
  evaluate (&toppar.repel_irexp = 2)
end if

!Procheck analysis:
evaluate (&analysis.procheckdir=&procheckdir)

set message on echo on end

```

```

!$Revision: 2.2 $
!$Date: 2002/07/23 16:19:27 $
!$RCSfile: generate.inp,v $

!      *****
!      * Authors and copyright:      *
!      * Michael Nilges, Jens Linge, EMBL *
!      * No warranty implied or expressed *
!      * All rights reserved          *
!      *****

@RUN:protocols/initialize.cns
(spectrum=$spectrum;iteration=$iteration;)

!!!@NEWIT:iteration.cns (iteration=$iteration;)

!!! @SPECTRUM:spectrum.cns (spectrum=$spectrum;)

@RUN:run.cns(
spectrum  =$spectrum;
iteration  =$iteration;
filenames =$filenames;
spectra   =$spectra;
data      =$data;
iterations=$iterations;
saprotocol=$saprotocol;
refine    =$refine;
relax     =$relax;
toppar    =$toppar;
analysis  =$analysis;
)

evaluate($prot_top_file= "RUN:toppar/" + $Toppar.prot_top)
topology
  @$prot_top_file
end

evaluate($prot_link_file= "RUN:toppar/" + $Toppar.prot_link)
evaluate($sequence_file= "RUN:data/sequence/" + $Toppar.prot_coor_1)
!for the case, that you don't want a SEGID:
if ($Toppar.prot_segid_1="") then
  evaluate($Toppar.prot_segid_1=" ")
end if

segment
  name=$Toppar.prot_segid_1
  chain
    @$prot_link_file
    if ($Toppar.pdb_or_sequence="sequence") then
      sequence @$sequence_file END      !for sequence file
    else
      coor @$sequence_file              !for PDB file
    end if
  end
end

! cis peptide patch necessary for werner
! this should be in run.cns
! patch cipp reference=NIL=(resid 34) end

!for the DIANA-XEASY-nomenclature:
if ($Toppar.xplortodiana=true) then
  @RUN:protocols/xplortodiana3.inp
end if

```



```

!for the disulfide bridge patches:
evaluate($di_counter=1)
while ($di_counter le 20) loop disulfide
  if ($Toppar.ss_bridge ge $di_counter) then
    patch disu reference=1=(resid $Toppar.ss_i_resid_$di_counter and
                          segid $Toppar.ss_i_segid_$di_counter )
    reference=2=(resid $Toppar.ss_j_resid_$di_counter and
                segid $Toppar.ss_j_segid_$di_counter ) end
  end if
  evaluate($di_counter=$di_counter + 1)
end loop disulfide

!for the histidine patches:
evaluate($hisd_counter=1)
while ($hisd_counter le 10) loop hisd
  if ($Toppar.hisd_resid_$hisd_counter > 0) then
    patch hisd reference=nil=(resid $Toppar.hisd_resid_$hisd_counter and
                             segid $Toppar.hisd_segid_$hisd_counter) end
  end if
  evaluate($hisd_counter=$hisd_counter + 1)
end loop hisd

evaluate($hise_counter=1)
while ($hise_counter le 10) loop hise
  if ($Toppar.hise_resid_$hise_counter > 0) then
    patch hise reference=nil=(resid $Toppar.hise_resid_$hise_counter and
                             segid $Toppar.hise_segid_$hise_counter) end
  end if
  evaluate($hise_counter=$hise_counter + 1)
end loop hise

!! This part is added by ESS 01/25/2007
!! This will allow to input two sequences, but ARIA will generate only one .psf file

evaluate($prot_link_file= "RUN:toppar/" + $Toppar.prot_link)
! edited by ESS 01/25/07
evaluate($sequence_file_2= "RUN:data/sequence/" + $Toppar.prot_coor_2)
!for the case, that you don't want a SEGID:
if ($Toppar.prot_segid_2="") then
  evaluate($Toppar.prot_segid_2="")
end if

segment
  name=$Toppar.prot_segid_2
  chain
    @$prot_link_file
    if ($Toppar.pdb_or_sequence_2="sequence") then
      sequence @@sequence_file END
    else
      coor @@sequence_file
    end if
  end
end

! cis peptide patch necessary for werner
! this should be in run.cns
! patch cipp reference=NIL=(resid 34) end

!for the DIANA-XEASY-nomenclature:
if ($Toppar.xplortodiana=true) then
  @@RUN:protocols/xplortodiana3.inp
end if

!for the disulfide bridge patches:
evaluate($di_counter=1)

```

```

while ($di_counter le 20) loop disulfide
  if ($Toppar.ss_bridge ge $di_counter) then
    patch disu reference=1=(resid $Toppar.ss_i_resid $di_counter and
                           segid $Toppar.ss_i_segid $di_counter )
    reference=2=(resid $Toppar.ss_j_resid $di_counter and
                 segid $Toppar.ss_j_segid $di_counter ) end
  end if
  evaluate($di_counter=$di_counter + 1)
end loop disulfide

!for the histidine patches:
evaluate($hisd_counter=1)
while ($hisd_counter le 10) loop hisd
  if ($Toppar.hisd_resid $hisd_counter > 0) then
    patch hisd reference=nil=(resid $Toppar.hisd_resid $hisd_counter and
                              segid $Toppar.hisd_segid $hisd_counter) end
  end if
  evaluate($hisd_counter=$hisd_counter + 1)
end loop hisd

evaluate($hise_counter=1)
while ($hise_counter le 10) loop hise
  if ($Toppar.hise_resid $hise_counter > 0) then
    patch hise reference=nil=(resid $Toppar.hise_resid $hise_counter and
                              segid $Toppar.hise_segid $hise_counter) end
  end if
  evaluate($hise_counter=$hise_counter + 1)
end loop hise

!! This is the end of part added by ESS 01/25/2007

!write psf file:
evaluate ($structure_file = "RUN:begin/" + $FileNames.structure)
write structure output=$structure_file end

!make table with equivalent protons and swapped protons:
set display RUN:begin/methyls.tbl end
@RUN:protocols/define_methyls_ini.cns
close RUN:begin/methyls.tbl end
@RUN:protocols/define_methyls_all.cns

set display RUN:begin/setup_swap_list.tbl end
if ($Toppar.float eq true) then
  @RUN:protocols/setup_swap_init.cns
else
  display do (store1 = 0) (all)
end if

stop

```

```
! newsymmetry.cns
! Requires the use of the version of refine.inp modified by Brad Jordan to apply symmetry
! in calculation

module (Toppar)
ncs
  restraints initialise
  group
    equivalence (segid "AAAA" and name CA and
      (resid 168
        or resid 169
        or resid 170
        or resid 171
        or resid 172
        or resid 173
        or resid 174
        or resid 175
        or resid 176
        or resid 177
        or resid 178
        or resid 179
        or resid 180
        or resid 181
        or resid 182
        or resid 183
        or resid 184
        or resid 185))
    equivalence (segid "BBBB" and name CA and
      (resid 168
        or resid 169
        or resid 170
        or resid 171
        or resid 172
        or resid 173
        or resid 174
        or resid 175
        or resid 176
        or resid 177
        or resid 178
        or resid 179
        or resid 180
        or resid 181
        or resid 182
        or resid 183
        or resid 184
        or resid 185))
    weight 400 sigb 1.0
  end
end

flags include ncs end
```

```

!$Revision: 2.1 $
!$Date: 2001/06/18 12:28:21 $
!$RCSfile: stereoassign.cns,v $

! Chemical shifts assignments to be swap/reversed:
!       HB's of: Leu166, Ser168, Leu169, Leu171, Leu174, Leu175, Leu179
!       CD's of: Leu170, Val164
!
! Chemical shifts assigned as is:
!       HB's of: Leu170, Ser172, His173
!       CD's of: Leu166, Leu169, Leu171, Leu174, Leu175, Leu179

!!! assignments swapped/reversed
!!! methylene protons
for $id in id
(
{====>} (name cb and segid "AAAA" and resid 169)
{====>}or (name cb and segid "BBBB" and resid 169)
{====>}or (name cb and segid "AAAA" and resid 168)
{====>}or (name cb and segid "BBBB" and resid 168)
{====>}or (name cb and segid "AAAA" and resid 171)
{====>}or (name cb and segid "BBBB" and resid 171)
{====>}or (name cb and segid "AAAA" and resid 174)
{====>}or (name cb and segid "BBBB" and resid 174)
{====>}or (name cb and segid "AAAA" and resid 166)
{====>}or (name cb and segid "BBBB" and resid 166)
{====>}or (name cb and segid "AAAA" and resid 175)
{====>}or (name cb and segid "BBBB" and resid 175)
{====>}or (name cb and segid "AAAA" and resid 179)
{====>}or (name cb and segid "BBBB" and resid 179)
)
loop Lmethy
  display aria revert (bondedto (id $id) and name h*) end
  display do (store1=0) (bondedto (id $id) and name h*)
end loop Lmethy
!
!!! isopropyle groups
for $id in id (name ca and (
{====>} (segid "AAAA" and resid 170)
{====>}or (segid "BBBB" and resid 170)
{====>}or (segid "AAAA" and resid 164)
{====>}or (segid "BBBB" and resid 164)
))
loop Vrev
  display do (store1=0) (byresid (id $id) and (name cg1 or name cd1))
  show (resn) (id $id)
  if ($result eq VAL) then
    evaluate ($name3 = cg1)
    evaluate ($name4 = cg2)
  elseif ($result eq LEU) then
    evaluate ($name3 = cd1)
    evaluate ($name4 = cd2)
  end if
  display aria revert (bondedto (byresid(id $id)
  display and (name $name3 or name $name4)) and name h*) end
end loop Vrev
!
!!! assignments as is
!!! methylene protons
for $id in id
(
{====>} (name cb and segid "AAAA" and resid 173)
{====>}or (name cb and segid "BBBB" and resid 173)
{====>}or (name cb and segid "AAAA" and resid 172)
{====>}or (name cb and segid "BBBB" and resid 172)
{====>}or (name cb and segid "AAAA" and resid 170)
{====>}or (name cb and segid "BBBB" and resid 170)
)
loop Lmethy

```

```

        display do (store1=0) (bondedto (id $id) and name h*)
    end loop Lmethy
    ! edited by ESS 02/20/2007
    !
    !!! isopropyle groups
    for $id in id (name ca and
    (
    {==>} (segid "AAAA" and resid 169)
    {==>}or (segid "BBBB" and resid 169)
    {==>}or (segid "AAAA" and resid 171)
    {==>}or (segid "BBBB" and resid 171)
    {==>}or (segid "AAAA" and resid 174)
    {==>}or (segid "BBBB" and resid 174)
    {==>} (segid "AAAA" and resid 166)
    {==>}or (segid "BBBB" and resid 166)
    {==>}or (segid "AAAA" and resid 175)
    {==>}or (segid "BBBB" and resid 175)
    {==>}or (segid "AAAA" and resid 179)
    {==>}or (segid "BBBB" and resid 179)
    ))
    loop Vrev
        display do (store1=0) (byresid (id $id) and (name cgl or name cd1))
    end loop Vrev

```

```
!$Revision: 2.1 $  
!$Date: 2001/06/18 12:28:21 $  
!$RCSfile: secondarystructure.cns,v $
```

```
!!! in this file, you have to define the secondary structure etc for the  
!!! calculation of the rms difference from the average structure  
!!! if this file remains empty, the default values are  
!!! (really) all residues, including any floppy histags  
!!! the ordered region defined by wellordered.inp for store1  
!!! delete the curly brackets to uncomment  
!  
!!! define the secondary structure:  
ident (store1) ! ESS 06/01/2007  
    (resid 168:185)  
!  
!!! define the range for all residues:  
ident (store2) (resid 154:188) ! ESS 06/01/2007
```

```

! Modified by ESS to prevent HB of the following residues from floating
!   Leu166, Leu169, Leu170, Leu171, His173, Leu174, Leu175, Leu179
!
do (q = 0) (all)
do (store1 = 0) (all)
  do (store1 = 2) (resn leu and name cdl)
  do (store1 = 2) (resn val and name cgl)
  do (store1 = 3) (resn arg and name nhl)
do (store1 = 1) (id 2)      ! "AAAA" G154 HA3
do (store1 = 1) (id 14)    ! "AAAA" G155 HA3
do (store1 = 1) (id 25)    ! "AAAA" I156 HG12
do (store1 = 1) (id 42)    ! "AAAA" F157 HB2
do (store1 = 1) (id 62)    ! "AAAA" S158 HB2
do (store1 = 1) (id 83)    ! "AAAA" E160 HB2
do (store1 = 1) (id 86)    ! "AAAA" E160 HG2
do (store1 = 1) (id 98)    ! "AAAA" F161 HB2
do (store1 = 1) (id 118)   ! "AAAA" L162 HB2
do (store1 = 1) (id 137)   ! "AAAA" K163 HB2
do (store1 = 1) (id 140)   ! "AAAA" K163 HG2
do (store1 = 1) (id 143)   ! "AAAA" K163 HD2
do (store1 = 1) (id 146)   ! "AAAA" K163 HE2
do (store1 = 1) (id 175)   ! "AAAA" F165 HB2
!do (store1 = 1) (id 195)   ! "AAAA" L166 HB2      ! commented by ESS 03/13/2007
do (store1 = 1) (id 213)   ! "AAAA" P167 HB2
do (store1 = 1) (id 216)   ! "AAAA" P167 HG2
do (store1 = 1) (id 219)   ! "AAAA" P167 HD2
do (store1 = 1) (id 228)   ! "AAAA" S168 HB2
!do (store1 = 1) (id 239)   ! "AAAA" L169 HB2      ! commented by ESS 03/02/2007
!do (store1 = 1) (id 258)   ! "AAAA" L170 HB2      ! commented by ESS 04/05/2007
!do (store1 = 1) (id 277)   ! "AAAA" L171 HB2      ! commented by ESS 04/05/2007
do (store1 = 1) (id 296)   ! "AAAA" S172 HB2
!do (store1 = 1) (id 307)   ! "AAAA" H173 HB2      ! commented by ESS 03/02/2007
!do (store1 = 1) (id 325)   ! "AAAA" L174 HB2      ! commented by ESS 04/05/2007
!do (store1 = 1) (id 344)   ! "AAAA" L175 HB2      ! commented by ESS 03/02/2007
do (store1 = 1) (id 375)   ! "AAAA" I177 HG12
do (store1 = 1) (id 390)   ! "AAAA" G178 HA3
!do (store1 = 1) (id 399)   ! "AAAA" L179 HB2      ! commented by ESS 03/02/2007
do (store1 = 1) (id 416)   ! "AAAA" G180 HA3
do (store1 = 1) (id 427)   ! "AAAA" I181 HG12
do (store1 = 1) (id 444)   ! "AAAA" Y182 HB2
do (store1 = 1) (id 467)   ! "AAAA" I183 HG12
do (store1 = 1) (id 482)   ! "AAAA" G184 HA3
do (store1 = 1) (id 491)   ! "AAAA" R185 HB2
do (store1 = 1) (id 494)   ! "AAAA" R185 HG2
do (store1 = 1) (id 497)   ! "AAAA" S185 HD2
do (store1 = 1) (id 515)   ! "AAAA" S186 HB2
do (store1 = 1) (id 518)   ! "AAAA" S186 HG2
do (store1 = 1) (id 521)   ! "AAAA" S186 HD2
do (store1 = 1) (id 539)   ! "AAAA" L187 HB2
do (store1 = 1) (id 567)   ! "BBBB" G154 HA2
do (store1 = 1) (id 579)   ! "BBBB" G155 HA2
do (store1 = 1) (id 591)   ! "BBBB" I156 HG12
do (store1 = 1) (id 608)   ! "BBBB" F157 HB2
do (store1 = 1) (id 628)   ! "BBBB" S158 HB2
do (store1 = 1) (id 649)   ! "BBBB" E160 HB2
do (store1 = 1) (id 652)   ! "BBBB" E160 HG2
do (store1 = 1) (id 664)   ! "BBBB" F161 HB2
do (store1 = 1) (id 684)   ! "BBBB" L162 HB2
do (store1 = 1) (id 703)   ! "BBBB" K163 HB2
do (store1 = 1) (id 706)   ! "BBBB" K163 HG2
do (store1 = 1) (id 709)   ! "BBBB" K163 HD2
do (store1 = 1) (id 712)   ! "BBBB" K163 HE2
do (store1 = 1) (id 741)   ! "BBBB" F165 HB2
!do (store1 = 1) (id 761)   ! "BBBB" L166 HB2      ! commented by ESS 03/13/2007
do (store1 = 1) (id 779)   ! "BBBB" P167 HB2
do (store1 = 1) (id 782)   ! "BBBB" P167 HG2
do (store1 = 1) (id 785)   ! "BBBB" P167 HD2
do (store1 = 1) (id 794)   ! "BBBB" S168 HB2
!do (store1 = 1) (id 805)   ! "BBBB" L169 HB2      ! commented by ESS 03/02/2007

```

```

!do (store1 = 1) (id 824)      ! "BBBB" L170 HB2      ! commented by ESS 04/05/2007
!do (store1 = 1) (id 843)      ! "BBBB" L171 HB2      ! commented by ESS 04/05/2007
do (store1 = 1) (id 862)      ! "BBBB" S172 HB2
!do (store1 = 1) (id 873)      ! "BBBB" H173 HB2      ! commented by ESS 03/02/2007
!do (store1 = 1) (id 891)      ! "BBBB" L174 HB2      ! commented by ESS 04/05/2007
!do (store1 = 1) (id 910)      ! "BBBB" L175 HB2      ! commented by ESS 03/02/2007
do (store1 = 1) (id 941)      ! "BBBB" I177 HG12
do (store1 = 1) (id 955)      ! "BBBB" G178 HA2
!do (store1 = 1) (id 965)      ! "BBBB" L179 HB2      ! commented by ESS 03/02/2007
do (store1 = 1) (id 981)      ! "BBBB" G180 HA2
do (store1 = 1) (id 993)      ! "BBBB" I181 HG12
do (store1 = 1) (id 1010)     ! "BBBB" Y182 HB2
do (store1 = 1) (id 1033)     ! "BBBB" I183 HG12
do (store1 = 1) (id 1047)     ! "BBBB" G184 HA2
do (store1 = 1) (id 1057)     ! "BBBB" R185 HB2
do (store1 = 1) (id 1060)     ! "BBBB" R185 HG2
do (store1 = 1) (id 1063)     ! "BBBB" R185 HD2
do (store1 = 1) (id 1081)     ! "BBBB" R186 HB2
do (store1 = 1) (id 1084)     ! "BBBB" R186 HG2
do (store1 = 1) (id 1087)     ! "BBBB" R186 HD2
do (store1 = 1) (id 1105)     ! "BBBB" L187 HB2
do (store1 = 1) (id 503)      ! "AAAA" R185 HH12
do (store1 = 1) (id 506)      ! "AAAA" R185 HH22
do (store1 = 1) (id 527)      ! "AAAA" R186 HH12
do (store1 = 1) (id 530)      ! "AAAA" R186 HH22
do (store1 = 1) (id 1068)     ! "BBBB" R185 HH11
do (store1 = 1) (id 1071)     ! "BBBB" R185 HH21
do (store1 = 1) (id 1092)     ! "BBBB" R186 HH11
do (store1 = 1) (id 1095)     ! "BBBB" R186 HH21

```