

9

# Anatomy of the Coiled Coil and Its Role in the Conformational Change of Influenza Hemagglutinin

by

Chavela Marguerite Carr

Submitted to the Department of Biology  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY  
in Biology  
at the

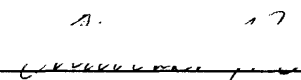
Massachusetts Institute of Technology

June 1995

© 1995 Chavela M. Carr. All rights reserved.

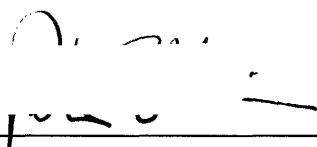
The author hereby grants to MIT permission to reproduce and to distribute publicly paper  
and electronic copies of this thesis document in whole or in part.

Signature of Author

  
\_\_\_\_\_

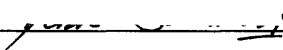
Department of Biology  
May 5, 1995

Certified by

  
\_\_\_\_\_

Peter S. Kim  
Associate Professor, Department of Biology  
Thesis Supervisor

Accepted by

  
\_\_\_\_\_

Frank Solomon  
Chairman, Biology Graduate Committee

Science

MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

1

MAY  
JUN 08 1995

LIBRARIES

Dedication

In memory of

Lawrence Silas Carr

and

Dwight Herrick Macduff

## Acknowledgments

For impressing upon me their appreciation of nature and the importance of critical thinking, I dedicate this thesis to my late grandfathers: Lawrence Silas Carr and Dwight Herrick Macduff. I also thank my living family members for being patient with me and forgiving my forgetfulness of birthdays and other special occasions during these seven years at MIT.

I have been influenced by a great number of excellent teachers who believed in me and pushed me to excel beyond what I imagined possible. Most notable for my scientific and analytical development are my human-genetics high-school teacher, Mr. Gordon Mendenhall; Mr. Dan Stiko, high-school teacher of several math courses; my undergraduate mentor, Professor Douglas Cavener; and, of course, my graduate advisor, Dr. Peter Kim. Toward my personal growth, I am indebted to my high-school choir and dance director Ms. Sandra Butz; Mr. Gary Meyers, my high-school theater director; and Mr. Michael Kim Howden, junior-high world-history teacher, enduring friend, counselor, and confidant.

The best years of my life have been spent in the exhilarating environment of MIT and the Whitehead Institute. The friends and colleagues I have met here I will treasure forever. For their generosity, I especially thank my talented and beloved friends outside the lab: Mattan Kamon, whose computer advice, encouragement and affection helped me through the best and the most difficult years; Joel Phillips, for good food, good wine, and a dear and lasting friendship; Christina Scherer, classmate and comrade in happiness and hardships from the first to the last day of graduate school; Shiufun Cheung, for his genuine friendship and work on the graphic rendering of the HA conformations, which was so widely publicized; Josephine Cheung, for her cheerful nature and friendship; Ignacio McQuirk, for his appreciation of spicy food, good tea, excellent chocolate, and friendship; Michelle and Bob Spina, that crazy couple who brought out my loud side and made sure I enjoyed myself thoroughly, whenever I was with them; Meg Winberg, who kept my musical and theatrical spirit alive; Amelia Shen, a terrific housemate and source of other lasting friendships; David Shia, for sharing his computer, printer, and good humor; and Brad Stewart, especially for his help with the computer search for coiled coils that led to the

discovery described in Chapter 2. Special thanks goes to Eve Nichols, the Whitehead's publicity director. Eve's enthusiasm and hard work in promoting my project has gone a long way to encourage and inspire me.

The familial friendships inevitably formed when people work closely together for several years added to the rich experience of graduate school. For their scientific and non-scientific contributions I thank "the old crew" and the "young turks" of the lab: Erin O'Shea, for the good-old days, the encouragement, and especially, the laughter. Jon Staley, ("Studley") for so much support and friendship when I was a young graduate student; Terry Oas, ("Oasage") for all manner of advice and so much help in the lab; Rheba Rutkowski (R<sup>2</sup>), for her critical, expressive and outspoken appreciation of human nature, and for her approval, which meant all the more for its difficulty to attain; David Litwack, for camaraderie during year-one in the lab, and especially for partnership in the outstanding prank on Staley; Jonathan Weissman, for his great sense of humor, so much scientific advice and a supportive friendship to endure all; Pehr Harbury (Harb) for numerous lessons, for his generous help in reading and correcting numerous sections of this thesis, and for a lasting friendship; Dan Minor (Major Minor) for practical and scientific advice and for his efforts to unify the lab in group outings, like the hiking trip from Hell; Brenda Schulman, (N\*) for sister-like friendship, for critical advice about my written work and for help digging through the literature for references for the introduction of this thesis; Bob Talanian (BT) for expert advice on DNase I footprinting (Appendix I); David Lockhart, (Lock-jaw) my role model, my friend, my confidant; Jamie McKnight, whose help with vesicles and fluorescence was essential to chapter 3; Tom Alber and his group for their large contributions as the structural collaborators in the GCN4 project (Appendix I); Zeng-Yu Peng, for his candid and critical advice plus his humble and good-natured spirit; Kevin Lumb whose contributions to this thesis deserve special consideration, as the work of Appendix II is primarily his; Andrea Cochran, who provided me with an alternative perspective on the world; Martha Oakley, for steadfast support, especially in my darkest hour; Deborah Fass, for her inspirational zeal for everything she does; Lawren Wu, whose calm manner and easy sense of humor is good medicine in the hectic environment of the lab; Charu Chaudhry, my hard-working, talented undergraduate research student (UROP), who worked out the proteinase K assay used in Chapter 3, and who was a great source of pride and joy; Steve Blacklow, for his enthusiasm; Min Liu, for his personal and practical advice; Paul Matsudaira and his students Kathy Collins and Don

Doering for many interesting scientific conversations and help during my early years; and Mitch Sanders, Michael Way and Navin Pokala for help with protein gels and immunoblot protocols.

For expert technical assistance, I thank Mike Milhollen, who made the expression vectors for prHMG and purified rHA2-X117; Mike Burgess, who made and analyzed the masses of many peptides used in the studies described here and some not described here; Rheba Rutkowski, for the reasons mentioned above and for peptide synthesis; Sean Britt for assistance with the GCN4 peptides; Christina Doetsch, the newest technician to the lab, a true and lasting friend and treasure; Paul Matsudaira's technical associate Mark Chaffel, for help with tissue culture and use of photography supplies and Matt Footer for his professional advice and for use of his photography skills and tools in preparing figures used in Chapter 3.

The most difficult acknowledgment to write is the one owed to my advisor, Peter S. Kim. Though these words cannot convey the weight of my admiration and gratitude, I thank you, Peter. I appreciate the exceptional facilities and opportunities that Peter has provided for the members of the lab, including powerful and plentiful, high-tech laboratory equipment, the large variety of scientific journals, expensive Silicon Graphics, Inc. (SGI) computers for structural analysis; and especially the diverse group of highly motivated students and post-docs who interacted to form a stimulating scientific community. My success is a direct result of Peter's willingness to take risks in starting new projects and his commitment to provide everything I thought necessary for the experiments, including new equipment, a hired computer consultant, and large quantities of infectious flu virus! Peter's good judgment in hiring Charu Chaudhry, the talented undergraduate student, who worked with me on the HA conformational assay (Chapter 3), led to a deeply rewarding teaching experience. I am proud to be among the ever-growing league of distinguished students that have graduated from this lab.

The 6 years I have spent in Peter's lab have confirmed the advice of sage professors from both my undergraduate and graduate years, and I have since passed on this advice to prospective graduate students, who are hoping to build a career in science: You have to love it; its too hard!. So why did I spend 7 years of my young life in the laboratory? I did it for love.

# **Anatomy of the Coiled Coil and Its Role in the Conformational Change of Influenza Hemagglutinin**

by

Chavela Marguerite Carr

Submitted to the Department of Biology on May 5, 1995 in  
partial fulfillment of the requirements for the  
Degree of Doctor in Philosophy in Biology

## **Abstract**

The identification of an unexpected coiled-coil sequence in influenza hemagglutinin (HA) led to a model for the conformational change required for the membrane-fusion function of this protein (Chapter 2). Analysis of coiled-coil peptides corresponding to the fusion-active conformation of HA not only support the model, but led to a prediction for the mechanism of the conformational change. Chapter 3 describes evidence that the "native," fusion-inactive conformation of HA is trapped in a metastable state, which, when destabilized, undergoes the conformational change and induces membrane fusion. Chapter 4 discusses recent advances that may shed some light on the steps for HA-mediated membrane fusion.

The appendices are devoted to the leucine-zipper motif, a short coiled-coil region that forms the dimerization domain of bZIP and bHLH-ZIP proteins. A conserved, buried hydrogen bond in the hydrophobic coiled-coil interface destabilizes the GCN4 leucine zipper, yet it provides specificity for the dimer conformation (Appendix I A; see also Harbury et al., 1993). The stabilizing effects of increasing the hydrophobicity of a buried amino acid is quantified using non-natural amino acids for the incremental addition of methylene groups to the central position of the GCN4 leucine-zipper interface (Appendix I B). The identification of a stable subdomain in the GCN4 leucine zipper has implications for the mechanism of protein folding (Appendix II).

Thesis Supervisor: Dr. Peter S. Kim,  
Titte: Associate Professor of Biology

## Table of Contents

Dedication	2
Acknowledgments	3
Abstract	6
Table of Contents	7
Chapter 1: Introduction	9
Chapter 2:	23
Chapter 2 has been published as C. M. Carr and P. S. Kim, “A Spring-Loaded Mechanism for the Conformational Change of Influenza Hemagglutinin.” Cell 73, 823-832 (1993). © Cell Press.	
Chapter 3: Evidence that Influenza Hemagglutinin is Trapped in the Native Conformation	56
Chapter 4:	85
Chapter 4 has been published as C. M. Carr and P. S. Kim, “Flu Virus Invasion: Halfway There.” Science 73, 234-236 (1994). © American Association for the Advancement of Science.	

<u>Appendix</u>	95
Appendix I: Structural and Functional Studies of the Coiled Coil	98
A. A Conserved Asparagine Destabilizes the Leucine Zipper of GCN4	99
B. Effects of Increased Hydrophobicity at the Coiled Coil Interface	104
Appendix II:	155
Appendix 2 has been published as K. J. Lumb, C. M. Carr and P. S. Kim, "Subdomain Folding of the Coiled Coil Leucine Zipper from the bZIP Transcriptional Activator GCN4." <i>Biochemistry</i> <b>33</b> , 7361-7367 (1994). © American Chemical Society.	
Appendix III: Bimolecular Equilibrium Equation	181
Biographical Profile	185



# Chapter 1

## Introduction

Proteins are the workhorses of life. The diversity of functions performed by proteins is enormous. For example, proteins are required for the structural integrity of skin and bones, the mechanical work performed by muscles, the efficient conversion of food into energy, the transduction of signals during sensory perception, and the transcription and translation of genes for the synthesis of other proteins. Fundamental to the description of any one of these processes is a "nuts and bolts" understanding of the way proteins work. Equipped with this knowledge, we may then begin to understand the nature of diseases caused either by cellular proteins with altered or abolished functions, or by parasites such as viruses, whose proteins function to invade and overtake the cell.

This chapter begins by addressing the relationship between protein structure and function and by outlining the approaches used to discover how a protein works. Following a general introduction, this chapter proposes how these approaches can be applied to specific questions concerning the role of the coiled-coil motif in the structure and function of hemagglutinin (HA), the membrane fusion protein from influenza virus. Chapters 2 - 4 describe how the application of our knowledge about the sequence determinants of coiled-coil structure leads to the discovery of a novel role for the coiled coil in membrane fusion. The appendices describe the identification of specific determinants for the structure and stability of the coiled-coil dimerization domain of GCN4, a transcription factor from the bZIP family of proteins.

## Protein Structure and Function

*"Conformation---exemplified by the relationship between the three-dimensional structure of proteins and their biological activity." --Lubert Stryer (Stryer, 1988).*

The function of a protein is dictated by its conformation. Any attempt to reveal the way a protein works requires knowledge of its three-dimensional structure. Yet interpretation of the protein structure often requires additional information to highlight the amino-acid residues of particular interest for protein function. Much progress has been made toward understanding protein function by combining 2 complementary approaches:

structural analysis and mutagenesis. Researchers have also had success in reconstituting protein function from protein fragments, a method hereafter referred to as "protein dissection." Ultimately, our understanding of protein structure and function will be tested by our ability to design functional proteins.

### *Structure Determination*

The most intuitive approach toward understanding a protein's function is to look at its conformation. By far the best methods for high-resolution protein-structure determination are x-ray crystallography (Blundell and Johnson, 1990), and more recently, nuclear magnetic resonance spectroscopy (Wüthrich, 1986). But how does the structure explain the function? For those proteins that bind to ligands or substrates (i.e. enzyme-substrate complexes or transcription factor-DNA complexes), a co-crystal can reveal the details of specific interactions, explaining specificity of binding. However, in many cases the information conveyed by the structure is difficult to interpret in terms of protein function or dynamics. The classic analogy is that a polaroid of a car's engine does not explain how the car works. Nonetheless, examination of a protein's structure inspires hypotheses for the mechanism of function, hypotheses which can be tested using a variety of methods to perturb the protein and examine the effect of the perturbation on protein structure and function.

While the first step to understanding function is to look at the structure, proteins that share the same conformation may have different functions (for a review, see (Orengo et al., 1994). For example, the four-helix bundle protein, Rop, is used for plasmid replication (Banner et al., 1987), cyt b562, for electron transport (Mathews et al., 1979), IL4, as a cytokine (Smith et al., 1992; Powers et al., 1992) and as a component of the coat protein of tobacco mosaic virus (Champness et al., 1976). The immunoglobulin superfold structure of antibodies is seen in many places, including the NF- $\kappa$ B transcription factor (Ghosh et al., 1995; Müller et al., 1995). The existence of protein families that share a structural motif or modular domain yet perform different specific functions suggests that there must be amino-acid sequence determinants for unique function superimposed on the sequence determinants for a protein's three-dimensional structure. Thus, in order to understand the specificity of protein function, these sequence determinants must be identified and their functional role assessed.

## *Mutagenesis*

A second and complementary approach toward discerning the functional regions of a protein is amino-acid substitution, or mutagenesis. Because the conformation of a protein is dictated by its amino-acid sequence (Epstein et al., 1963), alteration of the sequence can change or abolish the structure and function of the protein. Random mutagenesis, followed by a genetic screen or selection for functional protein, reveals that some sites in proteins are tolerant to substitution, while others are less so (Coplen et al., 1990; Hu et al., 1990). Those positions that are sensitive to substitutions may play an important role either in the structure or specificity of the protein function. Mutations which abolish function without disrupting structure reveal amino acids which may play an important role in the specificity of protein function. For example, mutations in the DNA-binding domain of the lambda repressor and cro transcription factors led to the discovery of how these proteins recognize and distinguish between different DNA sequences (Hochschild et al., 1986).

Similarly, structurally and functionally important residues are conserved between proteins of the same family, and the role of these residues can be assessed by site-directed mutagenesis (Reidhaar-Olson and Sauer, 1988). More recently, the introduction of “unnatural” amino acids into proteins has increased the repertoire of functional groups for substitution (Ellman et al., 1992; Mendel et al., 1992; Judice et al., 1993; Chung et al., 1993), allowing for a fine-tuned analysis of the functional groups important for protein structure and function. In combination with high-resolution structure determination, alanine-scanning mutagenesis is another powerful tool for uncovering the determinants of protein structure and function (Cunningham and Wells, 1989).

## *Protein Dissection*

A third approach that has met with surprising success is to reduce the structure/function problem into parts by protein dissection (Oas and Kim, 1988; Goodman and Kim, 1989; Peng and Kim, 1994). At first sight, the structure of proteins appears to be unapproachably complex: a composite of interdependent interactions between amino acids both near and far from each other in sequence determines the overall 3-dimensional

structure that is so essential to the function of the protein. Dissecting the protein into pieces by breaking the polypeptide chain would seem an unproductive approach.

Nonetheless, several observations suggest that proteins may be composed of autonomous structural domains that can be studied separately: 1. Deletions or truncations of proteins yield functional domains for some proteins, for example, DNA-binding domains of bZIP proteins (for a review see Hu and Sauer, 1992). 2. Proteolytic digestion of proteins into protease-resistant fragments has revealed that many proteins are composed of more than one stable domain or subdomain, which folds as an autonomous unit (Wu et al., 1994; Matsudaira, 1992). 3. The discovery of structural motifs shared between different families suggested that these evolutionarily conserved modules may be able to form independent structural units (for a review, see Cohen et al., 1995). Several structural motifs have been confirmed as autonomous folding units by protein dissection, for example, the leucine-zipper dimerization domain of the bZIP family of transcription factors (O'Shea et al., 1989a, b) and the SH2 and SH3 protein-binding domains of signalling proteins (Cohen et al., 1995).

Peptide models have been used successfully to study protein structure/function and protein folding. Reconstitution of RNase A activity by mixing fragments of this protein suggests that the whole is, in some cases, the sum of its parts (Richards and Vithayathil, 1959). In addition, reconstitution of trypsin inhibition and native-like structure with peptide models of bovine pancreatic trypsin inhibitor illustrate that the connectivity of the polypeptide chain need not remain intact for correct folding of single-domain proteins (Oas and Kim, 1988; Staley and Kim, 1990). The same conclusion comes from the result that circularly permuted T4 lysozyme retains its native, stable fold (Zhang et al., 1993).

The implication of these discoveries is that otherwise intractable, insoluble or overly large proteins can be manageably dissected into parts for studies of structure and function in solution. Quick and easy methods to assay protein structure and stability, such as circular dichroism, fluorescence and absorbance spectroscopy can then be applied. Trimming proteins into their minimal structural units also aids detailed structural analysis by crystallography and NMR. Of course, not all proteins will withstand dissection and reconstitution, but progress has already been made on biologically important molecules

such as SH3 proteins (Lim et al., 1994; Yu et al., 1994), tumor suppressor p53 (Jeffrey et al., 1995), and Max (Ferre-D'Amare et al., 1993), the partner of the protooncogene, c-Myc.

### *Protein Design*

The ultimate and most ambitious approach toward understanding a protein's structure and function is protein design. Hypotheses for sequence patterns that specify a given conformation can be tested by *de novo* design of a polypeptide chain, which should fold into the predicted structure if the hypothesis is correct. This approach has met with limited success (O'Shea et al., 1993), but as more is learned about determinants for uniqueness of structure, this will be the ultimate test of our understanding of protein structure and function.

### **Purpose**

This thesis examines the role of the coiled-coil structural motif in the structure and function of two very different proteins: hemagglutinin (HA), the membrane-fusion protein from influenza virus and GCN4, a bZIP transcription factor. The primary approach used in both studies is protein dissection: coiled-coil peptides were studied in solution to test predictions of protein stability and conformation. In addition to structural studies, the functions of these proteins have been reconstituted *in vitro* to test the effects of perturbation on protein function.

#### **I. Role of a Coiled Coil in the Conformational Change of Influenza Hemagglutinin Required for Membrane Fusion**

The fact that specialized proteins are required for membrane fusion events, such as fertilization, neurotransmitter release and viral infection, has exciting implications for the control of these events by inhibition of the fusogenic function of the proteins responsible. Nonetheless, progress toward the rational design of inhibitors requires an understanding of the mechanism of membrane fusion and the role of these specialized proteins in bilayer mixing. In the simplest model for membrane fusion, there is one basic mechanism of fusion with variation at the level of regulation of fusion activation for specificity of

targeting and timing of the fusion event. With this in mind, we hoped to learn some general principles about the mechanism of membrane fusion by studying the best characterized membrane fusion protein, influenza hemagglutinin (HA).

Several key questions are central to understanding membrane fusion: What is the structure of the fusogenic conformation? What is the mechanism of the conformational change? What is the mechanism of membrane fusion? These three questions are addressed in a discussion of experimental results, which shed light on these issues while raising new questions. In addition, the possible generality of the principles learned from HA and implications for the control of infection are discussed.

### *A Model for the Fusogenic Conformation of HA*

The HA envelope protein is responsible for the membrane fusion event that results in entry of influenza virus into the host cell. HA is held together in the native conformation by a long, three-stranded coiled-coil core (Fig. 1). In response to the acidic environment of the cellular endosome, HA undergoes an irreversible conformational change that is required for its function in membrane fusion. We propose that the dynamic nature of the HA coiled coil is critical for this conformational change (Chapter 2).

Chapter 2 describes the identification of a "latent" coiled-coil region in HA. The coiled coil is latent because the structure of this region is extended to form a loop conformation in the native state, though the sequence suggests that this region has a high propensity to form a coiled coil. We proposed that this region adopts a coiled coil as a consequence of the conformational change required for membrane fusion. Evidence is presented for the dramatically different "fusogenic" conformation of HA, and the relevance of this conformation to its function in membrane fusion is discussed.

### *The Mechanism of the HA Conformational Change*

Though mildly acidic pH induces the conformational change *in vivo*, several observations suggested to us that the mechanism for HA-mediated membrane fusion is not necessarily pH-dependent (see Chapter 2). These considerations led us to propose that the low pH *in vivo* acts primarily to destabilize a metastable native conformation, causing the

conformational change that induces membrane fusion. The major prediction of this hypothesis is that the conformational change and membrane fusion can be induced at neutral pH in response to destabilizing conditions. In Chapter 3 we analyze the conformation and function of HA in intact influenza virus to test this prediction. Evidence is presented for a mechanism of the HA conformational change that may be relevant for other membrane fusion events.

### *How do Membranes Fuse?*

How the conformational change of HA is utilized for the final steps in membrane fusion remains a mystery. Chapter 4 raises questions about the mechanism of membrane fusion and summarizes some recent discoveries plus a model for the next step in HA-mediated membrane fusion.



## **II. (Appendix) The Coiled Coil: Stability and Specificity of the GCN4 Leucine Zipper**

The "leucine-zipper," is a short coiled-coil dimerization motif of bZIP and bHLH-ZIP transcription factors and a model system for the analysis of coiled coils. The structural requirements of the coiled coil result in a conserved pattern of hydrophobic and hydrophilic residues in the amino-acid sequence, known as the heptad repeat. The leucine-zipper sequences also display this pattern, and in the case of GCN4 and the Fos/Jun heterodimer, peptides corresponding to the leucine-zipper sequences form specific coiled-coil dimers in solution (O'Shea et al., 1989a, b).

The short coiled coil of the leucine-zipper motif faces the challenge of forming a stable, coiled-coil dimer while maintaining specificity for its correct partner to form a functional transcription factor. For the Fos/Jun heterodimer, the determinants of partner specificity have been localized to the predominantly charged amino-acid residues that flank the hydrophobic interface of the leucine zipper dimer (O'Shea et al., 1992). We sought to uncover additional determinants of leucine-zipper structure and function using the coiled coil from GCN4. Appendix I addresses the role of a conserved asparagine in the leucine-zipper coiled coil and the contribution of hydrophobicity to coiled-coil stability. Detailed crystallographic analyses of the structures of GCN4 peptides were made possible by an ongoing collaboration with Tom Alber's lab (O'Shea et al., 1991 and Appendix I). Identification of a stably folded subdomain in an N-terminal deletion mutant of the GCN4 leucine-zipper peptide (Appendix II), suggests that leucine-zipper folding may proceed through this conformation as an intermediate in folding.

## References

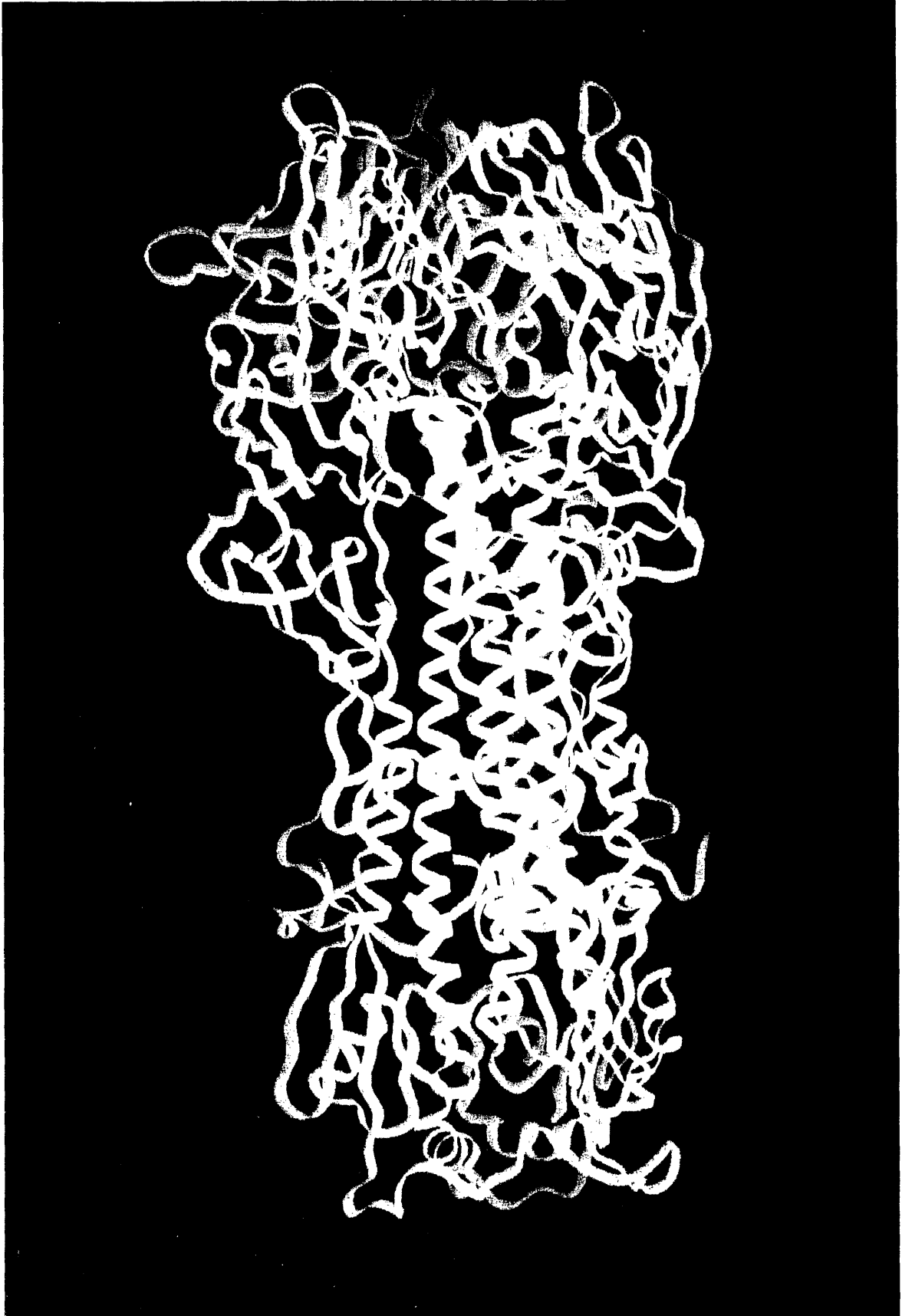
- Banner, D. W., Kokkinidis, M., and Tsernoglou, D. (1987). Structure of the ColE1 rop protein at 1.7 Å resolution. *Journal of Molecular Biology* 196, 657-675.
- Blundell, T. L., and Johnson, L. N. (1990). *Protein Crystallography*. (London: Academic Press).
- Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. G., and Klug, A. (1976). The structure of the protein disk of tobacco mosaic virus to 5Å resolution. *Nature* 259, 20-4.
- Chung, H. H., Benson, D. R., Cornish, V. W., and Schultz, P. G. (1993). Probing the role of loop 2 in Ras function with unnatural amino acids. *Proc Natl Acad Sci U S A* 90, 10145-9.
- Cohen, G. B., Ren, R., and Baltimore, D. (1995). Modular Binding Domains in Signal Transduction Proteins. *Cell* 80, 237-248.
- Coplen, L. J., Frieden, R. W., and Goldenberg, D. P. (1990). A genetic screen to identify variants of bovine pancreatic trypsin inhibitor with altered folding energetics. *Proteins* 7, 16-31.
- Cunningham, B. C., and Wells, J. A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 244, 1081-5.
- Ellman, J. A., Mendel, D., and Schultz, P. G. (1992). Site-specific incorporation of novel backbone structures into proteins. *Science* 255, 197-200.
- Epstein, C. J., Goldberger, R. F., and Anfinsen, C. B. (1963). *Cold Spring Harbor Symp. Quant. Biol.* 27, 439-449.
- Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993). Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain [see comments]. *Nature* 363, 38-45.
- Ghosh, G., Van Duyne, G., Ghosh, S., and Sigler, P. B. (1995). Structure of NF-κB p50 homodimer bound to a κB site. *Nature* 373, 303-310.
- Goodman, E. M., and Kim, P. S. (1989). Folding of a peptide corresponding to the alpha-helix in bovine pancreatic trypsin inhibitor. *Biochemistry* 28, 4343-7.
- Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* 262, 1401-1407.
- Hochschild, A., Douhan, J. 3., and Ptashne, M. (1986). How lambda repressor and lambda Cro distinguish between OR1 and OR3. *Cell* 47, 807-16.

- Hu, J. C., O'Shea, E. K., Kim, P. S., and Sauer, R. T. (1990). Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* 250, 1400-3.
- Hu, J. C., and Sauer, R. T. (1992). The basic-region leucine-zipper family of DNA binding proteins. *Nucleic Acids and Molecular Biology* 6, 82-101.
- Jeffrey, P. D., Gorina, S., and Pavletich, N. P. (1995). Crystal Structure of the Tetramerization Domain of the p53 Tumor Suppressor at 1.7 Angstroms. *Science* 267, 1498-1502.
- Judice, J. K., Gamble, T. R., Murphy, E. C., de Vos, A. M., and Schultz, P. G. (1993). Probing the mechanism of staphylococcal nuclease with unnatural amino acids: kinetic and structural studies. *Science* 261, 1578-81.
- Lim, W. A., Richards, F. M., and Fox, R. O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* 372, 375-9.
- Mathews, F. S., Bethge, P. H., and Czerwinski, E. W. (1979). The structure of cytochrome b562 from *Escherichia coli* at 2.5 Å resolution. *J Biol Chem* 254, 1699-706.
- Matsudaira, P. (1992). Mapping structural and functional domains in actin-binding proteins. (Oxford: IRL Press, Oxford University Press).
- Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., and Schultz, P. G. (1992). Probing protein stability with unnatural amino acids. *Science* 256, 1798-802.
- Müller, C. W., Rey, F. A., Sodeoka, M., Verdine, G., and Harrison, S. C. (1995). Structure of the NF-κB p50 homodimer bound to DNA. *Nature* 373, 311-317.
- O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989a). Evidence that the leucine zipper is a coiled coil. *Science* 243, 538-542.
- O'Shea, E. K., Rutkowski, R., Stafford, W. F., and Kim, P. S. (1989b). Preferential heterodimer formation by isolated leucine zippers from Fos and Jun. *Science* 245, 646-648.
- O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. A. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* 254, 539-544.
- O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992). Mechanism of Specificity in the Fos-Jun Oncoprotein Heterodimer. *Cell* 68, 699-708.
- O'Shea, E. K., Lumb, K. J., and Kim, P. S. (1993). Peptide "Velcro": design of a heterodimeric coiled coil. *Current Biology* 3, 658-667.
- Oas, T. G., and Kim, P. S. (1988). A peptide model of a protein folding intermediate. *Nature* 336, 42-48.
- Orengo, C. A., Jones, D. T., and Thornton, J. M. (1994). Protein superfamilies and domain superfolds. *Nature* 372, 631-4.

- Peng, Z. Y., and Kim, P. S. (1994). A protein dissection study of a molten globule. *Biochemistry* 33, 2136-41.
- Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., and Clore, G. M. (1992). Three-dimensional solution structure of human interleukin-4 by multidimensional heteronuclear magnetic resonance spectroscopy. *Science* 256, 1673-7.
- Reidhaar-Olson, J. F., and Sauer, R. T. (1988). Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* 241, 53-7.
- Richards, F. M., and Vithayathil, P. J. (1959). The Preparation of Subtilisin-modified Ribonuclease and the Separation of the Peptide and Protein Components. *J. Biol. Chem.* 234, 1459-1464.
- Smith, L. J., Redfield, C., Boyd, J., Lawrence, G. M., Edwards, R. G., Smith, R. A., and Dobson, C. M. (1992). Human interleukin 4. The solution structure of a four-helix bundle protein. *J Mol Biol* 224, 899-904.
- Staley, J. P., and Kim, P. S. (1990). Role of a subdomain in the folding of bovine pancreatic trypsin inhibitor. *Nature* 344, 685-8.
- Stryer, L. (1988). *Biochemistry*. (New York: W. H. Freeman and Company).
- Wu, L. C., Grandori, R., and Carey, J. (1994). Autonomous subdomains in protein folding. *Protein Sci* 3, 369-71.
- Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*. (New York: Wiley).
- Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* 76, 933-45.
- Zhang, T., Bertelsen, E., Benvegna, D., and Alber, T. (1993). Circular Permutation of T4 Lysozyme. *Biochemistry* 32, 12311-8.

## Figure 1

A three-stranded coiled coil forms the core of native HA. A ribbon-representation of the trimeric, native HA ectodomain (BHA) is depicted in purple, with the three-stranded coiled coil highlighted in white. This structure represents HA in the native, fusion-inactive conformation. A conformational change is required for the membrane fusion activity of HA, and the coiled coil plays a central role (see Chapter 2).



## Chapter 2

## **A Spring-Loaded Mechanism for the Conformational Change of Influenza Hemagglutinin**

### Summary

Influenza hemagglutinin (HA) undergoes a conformational change that induces viral fusion with the cellular membrane. The structure of HA in the fusogenic state is unknown. We have identified a sequence that has a high propensity for forming a coiled coil. Surprisingly, this sequence corresponds to a loop region in the x-ray structure of native HA: the loop is followed by a three-stranded, coiled-coil stem. We find that a 36-residue peptide (LOOP-36), comprised of the loop region and the first part of the stem, forms a three-stranded coiled coil. This coiled coil is extended and stabilized in a longer peptide, corresponding to LOOP-36 plus the residues of a preceding, short  $\alpha$ -helix. These findings lead to a model for the fusogenic conformation of HA: the coiled-coil stem of the native state is extended, relocating the hydrophobic fusion peptide, by 100 Å, toward the target membrane.



## Introduction

Membrane fusion is necessary for a large number of diverse processes in biology. For example, protein trafficking, protein secretion, fertilization and neurotransmission require the fusion of distinct membranes to form a single lipid bilayer. In general, membrane fusion is very slow in the absence of specific proteins.

The best characterized membrane fusion event occurs in the infection of animal cells by influenza virus. Infection begins with the binding of virus to sialic acid, a component of membrane proteins and lipids. The bound virion is then internalized into a cellular endosome by receptor-mediated endocytosis. Ultimately, the viral envelope fuses with the membrane of the mature endosome in response to a drop in pH (for reviews, see Wiley and Skehel, 1987; Stegmann and Helenius, 1993). Membrane fusion is promoted by the trimeric, viral-envelope glycoprotein, hemagglutinin (HA), which also functions in cell attachment (at neutral pH) by binding to sialic acid. At low pH, HA is sufficient for the fusion of membranes *in vivo* and *in vitro* (White et al., 1982).

Each HA monomer is synthesized as a protein precursor, denoted HA0. The HA0 polypeptide is processed proteolytically to form a pair of disulfide-bonded peptides, denoted HA1 and HA2 (Lazarowitz et al., 1971; Skehel and Waterfield, 1975). This proteolytic processing is essential for the infectivity of the virus (Klenk et al., 1975; Lazarowitz and Choppin, 1975). Native HA binds sialic acid, but is dormant for membrane-fusion activity.

There is considerable evidence for a large, irreversible conformational change in HA which is required for membrane fusion (for reviews, see Wiley and Skehel, 1987; White, 1992; see however, Stegmann et al., 1990). The conformational change is induced by mildly acidic conditions (e.g., pH ~5, the pH of the mature endosome) and results in a trimeric structure that is remarkably thermostable (Doms and Helenius, 1986; Ruigrok et al., 1986; 1988). At neutral pH, a similar conformational change, leading to a stable, fusion-active state of HA, can be induced at temperatures above 60° C (Ruigrok et al., 1986). The conformation of HA that is active for membrane fusion will be referred to as the "fusogenic" state.

## Structure of the Native State of HA

HA can be proteolytically cleaved with bromelain to release the exoplasmic domain from the viral membrane. The resulting soluble domain, termed BHA, has sialic acid-binding and low-pH, liposome-binding activities similar to intact HA (Brand and Skehel, 1972; Doms et al., 1985). The x-ray crystal structure of the BHA trimer, in the native state, has been determined with and without bound sialic acid (Wilson et al., 1981; Weis et al., 1988; 1990a). The interaction between HA and sialic acid has been well described (Weis et al., 1988; Glick et al., 1991; Sauter et al., 1992), and this interaction is separate from the function of HA in membrane fusion.

The central region of the HA2 polypeptide folds into a helical-hairpin structure (see Fig. 1A): a short  $\alpha$ -helix (red) is connected to a long  $\alpha$ -helix (magenta) by an extended loop region (yellow). The long  $\alpha$ -helix interacts with the corresponding long  $\alpha$ -helices from two other HA2 polypeptides to form an interwound rope of three helices, called a three-stranded coiled coil. The three shorter  $\alpha$ -helices are displayed on the outside of the coiled coil. The sialic acid-binding domains of the HA1 subunits assemble on top of the fibrous stem, which is formed primarily by the three-stranded HA2 coiled coil (Fig. 1B). The length of the trimer, from the junction with the membrane to the distal tip of the HA1 subunits, is  $\sim 135 \text{ \AA}$  (Wilson et al., 1981).

HA2 is the transmembrane subunit, which spans the envelope membrane once. At the amino terminus of HA2 is a highly conserved, hydrophobic sequence of  $\sim 25$  residues, known to be necessary for membrane fusion (Daniels et al., 1985; Gething et al., 1986a). Although buried in the hydrophobic interior in the native state (Fig. 1; light blue), these residues become exposed in the fusogenic state (for reviews, see Wiley and Skehel, 1987; Stegmann and Helenius, 1993) and insert into the endosomal membrane (Stegmann et al., 1991). For these reasons, this sequence is referred to as the "fusion peptide" (for a review, see White, 1992).

The structure of HA in the fusogenic conformation is not known. An immediate puzzle is to understand how the exposed fusion peptide facilitates fusion of the viral envelope with the endosomal membrane since, in the native conformation, the fusion

peptide is buried near the viral envelope, ~100 Å from the distal tip of the HA molecule (Fig. 1).

### **Model for the Structure of the Fusogenic State**

We present a model for the fusogenic structure of HA. The model evolved from our interest in the coiled-coil motif found in many proteins (for a review, see Cohen and Parry, 1990) including the "leucine zipper" domain of some transcription factors (Landschulz et al., 1988; O'Shea et al., 1989; 1991; Ellenberger et al., 1992). The coiled-coil motif consists of  $\alpha$ -helices wrapped around each other with a left-handed superhelical twist. In general, coiled coils contain either two (Crick, 1953) or three (Pauling and Corey, 1953) parallel  $\alpha$ -helices.

Coiled-coil sequences contain hydrophobic and hydrophilic amino-acid residues in a repeating, heptad pattern (denoted positions a through g). Hydrophobic residues tend to occur at positions a and d of the heptad repeat, and these residues form the interface between helices. This feature is known as the "4-3 hydrophobic repeat" and is a hallmark of coiled-coil sequences (Hodges et al., 1972; McLachlan and Stewart, 1975).

We evaluated the coiled-coil propensity for the amino-acid sequences of proteins with known three-dimensional structures. Each 28-residue segment was scored (Lupas et al., 1991) according to the statistical preference of different residues for a specific position in the heptad (Parry, 1982). This analysis (see Experimental Procedures) revealed a sequence in HA2 that has an unusually high score (Fig. 2). Surprisingly, this high-scoring region does not correspond to the known three-stranded coiled coil of HA, but to the adjacent, extended loop region (Fig. 1; yellow).

A closer examination of the HA2 sequence revealed a continuous, 88-residue sequence with the 4-3 hydrophobic repeat characteristic of coiled coils (Fig. 3). The sequence begins at the N-terminus of the short  $\alpha$ -helix, includes the loop and continues through the long  $\alpha$ -helix of the coiled coil (Fig. 1; red, yellow and magenta). The heptad repeat is maintained, in register, throughout the 88-residues. This sequence is highly conserved among different strains of influenza virus (reviewed in Wiley and Skehel, 1987).

These considerations lead to the hypothesis that HA2 folds into a long, three-stranded coiled coil in the fusogenic state (Fig. 4): the existing three-stranded coiled coil (magenta) is extended to include the loop region (yellow) and the short, external  $\alpha$ -helix (red). Thus, the 80 Å coiled coil that exists in the native state forms a 135 Å coiled coil in the fusogenic state.

A protein conformational change of the magnitude proposed here, while rare, is not unprecedented. A striking example, demonstrated by x-ray crystallography, involves the serpin family of protease inhibitors. Upon cleavage of the reactive center peptide bond, a segment of these proteins can undergo a dramatic conformational change in secondary and tertiary structure, resulting in the relocation of a short region by 70 Å (Wright et al., 1990; Stein et al., 1990; Mottonen et al., 1992). In addition, a conformational change with aspects similar to our model for HA has been proposed recently for the monomer to trimer transition of yeast heat shock factor (Rabindran et al., 1993).

## Results

In order to evaluate this model, we studied a 36-residue peptide called LOOP-36, which corresponds to the 28-residue loop region, plus 8 residues of the long  $\alpha$ -helix (Fig. 3; Fig. 5A). Most short peptides do not fold into stable structures in aqueous solution (for reviews, see Wright et al., 1988; Kim and Baldwin, 1990), although coiled-coil peptides containing four or five heptad repeats are a notable exception (Hodges et al., 1981; O'Shea et al., 1991). Thus, a stringent test of the model is to determine if the LOOP-36 peptide forms a three-stranded coiled coil.

As determined by circular dichroism (CD) spectroscopy, LOOP-36 is highly helical (>90% at 100  $\mu$ M) at pH 4.8, although it is unfolded at neutral pH (Fig. 5B). CD experiments also reveal a reversible thermal unfolding transition for LOOP-36 at pH 4.8, but there is no stable structure at neutral pH (Fig. 5C). Sedimentation equilibrium experiments indicate that LOOP-36 is trimeric at pH 4.7, but primarily monomeric at pH 7.2 (Fig. 5D). We conclude that LOOP-36 folds into a trimeric,  $\alpha$ -helical coiled coil at the pH of membrane fusion.

There is a sharp transition in the  $\alpha$ -helical CD signal of LOOP-36 between pH 5 and pH 7 (Fig 5E). The predominance of acidic residues at particular heptad positions in the LOOP-1 sequence (see Fig. 2) is reminiscent of the arrangement seen in the Fos leucine zipper homodimer, which displays a similar pH-dependence of stability (O'Shea et al., 1992). Protonation of acidic sidechains at low pH alleviates electrostatic repulsion that destabilizes the folded conformation at neutral pH. There is a second structural transition in LOOP-36 helicity at even lower pH, but the conformation at pH 2 probably is not relevant for membrane fusion.

At 37° C, however, LOOP-36 is unstable, even at pH 5. In addition, as mentioned earlier, the fusogenic conformation of HA can be induced at neutral pH, with elevated temperature. We therefore suspected that coiled-coil structure in the loop region would be stabilized in the context of the longer coiled coil. To test this notion, we studied a 52-residue peptide, LOOP-52, which begins at the N-terminus of the short  $\alpha$ -helix and ends at the C-terminus of LOOP-36 (Fig. 3; Fig. 6A).

LOOP-52 forms a fully helical (~100%) structure at pH 7.0 and pH 4.8 (Fig. 6B). At both neutral pH and pH 4.7, LOOP-52 is a trimer, as determined by equilibrium sedimentation (Fig. 6C). The trimeric, helical structure in LOOP-52 is very stable: at pH 7.0, LOOP-52 unfolds with a transition midpoint ( $T_m$ ) of 52 °C, and at pH 4.8, the  $T_m$  is 72 °C (Fig. 6D).

## Comparison with Other Results

It is extremely unusual for a small protein fragment, in aqueous solution, to fold into a stable structure which is different from that found in the native protein. Our results: (i) demonstrate that the loop region can fold as a three-stranded coiled coil, (ii) provide strong support for the proposal that, in the fusogenic state, the coiled coil includes the external  $\alpha$ -helix in addition to the loop region and (iii) indicate that the longer coiled coil is very stable, even at neutral pH.

Consistent with our model for the conformation of the fusogenic state, HA is trimeric at pH 5 (Doms and Helenius, 1986). In addition, electron microscopic studies suggest that an extended fibrous structure is formed on the surface of influenza viral

membranes following treatment at low pH, or elevated temperatures at neutral pH (Ruigrok et al., 1986; 1988).

Earlier proteolysis experiments provide strong biochemical support for our model. Although certain regions of the fusogenic state of HA are sensitive to proteolysis, much of the HA2 subunit is resistant to degradation (Skehel et al., 1982; Ruigrok et al., 1988). The final tryptic digestion product of HA2 starts at residue 40, and that obtained with thermolysin starts at residue 38. In both cases, the residues of the short  $\alpha$ -helix, loop region and the long  $\alpha$ -helix remain intact. The native state of HA is resistant to proteolysis. Nonetheless, if HA2 maintained the helical-hairpin structure in the fusogenic state, one would expect the loop region to be sensitive to proteolysis upon dissociation and degradation of the HA1 subunits. In contrast, the loop region would be protected from proteolysis if it were part of a folded structure, such as the coiled coil in our model.

### **A Mechanism for the Conformational Change**

One hypothesis for the mechanism of the conformational change is that a decrease in pH shifts the thermodynamic equilibrium between the native and fusogenic states. It is striking that the following pH transitions coincide: (i) LOOP-36 helicity (and LOOP-52 stability), (ii) the known pH dependences for the conformational changes and (iii) the onset of membrane fusion activity in HA (see Fig. 7). All of these transitions occur abruptly between pH 7 and pH 5. If the mechanism for the conformational change is under thermodynamic control, it should be possible to design fusion mutants that alter the stability of the coiled coil, and hence the fusogenic state, without affecting the stability of the native state.

For several reasons, however, we favor a kinetically controlled mechanism for the conformational change in HA, in which the native state of HA is metastable (i.e., although it is stably folded, it has the potential to form a thermodynamically more stable state). First, the conformational change, induced either by a decrease in pH or an increase in temperature at neutral pH, is irreversible (Ruigrok et al., 1986; for reviews, see Wiley and Skehel, 1987 and Stegmann and Helenius, 1993). Second, since LOOP-52 is very stable, even at neutral pH (Fig. 6C), we expect that the 88-residue coiled coil of the fusogenic conformation will be extremely stable. Third, there is a good correlation between the

decrease in stability of the native state and the increase in the pH of fusion by the HA mutants (Ruigrok et al., 1986).

In addition, previous biophysical studies provide strong evidence for a mechanism that is under kinetic control. Circular dichroism experiments reveal an irreversible thermal transition at ~63° C, correlated with the temperature of HA-mediated fusion at neutral pH, followed by a second, reversible transition at even higher temperature (Ruigrok et al., 1986; see, however, Ruigrok et al., 1988). The irreversibility of the first transition is consistent with a transition from a metastable initial state to a more stable final state. We would assign the second transition to the unfolding of the long, stable coiled coil.

The notion of a metastable structure raises the question: how does HA fold into a conformation that is not the thermodynamically most stable structure? The *in vivo* folding of HA0 (Braakman et al., 1992) is known to be dependent on metabolic energy. A minority of the HA0 molecules fail to fold into the native conformation, but instead form stable, aberrant trimer species that do not leave the endoplasmic reticulum (Gething et al., 1986b; Copeland et al., 1986). These species may have an extended coiled coil structure similar to our proposed fusogenic conformation. Thus, the native conformation of HA may be the result of the *in vivo* folding pathway of the HA0 precursor. In addition, the relative instability of the loop region of HA2 at neutral pH (see Figs. 5C and 6C) may be important for *in vivo* folding of HA0.

Regardless of the mechanism, there appear to be two types of interactions that stabilize the native state and thereby inhibit formation of the stable fusogenic state: (i) intersubunit HA1 protein-protein interactions and (ii) burial of the fusion peptide in the hydrophobic core of the trimer. As mentioned previously, in the fusogenic conformation, the HA1 subunits dissociate and the fusion peptide becomes exposed (see, however, Stegmann and Helenius, 1990).

In the structure of the native state (Wilson et al., 1981), extensive interactions are apparent between the loop region of HA2 and its corresponding HA1 subunit (Fig. 1B). Thus, the HA1 subunit may act as an inhibitor or “clamp” that binds to the loop region, preventing the conformational change. In addition, multiple HA1-HA1 interactions are likely to prevent the conformational change. Indeed, introducing disulfide bonds between

HA1 subunits inhibits the conformational change in HA and prevents membrane fusion (Godley et al., 1992). Furthermore, many mutations that alter the pH of fusion appear to destabilize HA1-HA1 or HA1-HA2 interactions in the native state (reviewed in Wiley and Skehel, 1987).

The native conformation may also be stabilized by the fusion peptide, which makes significant hydrophobic interactions in the core of the native structure. The buried fusion peptide resembles a hydrophobic “hook,” which holds the helices together in a helical-hairpin conformation (Fig. 1A). Many of the HA mutants that fuse under less acidic conditions (Daniels et al., 1985; Gething et al., 1986a) contain amino acid substitutions near the region of the fusion peptide. For example, the x-ray crystal structure of the HA2 fusion mutant, D112G, reveals that 4 hydrogen bonds are lost (per monomer) between the aspartate side chain and residues of the fusion peptide. The pH of fusion for the mutant is elevated by 0.4 pH units, presumably because loss of the hydrogen bonds destabilizes the native state (Weis et al., 1990b).

### **Speculation on the Relevance to Other Fusion Events**

Although influenza hemagglutinin is the best characterized membrane fusion protein, other viral (and non-viral) fusion proteins are known to be oligomeric (reviewed in White, 1992), and some of these are candidates for further study in light of our model. Like HA, these glycoproteins are derived proteolytically from a precursor; this proteolysis creates a new N-terminus with an adjacent fusion peptide sequence (for reviews, see White, 1990; Stegmann and Helenius, 1993). Moreover, several of these fusion peptide sequences are followed by a stretch of amino acids with a 4-3 hydrophobic repeat, suggesting a coiled-coil structure (Chambers et al., 1990). Our results suggest that some of these 4-3 hydrophobic repeats may correspond to a coiled-coil structure that is found only in the fusogenic conformation of the glycoprotein.

Proteins that mediate fusion at neutral pH (for example those involved in fertilization, protein trafficking or HIV infection; reviewed in White, 1992) may also unleash a latent structural conformation in response to an external stimulus, such as binding of a receptor or interaction with another protein. For example, the binding of HIV to CD4 is known to induce conformational changes in the envelope glycoprotein that result in



release of the extracellular subunit (gp120) from the transmembrane subunit (gp41), exposure of the fusion peptide, and fusion of the viral membrane with the cell (reviewed in Vaishnav and Wong-Staal, 1991). A peptide from gp41, corresponding to a sequence adjacent to the N-terminal fusion peptide, has recently been shown to form a coiled coil (Wild et al., 1992). Moreover, expression of gp41 in the absence of gp120 results in syncytium formation (Perez et al., 1992), demonstrating that the interaction between gp120 and gp41 inhibits fusogenic activity, possibly in a manner similar to our proposed inhibitory interaction of the HA1 and HA2 subunits in the native state of HA.

Finally, if our model is correct, it may lead to new approaches for the discovery of anti-viral drugs. For example, therapeutic agents might be designed to prevent the low-pH activation of HA. Alternatively, since acid pretreatment of influenza virus can abolish infectivity (Doms et al., 1985; Stegmann et al., 1990; Puri et al., 1990), agents might be designed that cause the conformational change to occur prematurely. It is also interesting, in light of our results, that a synthetic coiled-coil peptide from gp41 of HIV inhibits viral replication (Wild et al., 1992).

#### **Note added in Proof**

Recently, a connection was made between fusion proteins involved in cellular vesicle transport and vesicle-synaptosomal membrane fusion (Söllner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J. E. (1993). SNAP receptors implicated in vesicle targetting and fusion. *Nature* 362, 318-324). The amino-acid sequences for some of these membrane proteins contain a long, continuous region of 4-3 hydrophobic repeats.

#### **Acknowledgments**

We thank Rheba Rutkowski and Mike Burgess for synthesis of the LOOP-36 and LOOP-52 peptides; Brad Stewart for the computer-aided sequence search; Jeff Stock for the "coiledcoil" computer program; Shiufun Cheung for figure 4; Fred Hughson for help with the HA literature and for in-depth discussions; Jonathan Weissman for discussion; Pehr Harbury and Tom Alber for discussions about oligomerization states of coiled coils; Stan Watowich for the compilation of HA sequences for various strains of influenza A; and members of the Kim lab and several colleagues for helpful comments on earlier

versions of the manuscript. C. M. C. is supported by an N. I. H. Training Grant (AI07348). P. S. K. is a Pew Scholar in the Biomedical Sciences. This research was supported by the Howard Hughes Medical Institute.

## **Experimental Procedures**

### **Coiled Coil Prediction**

An algorithm (Parry, 1982) was used to predict coiled-coil propensities based on the statistical preference of different amino acids for each position in the heptad repeat. Sequences from the Protein Data Bank (Brookhaven) were analyzed using a computer program ("coiledcoil"; Lupas et al., 1991) based on the algorithm. With a window size of 28 residues, sequences with scores above 1.3 are thought to have a good probability for coiled-coil structure, while scores between 1.1 and 1.3 generally correspond to amphipathic  $\alpha$ -helices in globular proteins (Lupas et al., 1991). A score of 1.6 was obtained for residues 54-81 of HA2.

Inspection of the complete HA2 sequence revealed a continuous heptad repeat from residues 38 to 125. Although an earlier analysis had predicted a coiled coil structure for much of the 88-residue sequence of HA2 identified here, the register of the heptad repeat was not maintained throughout this sequence, and several interruptions in the coiled coil were also predicted (Ward and Dopheide, 1980; see also, Chambers et al., 1990).

### **Peptide Synthesis and Purification**

Peptides were synthesized on an Applied Biosystems model 430A peptide synthesizer using Fmoc chemistry with Fastmoc reaction cycles modified to include acetic anhydride capping (Fields et al., 1991). LOOP-36 corresponds to residues 54-89 and LOOP-52 corresponds to residues 38-89 of HA2 from the X-31 strain of influenza virus (Kilbourne, 1969). In each peptide, the N-terminus is acetylated and the C-terminus is amidated. The peptides were cleaved using standard Fmoc protocols and desalted on a Sephadex G-10 or G-25 column (Pharmacia) in 5% acetic acid. Final purification was by reverse-phase, high-performance liquid chromatography (HPLC, Waters, Inc.) at 25° C using a Vydac preparative or semipreparative C18 column (stock# 218TP1022 or 218TP510, respectively). A linear acetonitrile-H<sub>2</sub>O gradient of 0.1% buffer B increase per minute was used with a flow rate of 10 ml/min (5 ml/min for the semipreparative column). Buffer A is 0.1% trifluoroacetic acid (TFA) in water; buffer B is 90% acetonitrile and 0.1% TFA in water. The identity of the peptides was confirmed by laser desorption mass spectrometry (Finnigan MAT LASERMAT). The measured mass for LOOP-36 was 4451 Da (expected 4450 Da), and the measured mass for LOOP-52 was 6147 Da (expected 6146).

### **Circular Dichroism Spectroscopy**

Circular dichroism (CD) spectroscopy (for a review, see Woody, 1985) was performed on an AVIV CD spectrophotometer (model 62DS) equipped with a thermoelectric temperature controller. The cuvettes used for thermal unfolding studies and pH studies were 1 cm and 0.1 mm in pathlength and the cuvettes used for wavelength spectra were 1 mm and 0.1 mm in pathlength. Peptide concentration was determined by tyrosine absorbance at 275.5 nm (Edelhoch, 1967) in 6 M guanidinium chloride (Schwarz/Mann Biotech, Ultra-Pure grade).

The pH-dependence of LOOP-36 structure was determined by monitoring the CD signal at 222nm. Measurements were made at 0° C and 32 μM peptide in a buffer of 150 mM NaCl, 20 mM (each) of sodium borate, sodium citrate and sodium phosphate. Both transitions observed in the pH-dependence studies are >95% reversible (data not shown).

For thermal unfolding experiments, the CD signal was monitored at 222nm as a function of temperature. Samples at pH 7.0 contained 32 μM LOOP-36, 150 mM NaCl and 10 mM sodium phosphate; or 500 μM LOOP-52, 50 mM sodium phosphate and 150 mM NaCl. Samples at pH 4.8 also contained 10 mM sodium citrate and 10 mM sodium borate for both LOOP-36 and LOOP-52. All thermal unfolding experiments are reversible (>95% for LOOP-36 samples and >85% for LOOP-52 samples) in the temperature range from 0-60° C for LOOP-36, and 0-85° C for LOOP-52 (data not shown).

CD spectra were obtained at 0° C, in the same conditions as used in thermal unfolding experiments, except that the peptide concentration was 100 μM for LOOP-36. In Fig. 7, percent helicity for LOOP-36 was calculated assuming that 100% helicity corresponds to  $-33,000 \text{ deg cm}^2 \text{ dmol}^{-1}$ , as has been found in studies of helical peptides of this length (Chen et al., 1974)

### **Equilibrium Sedimentation**

Molecular weights were determined at 1° C for LOOP-36 and 4° C for LOOP-52, by analytical ultracentrifugation (reviewed in Laue et al., 1992) with a Beckman XL-A Optima Analytical Ultracentrifuge equipped with absorbance optics. An An-60Ti rotor was used at 22,000 and 27,000 rpm. Experiments were performed at pH 7.2 and pH 4.7. LOOP-36 samples were made at three concentrations (300, 100 and 33 μM) per pH and LOOP-52 samples were made at six concentrations (500, 167, 56, 130, 43 and 14 μM) per pH, and all samples were dialyzed exhaustively (~24 hours) against buffer prior to the experiments. Buffer conditions were 150 mM NaCl, 50 mM sodium phosphate, pH 7.2; or 150 mM NaCl, 5 mM sodium acetate, pH 4.7. The molecular weights were determined by a simultaneous fit of each data-set using a non-linear least squares fit algorithm, HID-4000 (Johnson et al., 1981); no systematic residuals were observed.

## References

- Braakman, I., Helenius, J., and Helenius, A. (1992). Role of ATP and disulphide bonds during protein folding in the endoplasmic reticulum. *Nature* 356, 260-262.
- Brand, C. M., and Skehel, J. J. (1972). Crystalline antigen from the influenza virus envelope. *Nature New Biology* 238, 145-147.
- Chambers, P., Pringle, C. R., and Easton, A. J. (1990). Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *J. Gen. Virol.* 71, 3075-3080.
- Chen, Y.-H., Yang, J. T., and Chau, K. H. (1974). Determination of the helix and  $\beta$  form of proteins in aqueous solution by circular dichroism. *Biochemistry* 13, 3350-3359.
- Cohen, C., and Parry, D. A. D. (1990).  $\alpha$ -Helical coiled coils and bundles: how to design an  $\alpha$ -helical protein. *Proteins* 7, 1-15.
- Copeland, C. S., Doms, R. W., Bolzau, E. M., Webster, R. G., and Helenius, A. (1986). Assembly of influenza hemagglutinin trimers and its role in intracellular transport. *J. Cell Biol.* 103, 1179-1191.
- Crick, F. H. C. (1953). The packing of  $\alpha$ -helices: simple coiled coils. *Acta. Cryst.* 6, 689-697.
- Daniels, R. S., Downie, J. C., Hay, A. J., Knossow, M., Skehel, J. J., Wang, M. L., and Wiley, D. C. (1985). Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell* 40, 431-439.
- Doms, R. W., Helenius, A., and White, J. (1985). Membrane fusion activity of the influenza virus hemagglutinin. *J. Biol. Chem.* 260, 2973-2981.
- Doms, R. W., Gething, M.-J., Henneberry, J., White, J., and Helenius, A. (1986). Variant influenza virus hemagglutinin that induces fusion at elevated pH. *J. Virology.* 57, 603-613.
- Doms, R. W., and Helenius, A. (1986). Quaternary structure of influenza virus hemagglutinin after acid treatment. *J. Virol.* 60, 833-839.
- Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* 6, 1948-1954.
- Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992). The GCN4 basic-region-leucine zipper binds DNA as a dimer of uninterrupted  $\alpha$ -helices: crystal structure of the protein-DNA complex. *Cell* 71, 1223-1237.
- Fields, C. G., Lloyd, D. H., Macdonald, R. L., Otteson, K. M., and Noble, R. L. (1991). HBTU activation for automated Fmoc solid-phase peptide synthesis. *Peptide Research* 4, 95-101.

- Gething, M. -J., Doms, R. W., York, D., and White, J. (1986a). Studies on the mechanism of membrane fusion: site specific mutagenesis of the hemagglutinin of influenza virus. *J. Cell Biol.* 102, 11-23.
- Gething, M. -J., McCammon, K., and Sambrook, J. (1986b). Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport. *Cell* 46, 939-950.
- Glick, G. D., Toogood, P. L., Wiley, D. C., Skehel, J. J., and Knowles, J. R. (1991). Ligand recognition by influenza virus. *J. Biol. Chem.* 266, 23660-23669.
- Godley, L., Pfeifer, J., Steinhauer, D., Ely, B., Shaw, G., Kaufmann, R., Suchanek, E., Pabo, C., Skehel, J. J., Wiley, D. C., and Wharton, S. (1992). Introduction of intersubunit disulfide bonds in the membrane-distal region of the influenza virus hemagglutinin abolishes membrane fusion activity. *Cell* 68, 635-645.
- Hodges, R. S., Sodak, J., Smillie, L. B., and Jurasek, L. (1972). Tropomyosin: amino acid sequence and coiled-coil structure. *Cold Spring Harbor Symp. Quant. Biol.* 37, 299-310.
- Hodges, R. S., Saund, A. K., Chong, P. C. S., St.-Pierre, S. A., and Reid, R. E. (1981). Synthetic model for two-stranded  $\alpha$ -helical coiled-coils. *J. Biol. Chem.* 256, 1214-1224.
- Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* 36, 575-588.
- Kim, P. S., and Baldwin, R. L. (1990). Intermediates in the folding reactions of small proteins. *Ann. Rev. Biochem.* 59, 631-660.
- Kilbourne, E. D. (1969). Future influenza vaccines and the use of genetic recombinants. *Bull. Wrl. Hlth. Org.* 41, 643-645.
- Klenk, H.-D., Rott, R., Orlich, M., and Blödorn, J. (1975). Activation of influenza A virus by trypsin treatment. *Virology* 68, 426-439.
- Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759-1764.
- Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Royal Society of Chemistry, Cambridge (editors, Harding, S. E., Rowe, A J. and Horton, J. C.), pp. 90-125.
- Lazarowitz, S. G., Compans, R. W., and Choppin, P. W. (1971). Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. *Virology* 46, 830-843.

- Lazarowitz, S. G., and Choppin, P. W. (1975). Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. *Virology* 68, 440-454.
- Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* 252, 1162-1164.
- McLachlan, A. D., and Stewart, M. (1975). Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *J. Mol. Biol.* 98, 293-304.
- Mottonen, J., Strand, A., Symersky, J., Sweet, R. M., Danley, D. E., Geoghegan, K. F., Gerard, R. D., and Goldsmith, E. J. (1992). Structural basis of latency in plasminogen activator inhibitor-1. *Nature* 355, 270-273.
- O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989). Evidence that the leucine zipper is a coiled coil. *Science* 243, 538-542.
- O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel, coiled coil. *Science* 254, 539-544.
- O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* 68, 699-708.
- Parry, D. A. D. (1982). Coiled-coils in  $\alpha$ -helix-containing proteins: analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. *Biosci. Rep.* 2, 1017-1024.
- Pauling, L., and Corey, R. B. (1953). Compound helical configurations of polypeptide chains: structure of proteins of the  $\alpha$ -keratin type. *Nature* 171, 59-61.
- Perez, L. G., O'Donnell, M. A., and Stephens, E. B. (1992). The transmembrane glycoprotein of human immunodeficiency virus type 1 induces syncytium formation in the absence of the receptor binding glycoprotein. *J. Virol.* 66, 4134-4143.
- Puri, A., Booy, F. P., Doms, R. W., White, J. M., and Blumenthal, R. (1990). Conformational changes and fusion activity of influenza virus hemagglutinin of the H2 and H3 subtypes: effects of acid pretreatment. *J. Virol.* 64, 3824-3832.
- Rabindran, S. K., Haroun, R. I., Clos, J., Wisniewski, J., and Wu, C. (1993). Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. *Science* 259, 230-234.
- Rott, R., Orlich, M., Klenk, H.-D., Wang, M. L., Skehel, J. J., and Wiley, D. C. (1984). Studies on the adaptation of influenza virus to MDCK cells. *EMBO J.* 3, 3329-3332.
- Ruigrok, R. W. H., Martin, S. R., Wharton, S. A., Skehel, J. J., Bayley, P. M., and Wiley, D. C. (1986). Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. *Virology* 155, 484-497

- Ruigrok, R. W. H., Aitken, A., Calder, L. J., Martin S. R., Skehel, J. J., Wharton, S. A., Weis, W., and Wiley, D. C. (1988). Studies on the structure of the influenza virus hemagglutinin at the pH of membrane fusion. *J. Gen. Virol.* *69*, 2785-2795.
- Sauter, N. K., Glik, G. D., Crowther, R. L., Park, S. -J., Eisen, M. B., Skehel, J. J., Knowles, J. R., and Wiley, D. C. (1992). Crystallographic detection of a second ligand binding site in influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA* *89*, 324-328.
- Skehel, J. J., and Waterfield, M. D. (1975). Studies on the primary structure of the influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA* *72*, 93-97.
- Skehel, J. J., Bayley, P. M., Brown, E. B., Martin, S. R., Waterfield, M. D., White, J. M., Wilson, I. A., and Wiley, D. C. (1982). Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc. Natl. Acad. Sci. USA* *79*, 968-972.
- Spruce, A. E., Iwata, A., White, J. M., and Almers, W. (1989). Patch-clamp studies of single cell-fusion events mediated by a viral fusion protein. *Nature* *342*, 555-558.
- Stegmann, T., White, J. M., and Helenius, A. (1990). Intermediates in influenza induced membrane fusion. *EMBO J.* *9*, 4231-4241.
- Stegmann, T., Delfino, J. M., Richards, F. M., and Helenius, A. (1991). The HA2 subunit of influenza hemagglutinin inserts into the target membrane prior to fusion. *J. Biol. Chem.* *266*, 18404-18410.
- Stegmann, T., and Helenius, A. (1993). Influenza Virus Fusion: From models toward a mechanism. *Viral Fusion Mechanisms*, J. Bentz, eds. (CRC Press: Boca Raton, Florida).
- Stein, P. E., Leslie, A. G. W., Finch, J. T., Turnell, W. G., McLaughlin, P. J., and Carrell, R. W. (1990). Crystal structure of ovalbumin as a model for the reactive centre of serpins. *Nature* *347*, 99-102.
- Steinhauer, D. A., Wharton, S. A., Skehel, J. J., Wiley, D. C., and Hay, A. J. (1991). Amantadine selection of a mutant influenza virus containing an acid-stable hemagglutinin glycoprotein: evidence for virus-specific regulation of the pH of glycoprotein transport vesicles. *Proc. Natl. Acad. Sci. U S A* *88*, 11525-11529.
- Vaishnav, Y. N., and Wong-Staal, F. (1991). The biochemistry of AIDS. *Ann. Rev. Biochem.* *60*, 577-630.
- Ward, C. W., and Dopheide, T. A. (1980). Influenza virus hemagglutinin. Structural predictions suggest that the fibrillar appearance is due to the presence of a coiled-coil. *Aust. J. Biol. Sci.* *33*, 441-447.
- Weis, W., Brown, J. H., Cusack, S., Paulson, J. C., Skehel, J. J., and Wiley, D. C. (1988). Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature* *333*, 426-431.
- Weis, W. I., Brünger, A. T., Skehel, J. J., and Wiley, D. C. (1990a). Refinement of the influenza virus hemagglutinin by simulated annealing. *J. Mol. Biol.* *212*, 737-761.

- Weis, W. I., Cusack, S. C., Brown, J. H., Daniels, R. S., Skehel, J. J., and Wiley, D. C. (1990b). The structure of a membrane fusion mutant of the influenza virus haemagglutinin. *EMBO J.* *9*, 17-24.
- Wharton, S. A., Skehel, J. J., and Wiley, D. C. (1986). Studies of influenza haemagglutinin-mediated membrane fusion. *Virology* *149*, 27-35.
- White, J., Helenius, A., and Gething, M.-J. (1982). Haemagglutinin of influenza virus expressed from a cloned gene promotes membrane fusion. *Nature* *300*, 658-659.
- White, J. M., and Wilson, I. A. (1987). Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the Influenza virus hemagglutinin. *J. Cell Biol.* *105*, 2887-2896.
- White, J. M. (1990). Viral and cellular membrane fusion proteins. *Ann. Rev. Physiol.* *52*, 675-697.
- White, J. M. (1992). Membrane fusion. *Science* *258*, 917-924.
- Wild, C., Oas, T., McDanal, C., Bolognesi, D., and Matthews, T. (1992). A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. *Proc. Natl. Acad. Sci. USA* *89*, 10537-10541.
- Wiley, D. C., and Skehel, J. J. (1987). The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Ann. Rev. Biochem.* *56*, 365-394.
- Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* *289*, 366-373.
- Woody, R. W. (1985). Circular Dichroism of Peptides. *The Peptides*. vol. 7 Udenfriend, S., and Meienhofer, J. and Hruby, V. J., eds (Academic Press, Inc: Orlando, Florida) 15-114.
- Wright, P. E., Dyson, J., and Lerner, R. A. (1988). Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding. *Biochemistry* *27*, 7167-7175.
- Wright, H. T., Qian, H. X., and Huber, R. (1990). Crystal structure of plakalbumin, a proteolytically nicked form of ovalbumin. *J. Mol. Biol.* *213*, 513-528.

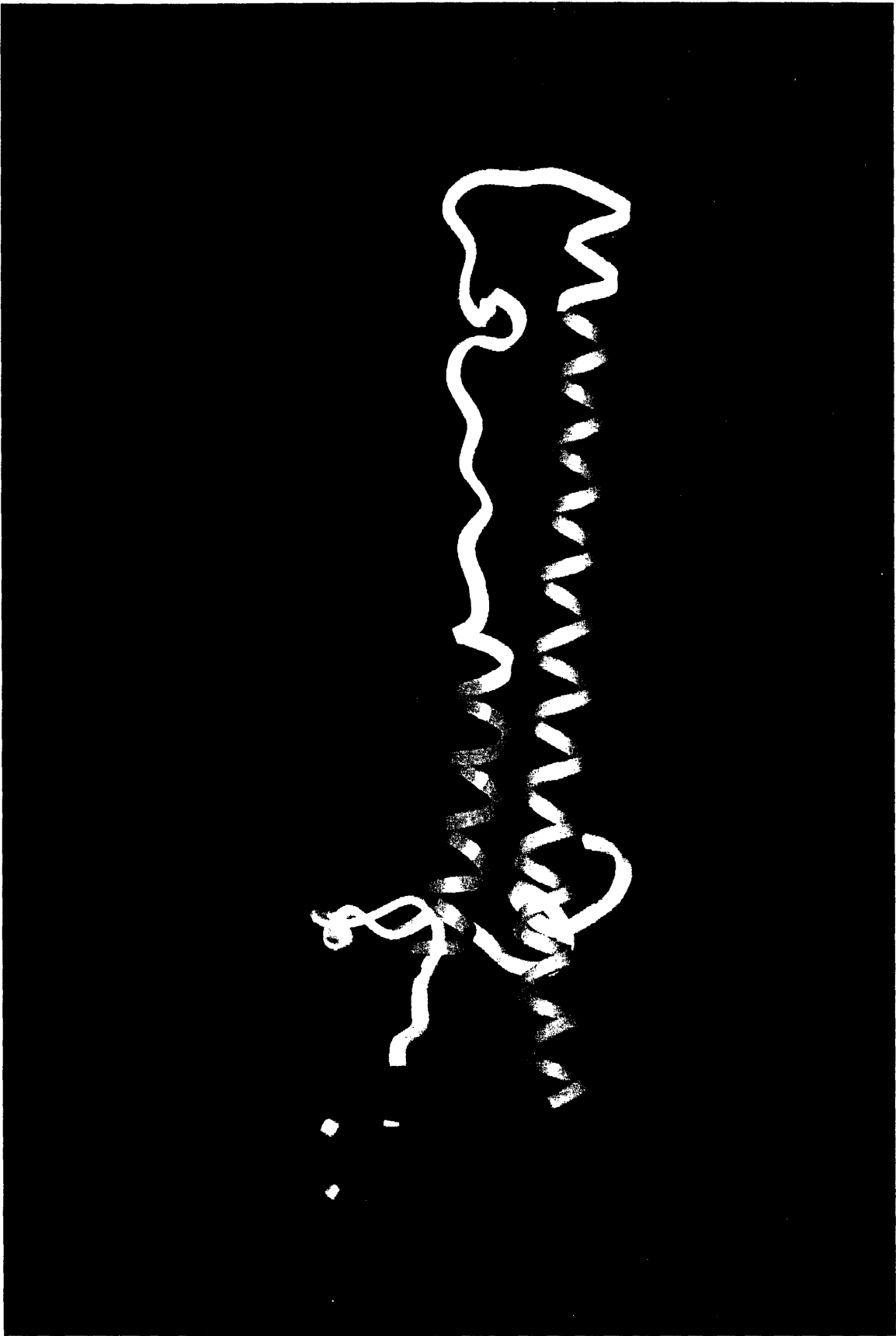


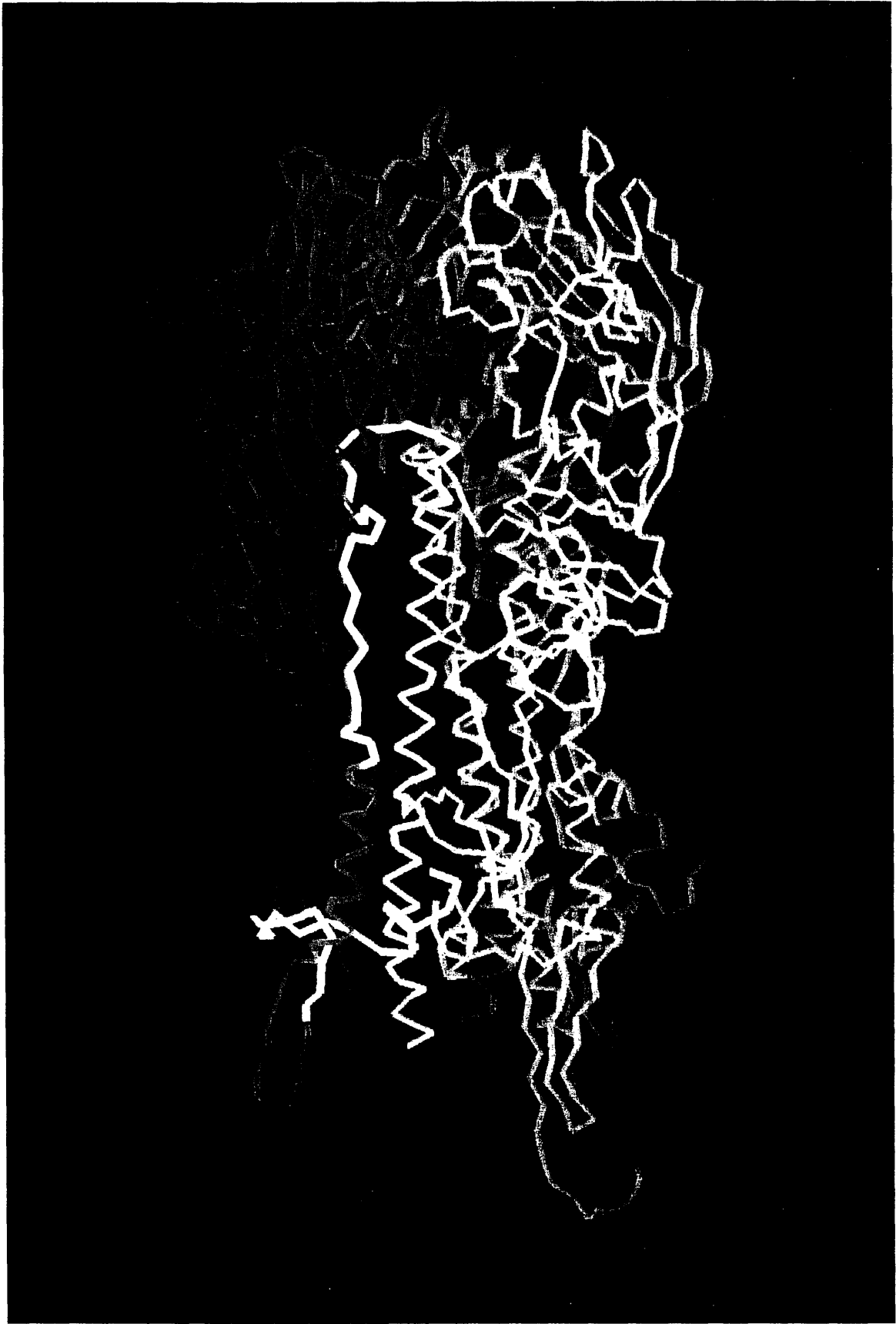
**Figure 1** The Helical-Hairpin Structure in Hemagglutinin (HA).

The computer graphics representations are based on the neutral-pH crystal structure of bromelain-cleaved HA ectodomain, called BHA (Wilson et al., 1981; Weis et al., 1990a).

**(A)** A ribbon representation of the helical-hairpin structure in the transmembrane subunit (HA2) of HA. The C-terminal residues of HA2 are omitted for clarity. The N-terminal fusion peptide (light blue; residues 1-25) resembles a "hook," which is followed by a short antiparallel  $\beta$ -sheet (purple; residues 26-34). The hook fastens two helices together in a helical-hairpin structure, which contains a short  $\alpha$ -helix (red; residues 35-53) connected to a long  $\alpha$ -helix (magenta; residues 82-125) by the loop region (yellow; residues 54-81). The long  $\alpha$ -helix associates with the corresponding long  $\alpha$ -helices of two other HA2 subunits, forming a three-stranded coiled coil (not shown).

**(B)** A backbone-trace representation of the interaction of the HA1 domain with the HA2 helical hairpin. The globular, sialic acid-binding domain (HA1) is at the top of the molecule. The viral membrane would be located at the bottom, and the target membrane of the endosome would be located at the top of the figure. The distance from the viral membrane to the distal tip of HA is 135 Å. The helical-hairpin structure of one HA2 subunit is colored as in (A); the C-terminal residues of HA2 are omitted, for clarity. The corresponding HA1 subunit is colored dark blue. The other HA monomers (purple) assemble to form a trimer, which buries the fusion peptides in the core of the coiled-coil interface.

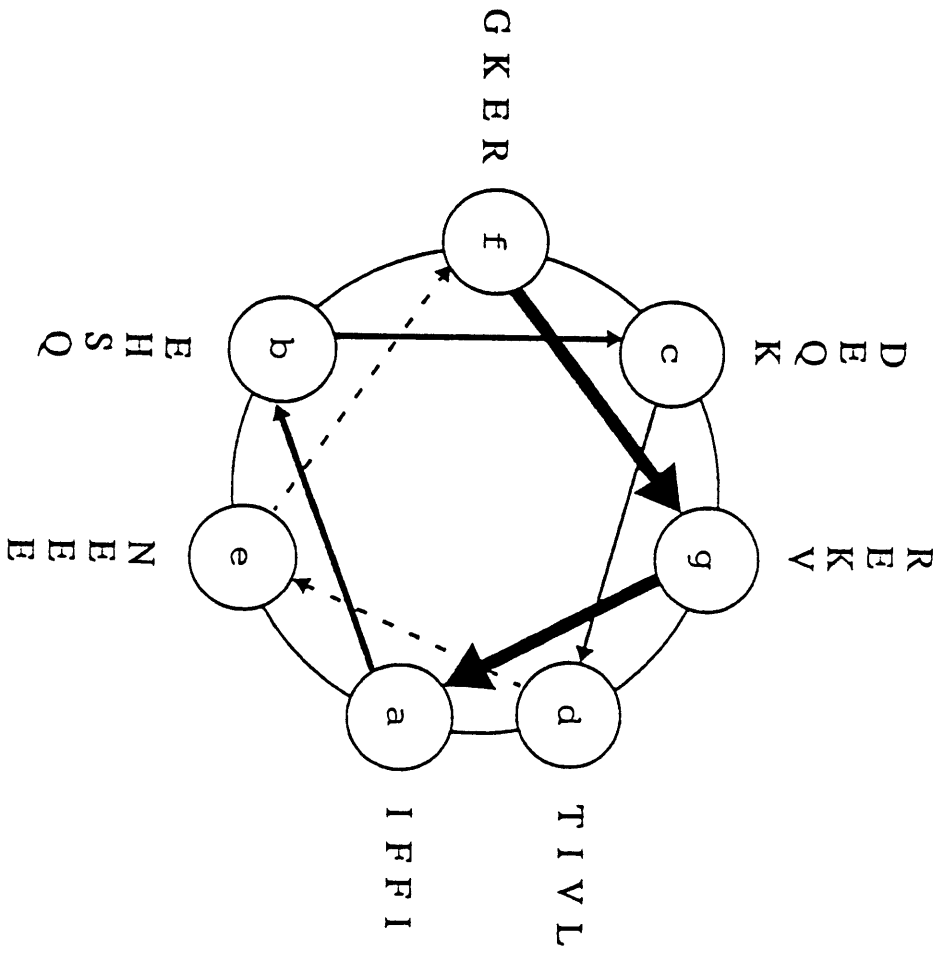




**Figure 2** The Loop Region of the Helical Hairpin as a Coiled Coil

The 28-residues with high coiled coil propensity (see Experimental Procedures) are projected onto a helical wheel. The view is down the helical axis starting at the N-terminus (position **f**). The alignment of predominantly hydrophobic amino acids at the **a** and **d** positions and the predominance of hydrophilic amino acids at other positions reveals the strikingly amphipathic nature of the sequence. Amino acid abbreviations: A = alanine; D = aspartic acid; E = glutamic acid; F = phenylalanine; G = glycine; H = histidine; I = isoleucine; K = lysine; L = leucine; N = asparagine; Q = glutamine; R = arginine; S = serine; T = threonine; V = valine; Y = tyrosine.

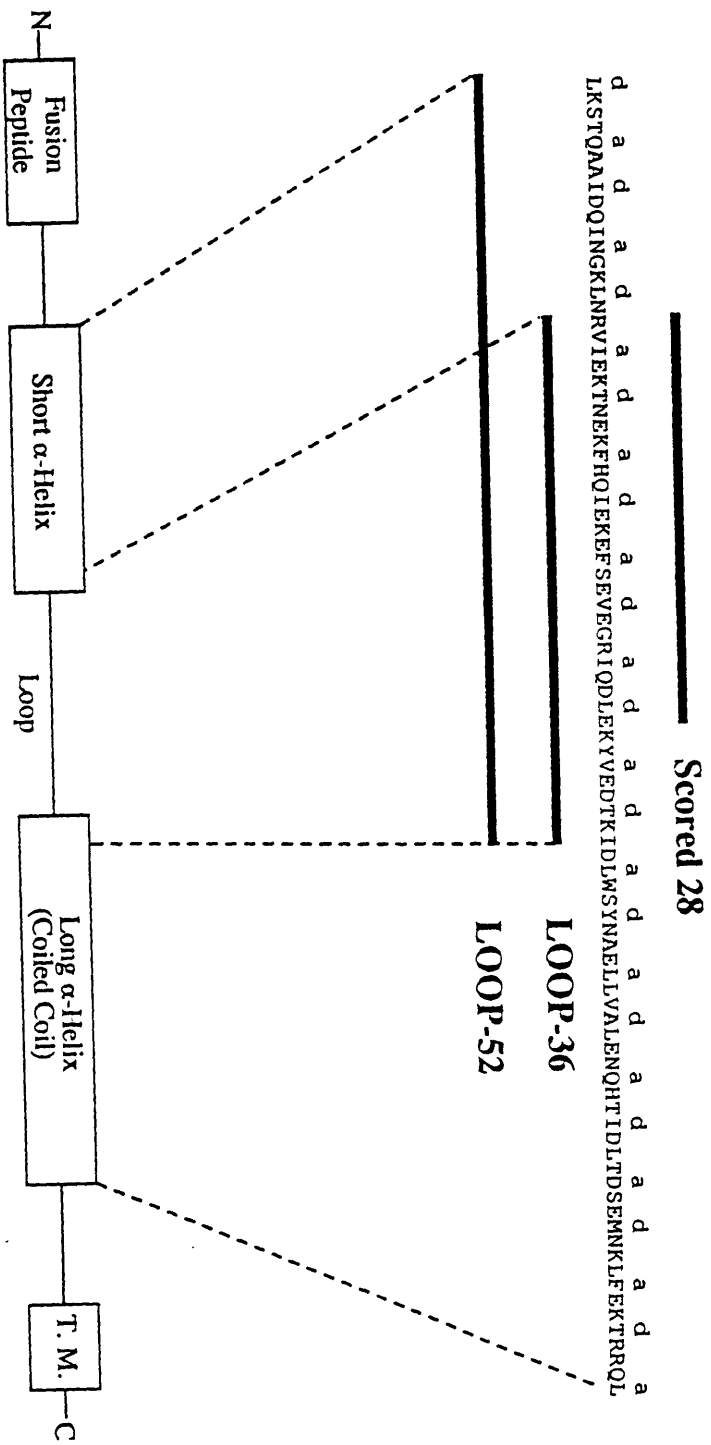
Figure 2



**Figure 3** Position of the LOOP-36 and LOOP-52 Peptides in HA2

A schematic diagram of the structural units in the native structure of HA2. The 88 amino-acid sequence (residues 38-125), proposed to form an extended coiled coil in the fusogenic state, is shown with the **a** and **d** positions of the heptad repeat indicated above the sequence. The high-scoring 28 residue region (residues 54-81) is indicated by a bar (Scored 28). The sequences of the LOOP-36 (residues 54-89) and LOOP-52 (residues 38-89) peptides are underlined. LOOP-36 contains the entire loop region (residues 54-81) plus part of the sequence of the long  $\alpha$ -helix. LOOP-52 contains the entire short  $\alpha$ -helix plus the residues of LOOP-36. The fusion peptide and transmembrane region (T.M.) are indicated.

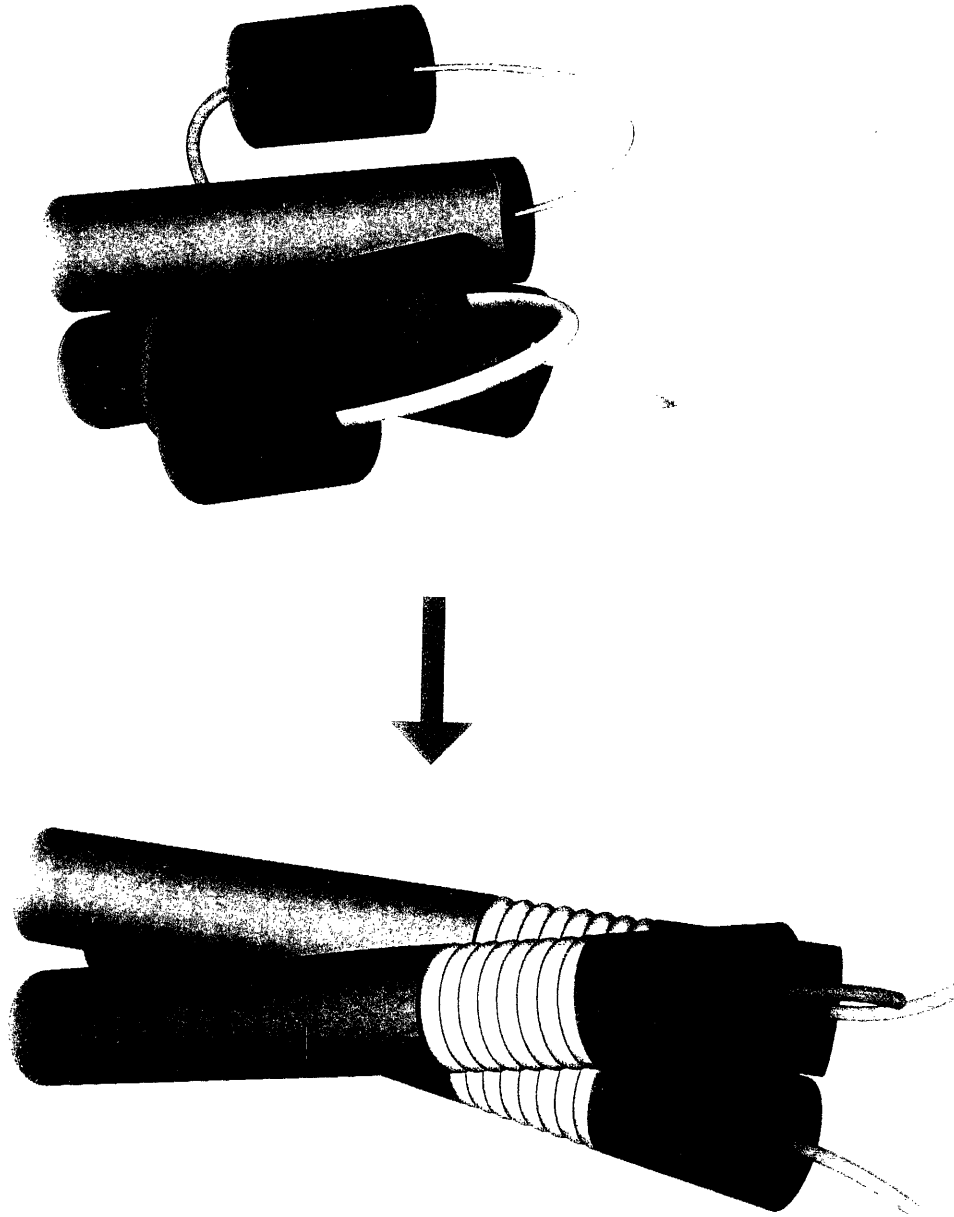
**Figure 3**



**Figure 4** A Model for the Fusogenic State of HA

A trimer of HA2 molecules in the native state (left) is depicted schematically, as a stem of three helical hairpins (red, yellow and magenta), each fastened by the N-terminal, fusion peptide hook (blue), which is buried in the core of the coiled coil. In the fusogenic state (right), the HA1 subunits (not shown) dissociate from the stem (see text), the fusion peptide is released from the protein interior, the loop "springs" into a helical conformation, and the helices associate to form an extended coiled coil that relocates the fusion peptide 100 Å toward the target membrane to promote membrane fusion.





**Figure 5** Folding of LOOP-36 as a Helical Trimer

LOOP-36 at neutral pH (○) and low pH (●).

(A) LOOP-36 as a coiled coil. The sequence of the LOOP-36 peptide is projected onto a helical wheel, as in Fig. 2.

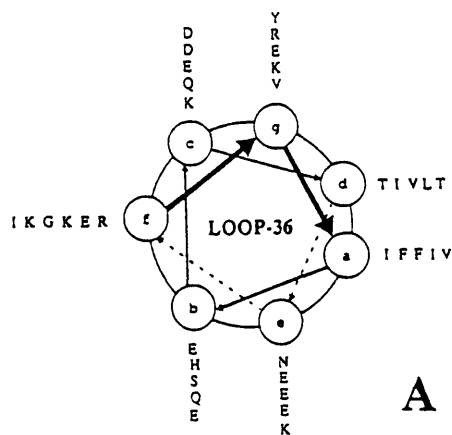
(B) Circular dichroism (CD) spectroscopy at 0 °C indicates a characteristic  $\alpha$ -helical spectrum (>90% helix content based on the value of  $[\theta]_{222}$ ) for LOOP-36 at pH 4.8. The CD spectrum for LOOP-36 at pH 7.0 indicates a random-coil conformation.

(C) The CD signal at 222nm for LOOP-36 shows a cooperative, thermal-unfolding transition at pH 4.8. No transition is seen for LOOP-36 at pH 7.0.

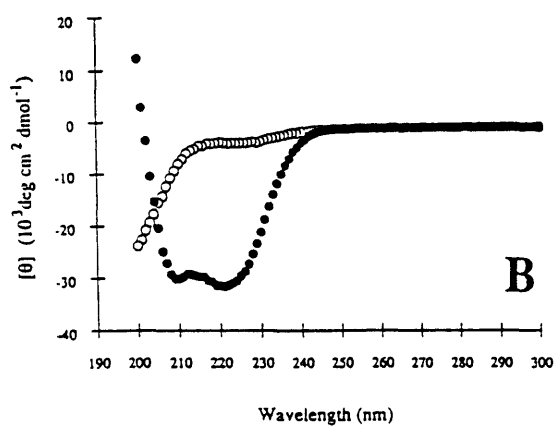
(D) The molecular weight of LOOP-36 as determined by sedimentation equilibrium experiments. Molecular weights were determined for samples at several peptide concentrations (see Experimental Procedures). For LOOP-36, the average molecular weight at pH 4.7 was 13.6 kD (expected for trimer = 13.4 kD), while the measured molecular weight at pH 7.2 was 5.2 kD (expected for monomer = 4.5 kD).

(E) The pH-dependence of helix content in LOOP-36. The CD signal at 222nm (▲) is monitored by circular dichroism (CD) spectroscopy at 0 °C as a function of pH. There is a sharp transition from an unfolded peptide at pH 7 to a helical, folded structure at pH 5. A second transition is seen at lower pH.

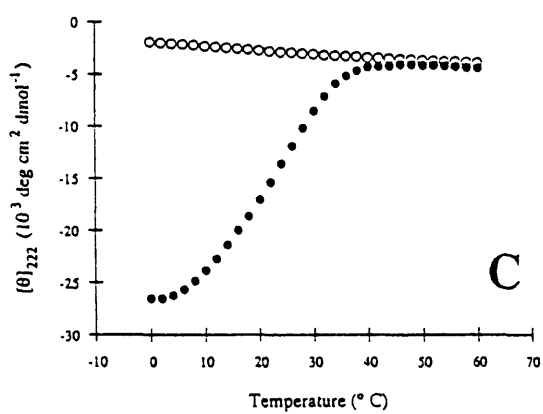
Figure 5



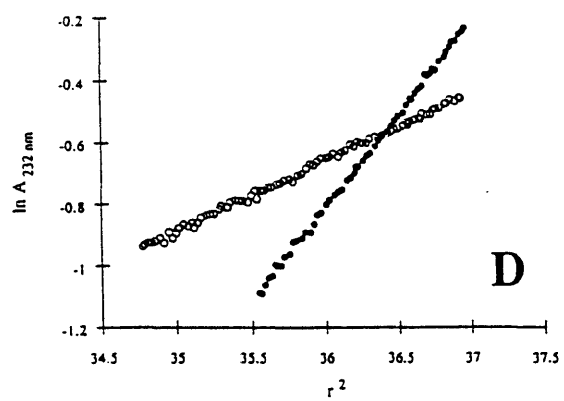
A



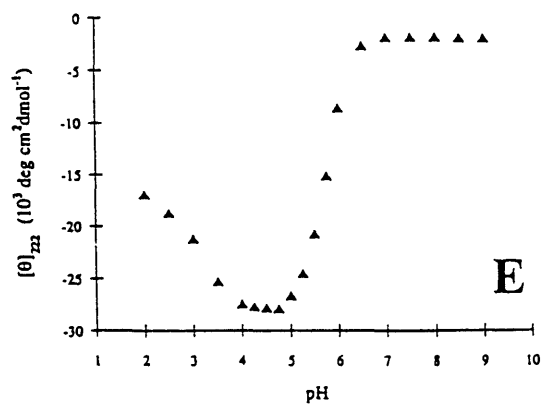
B



C



D



E

**Figure 6** Folding of LOOP-52 as a Helical Trimer

LOOP-52 at neutral pH (□) and low pH (■).

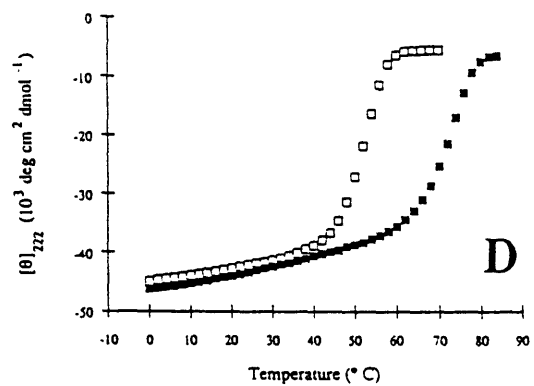
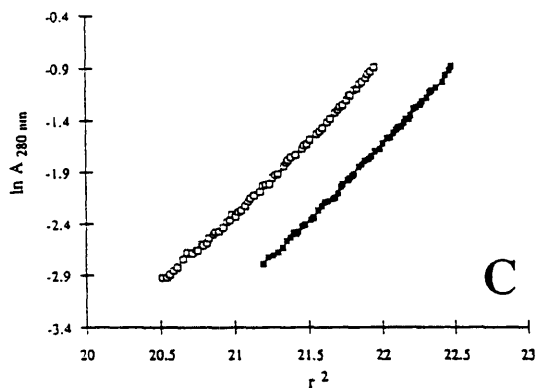
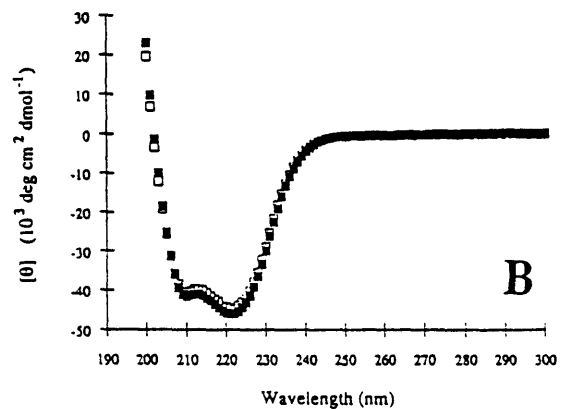
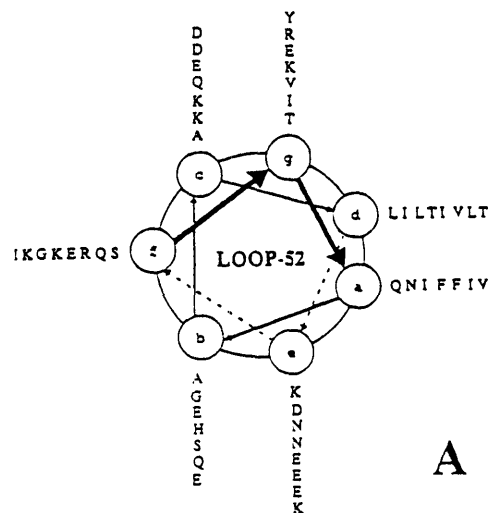
(A) LOOP-52 as a coiled coil. The sequence of the LOOP-52 peptide is projected onto a helical wheel, as in Fig. 2.

(B) Circular dichroism (CD) spectroscopy studies of LOOP-52 (0 °C) indicate characteristic  $\alpha$ -helical spectra (~100% helix content) at pH 4.8 and at pH 7.0.

(C) The molecular weight of LOOP-52 as determined by sedimentation equilibrium experiments. Molecular weights were determined for samples at several peptide concentrations (see Experimental Procedures). The average molecular weights for LOOP-52 were 18.4 kD at pH 4.7, and 19.5 kD at pH 7.2 (expected for trimer = 18.4 kD).

(D) The CD signal at 222nm for LOOP-52 shows a cooperative, reversible, thermal-unfolding transition, at both pH 4.8 and pH 7.0.

Figure 6



**Figure 7** The LOOP-36 Folding Transition Occurs in a Physiologically Relevant pH Range.

Normalized CD data for LOOP-36 ( $\blacktriangle$ ; from Fig. 5D) is superimposed on published results for the pH transitions of HA conformational change and membrane fusion activity. The transition at  $\sim$ pH 5.7 agrees well with other pH transitions reported for HA.

A sharp transition is seen between pH 7 and pH 5 for viral-cellular membrane fusion ( $\square$ ) (Wharton et al., 1986), exposure of a fusion peptide epitope to a monoclonal antibody

( $\dagger$ ) (White et al., 1987), protease sensitivity of BHA ( $\circ$ ) (Ruigrok et al., 1988),

liposome binding by BHA ( $\times$ ) (Doms et al., 1985), dissociation of HA1 subunits and

exposure of an epitope to a monoclonal antibody ( $\triangle$ ) (White et al., 1987), and cell fusion

by BHA ( $\diamond$ ) (Doms et al., 1985).

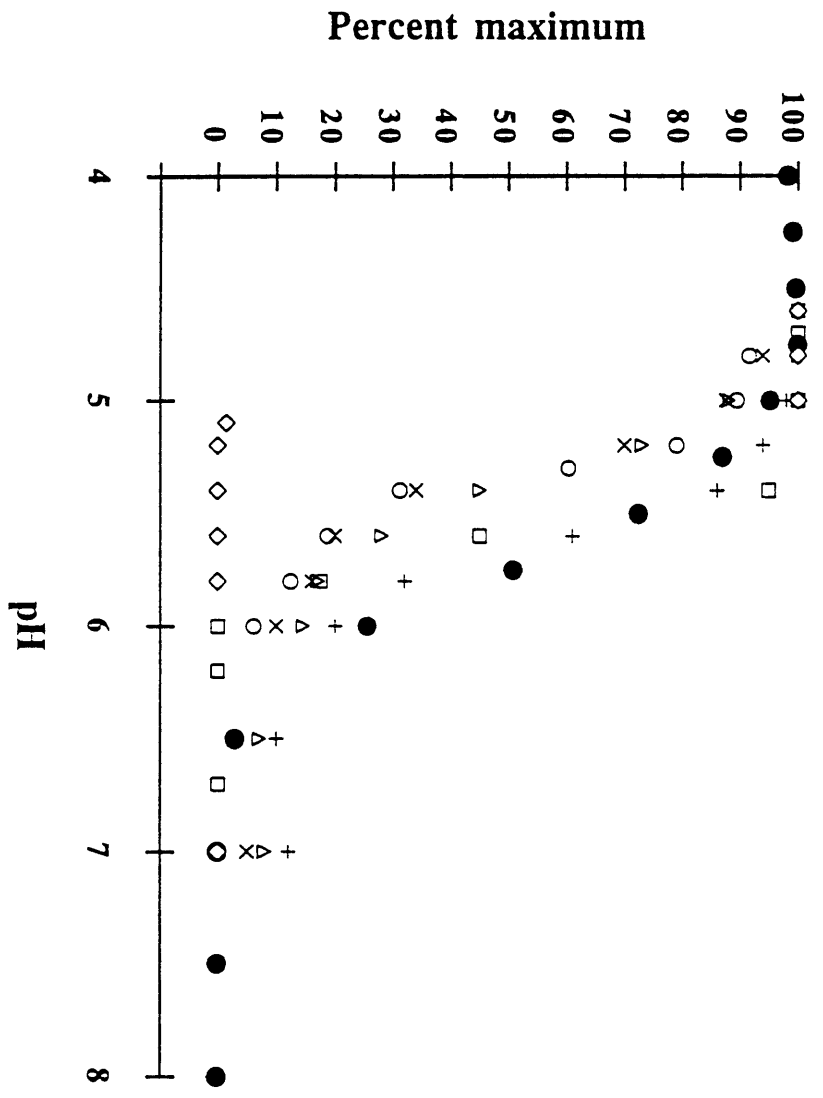


Figure 7

## Chapter 3



## **Evidence that Influenza Hemagglutinin is Trapped in the Native Conformation**

Enveloped viruses, such as influenza virus, human respiratory syncytia virus (hRSV) and human immunodeficiency virus (HIV), enter cells by protein-mediated membrane fusion. For influenza virus, membrane fusion is regulated by the conformational state of the HA (hemagglutinin) protein, which switches from the native structure to a fusogenic conformation when exposed to the acidic environment of the cellular endosome. Extensive proteolysis of HA in the acid-induced conformation results in a stable fragment (TBHA2), revealed by biochemical and crystallographic methods to be an extended, coiled-coil structure that relocates the essential fusion peptides 100Å toward the endosomal membrane (Carr and Kim, 1993; Bullough et al., 1994). Here we demonstrate that destabilization of the native structure at neutral pH, either with heat or the denaturant urea, triggers the conformational change to a fusogenic state with a structure and function that is indistinguishable from the acid-induced conformation. For each of the variables, pH, temperature or urea, the onset of the conformational change is coincident with induction of membrane fusion. These results provide strong support for the notion that in its native conformation, HA is trapped in a metastable state, and that the more stable fusogenic conformation, represented by the TBHA2 structure, is released by destabilization of the trap. This mechanism may be shared by other enveloped viruses, including those that enter the cell at neutral pH, such as HIV.

## Introduction

Efforts to discover the mechanism of membrane fusion have focused on hemagglutinin (HA), the envelope protein required for entry of influenza virus into its host cell (for a review, see Wiley and Skehel, 1990). The membrane-fusion activity of HA is controlled by a conformational switch. HA in the native conformation is not fusion active, but facilitates viral attachment to the host cell. Once the virus has been taken into the cell by endocytosis, the acidic pH of the cellular endosome triggers a conformational change in HA from the native to the fusogenic conformation. Although the structures of the HA ectodomain in its native (Wilson et al., 1981) and acid-induced conformations (Bullough et al., 1994) are known, the mechanism by which acid triggers this conformational change is poorly understood.

Previously we hypothesized that even at neutral pH, the fusogenic conformation is the most stable state, but that the native structure acts as a trap, holding native HA in a metastable conformation (i.e., although it is stably folded, it has the potential to form a thermodynamically more stable state; Carr and Kim, 1993), which is less stable than the fusogenic conformation. In support of this hypothesis, peptide studies indicated that the extended coiled-coil structure of the fusogenic conformation is highly stable, even at neutral pH (Carr and Kim, 1993). In this model, the role of acid in the conformational change is not specific; rather, the mildly acidic pH of the endosome acts to destabilize the metastable, native structure, thereby allowing the conformational change to the more stable, fusogenic state.

There are few examples of proteins that fold into a metastable conformation. Denaturation and renaturation studies have revealed a long-lived, metastable state for two proteins: the serpin family of protease inhibitors (Carrell et al., 1991; see however Carrell et al, 1993) and N\*, a folding intermediate of bovine pancreatic trypsin inhibitor (BPTI; Weissman and Kim, 1991). A few other examples may exist, if less rigorous criteria are accepted as evidence for a metastable conformation (for a review, see Baker and Agard, 1994).

For the metastable states of the serpins and BPTI, it has been demonstrated that moderate concentrations of the denaturant urea induce a conformational change to a second, stably folded state, which persists after removal of the denaturant. If acid acts to induce the HA conformational change by destabilizing the metastable native state, then any moderately denaturing conditions should destabilize the native structure and induce the identical conformational change, independent of pH. To test this prediction, we used two methods to destabilize native HA at neutral pH: heat and urea. Exposure of virus to either heat or the denaturant urea results in the same HA conformational change as that induced by acid, as judged by conformational and functional assays. These results provide strong evidence to support the hypothesis that HA is metastable in the native conformation and suggest that the major role of low pH, *in vivo*, is to destabilize the native state, converting HA to the fusogenic conformation. The mechanism of destabilizing a metastable conformation to induce membrane fusion may be shared by viruses that enter the cell at neutral pH, such as HIV.

## Results

The structure and function of the HA envelope protein are best described in terms of its component subunits: HA1 and HA2. Inside the host cell, HA is synthesized from the viral genome and assembled as a trimer of precursor polypeptides (HA0). Each HA0 is then proteolytically cleaved to form the disulfide-bonded subunits (HA1 and HA2) that form one monomer of the native HA trimer. HA1 is extracellular and responsible for attachment of the virus to the cell by binding to sialic-acid components of the host cell glycoproteins and glycolipids. HA2 is the transmembrane domain, which contains at its N-terminus the essential fusion-peptide region. Previous results suggest that only HA2 is necessary for membrane fusion: removal of HA1 (by proteolysis or disulfide reduction) does not inhibit membrane fusion, but proteolysis of HA2 abolishes fusion activity (Schoch et al., 1992).

### Protease Assay for the HA Conformational Change

A protease assay is used to distinguish between the two conformations of HA on the surface of intact influenza A/Beijing/32/92/X-117 (H3N2). In the native conformation HA is resistant to proteinase K digestion, but in the acid-induced conformation HA is

digested to specific, protease-resistant fragments (Fig. 1). Virus samples containing native HA are incubated for 15 minutes at pH 5 to allow for the conformational change. To stop the transition, samples are returned to neutral pH. Thus, in each experiment, extensive proteolysis by proteinase K is initiated under the same conditions (see methods). N-terminal sequence analysis of HA2 reveals that the 27 amino-acid residues of the fusion-peptide region are removed by proteolysis of the acid-induced conformation.

### **Denaturants Induce the Conformational Change at Neutral pH**

At neutral pH the conformational change is induced by heat or the denaturant urea. Proteolytic digestion of HA in the heat and urea-induced conformation results in the same, specific proteolytic pattern as that observed for the acid-induced conformation (Fig. 1; 65 °C and 4.5 M urea). The stable coiled-coil core of the acid-induced conformation (Bullough et al., 1994) is protected from proteolysis, and the same core is protected in the heat and urea-induced conformation. In all three cases, HA2 begins at residue 28 and extends to at least cysteine 138, which is disulfide-bonded to fragments of HA1. Disappearance of native HA is concurrent with appearance of the stable core (Fig. 2), which persists to temperatures of at least 70 °C or urea concentrations of at least 5M. Treatment at 100 °C, however, results in denatured HA that is sensitive to proteolytic digestion. In addition, a peptide corresponding to the coiled-coil core is stably folded at temperatures as high as 70 °C or urea concentrations as high as 6 M, confirming the stability of this conformation, even at neutral pH (Fig. 3).

### **Membrane Fusion is Induced by Denaturants at Neutral pH**

In order to detect HA in the fusogenic conformation, we monitored the membrane-fusion activity of virus under conditions that induce the HA conformational change. Membrane fusion is detected using a fluorescence energy transfer assay (Struck et al., 1981). As a result of membrane fusion, the lipid bilayer of the virus mixes with the lipid bilayer of the target vesicle, uncoupling the resonance energy transfer between fluorescent donor and acceptor molecules that are covalently attached to lipids in the target vesicle. Loss of energy transfer is monitored as an increase in donor fluorescence, allowing membrane fusion to be assayed by fluorescence spectroscopy as a function of time.

The onset of the HA conformational change is coincident with induction of membrane fusion. The pH of membrane-fusion activity and the pH of the HA-conformational change are known to co-vary (Wharton et al., 1986). We observe fusion at pH 6.0 with a maximum at pH 5.8 (Fig. 4a), which coincides with the onset of the conformational change, between pH 6.0 and pH 5.8 (Fig. 2a). Heat-induced membrane fusion is induced at neutral at 58 °C, with optimal fusion at 60 °C (Fig. 4b), in good agreement with the onset of the conformational change, observed at 58 °C (Fig 2b). At neutral pH, optimal fusion is induced by 3.5 M urea. Fusion is also observed at 3.0 M urea, although the rate is significantly reduced (Fig. 4c). The onset of the conformational change is ~ 3.25 M urea (Fig. 2c), which correlates well with induction of membrane fusion by urea. Fusion at neutral pH occurs with rates and extents comparable to those of acid-induced membrane fusion. Furthermore, the correlation between optimal conditions for fusion and the conformational change provides strong evidence that, like acid, the denaturants heat and urea induce the conformational change to the fusogenic species.

In the absence of target membranes, acidic pH is known to inactivate the membrane fusion function of HA (White et al., 1982; see also Ramalho-Santos et al., 1993; Weber et al., 1994; Wharton et al., 1995). In agreement with this, we observe inhibition of membrane fusion for virus pre-incubated at low pH (Fig. 4a). Likewise, at neutral pH, pre-heating virus (to either 58 or 60 °C) inactivates membrane fusion (Fig. 4b). Pre-incubating the virus in urea at neutral pH also abolishes membrane fusion (Fig. 4c). Thus, the conditions that induce the conformational change and membrane fusion also lead to viral inactivation, in the absence of target membranes. While the mechanism of viral inactivation is poorly understood, the simplest model to explain the available data is a transition from the native state (N) to the fusogenic state (F), which induces membrane fusion in the presence of target membranes and causes viral inactivation in the absence of target membranes (Fig. 5).

Native HA is trapped in a metastable conformation. At neutral pH the denaturants heat and urea each induce the conformational change to a specific, stable, fusogenic state. Because the native conformation is trapped in a metastable state, the primary role of acid in the mechanism of the conformational change *in vivo* is as a denaturant, inducing the conversion to the fusogenic state by destabilization of the metastable state.

The existence of a metastable state for HA raises an intriguing question: How does HA fold into a metastable conformation? The answer probably lies in the *in vivo* synthesis, folding and processing of HA. Trapping HA in a metastable conformation may be achieved in one of two ways: 1. HA0 folds into its most stable conformation, and maturation cleavage of HA0 into the HA1 and HA2 subunits renders native HA metastable. Consistent with this theory, it is known that the newly created N- and C-termini are ~ 22 Å apart in the native structure, indicating a substantial rearrangement after maturation cleavage of HA0 (Wilson et al., 1981). 2. Chaperonin-mediated folding, *in vivo*, assists in folding HA0 into a metastable state (see discussion, Carr and Kim, 1993). Whether or not HA0 folds into a metastable or stable structure remains to be tested.

## Implications

Destabilization of a metastable native state to induce a fusogenic conformation is a mechanism that may be shared by other enveloped viruses, including those that enter cells at neutral pH, such as HIV. Enveloped viruses have evolved two different routes of entry into their host cell: some viruses fuse with the plasma membrane at neutral pH while others fuse with the cellular endosome in response to a decrease in pH. Variations in the pH optima for membrane fusion might imply that viruses which fuse at the cell surface use a mechanism different from the acid-induced mechanism used by viruses that fuse with the endosome. Our results demonstrate that the conformational change of HA to a fusogenic state is pH-independent, in spite of the fact that influenza fuses with the endosome, *in vivo*. We conclude that the mechanism of membrane fusion for influenza virus is by acid-destabilization of native HA, which is trapped in a metastable state. Because destabilization of a metastable envelope protein could be achieved at neutral pH by binding to a receptor on the cell surface or by proteolysis of the envelope protein, this mechanism need not be restricted to influenza HA or acid-induced fusion events.

## Acknowledgments:

Thanks go to Charu Chaudhry for her work on the proteinase K assay; Mike Milhollen for preparing rHA2-X117; Dr. James McKnight for help with fluorescence assays; Dr. Pete Petillo for help with carboxyfluorescein recrystallization and lipid lyophilization; Drs. Toon Stegmann, Jim Lear and Joe Bentz for helpful discussions about

the fusion assay; Dr. Thomas Weber for helpful discussions about fusion and proteolysis of influenza virus, Parke-Davis, Rochester Operations, especially Dr. Armen Donabedian for his enthusiasm and generosity in supplying flu virus and split flu; Dr. Michael Shaw and Dr. Xiyan Xu, Influenza Branch, Center for Disease Control, for sequence information for the X-117 strain; and Christina Doetsch for technical assistance with virus and antibodies.

## **Experimental Procedures**

### **Materials:**

Dioleoyl phosphatidylcholine, N-(7-nitro-2-1,3-benzoxadiazol-4-yl) Diacyl-phosphatidylethanolamine ("NBD") and Lissamine-rhodamine-phosphatidylethanolamine ("Rho") were purchased from Avanti Polar Lipids, Inc., Alabaster AL. Bovine ganglioside GD1a was purchased from Calbiochem-NovaBiochem Inc., La Jolla, CA. 5 (6)-carboxy-fluorescein was purchased from Kodak Co., Rochester, NY. Purified Influenza A/Beijing/32/92 (H3N2) in 35% sucrose was a generous gift from Parke-Davis, Rochester, MI. Trypsin and the Inorganic Phosphorus Assay kit were purchased from Sigma, St. Louis, MO. Organic solvents were purchased from J. T. Baker Inc., Phillipsburg, NJ. The sequences of HA1 and HA2 from the X-117 strain were provided by Dr. Xiyan Xu, Influenza Branch, National Center for Infectious Disease, Center for Disease Control (Atlanta, GA).

### **Methods:**

#### **Design, expression and purification of rHA2-X117**

A 95 amino-acid peptide model of the coiled-coil core of the acid-induced conformation of HA (rHA2-X117) was designed to include the long helix of HA2, as seen in the x-ray crystal structure of TBHA2 (Bullough et al., 1994). The sequence of the peptide rHA2-X117 begins at residue G33 and ends at residue R127 of HA2 from Influenza A/Beijing/32/92/X-117 (H3N2). A synthetic gene of rHA2 from HA of the X-31 strain of Influenza A/Hong Kong/1968/X-31 (H3N2) was prepared with optimal codon usage for *E. coli* (Grosjean and Fiers, 1982), ligated into the NdeI/BamHI site of the high-copy expression plasmid (pAED4; Doering, 1992) using standard procedures (Sambrook et al., 1989) and named prHMG. For the HA2 sequence from Influenza A/Beijing/32/92/X-117 (H3N2), a double amino-acid substitution (V55L and R124K) was made from prHMG using single-stranded DNA mutagenesis (Kunkel, 1985) and named prHMGX117.

The recombinant plasmids were used to transform *Escherichia coli* strain BL21 (DE3) pLysS and induced to express protein at 0.8 O.D. by addition of IPTG to 400  $\mu$ M, using the T7 expression system (Studier et al., 1990). Cells expressing the peptides were lysed in 50 mM TRIS, pH 8.0, 200 mM KCl, 0.2 mM EDTA and 20% glycerol by sonication on ice to visual clarity. Concentrated HCl was added to pH 2, and the precipitate was pelleted by centrifugation at 12,000 rpm for 20 minutes in a Sorvall SS-34 rotor. The supernatant fraction was adjusted to pH 8.0, centrifuged to remove debris, and dialyzed against 10 mM Tris-HCl, pH 8, 150 mM sodium chloride. The precipitate from dialysis was centrifuged, and the supernatant fraction (40ml) was loaded onto a 40 ml bed-

volume, DEAE sephadex gravity column and eluted with a linear salt gradient from 0 M to 1 M sodium chloride in 10 mM Tris, 4 M urea, pH 8.0. Fractions taken every 4 minutes were monitored at 280 nm with an absorbance detector connected to a chart recorder, and by SDS-PAGE to identify peaks containing the peptide. Fractions containing the peptide were pooled, dialyzed against 5% acetic acid, and further purified by high performance liquid chromatography on a Vydac C-18, reverse-phase column in 0.1% trifluoroacetic acid (TFA) with a linear water/acetonitrile gradient from 10-45% B at 4.5 % min<sup>-1</sup> and 45-48% B at 0.1% min<sup>-1</sup> (where B is 90% acetonitrile, 0.1% TFA) at a 10 ml min<sup>-1</sup> flow rate. The identity of the peptide was confirmed to be within one mass unit, by laser desorption mass spectrometry (Finnigan LASERMAT): rHA2-X117=11,101.3 Da, expected 11,100 Da. Peptide concentration was determined by tyrosine absorbance at 275.5 nm with a molar extinction coefficient of 8, 250 M<sup>-1</sup> cm<sup>-1</sup> (Edelhoch, 1967) in 6 M guanidinium chloride (Schwarz/Mann Biotech, Ultra-Pure grade).

### **Circular Dichroism of Peptides**

Circular dichroism (CD) spectroscopy was performed on an AVIV CD spectrometer (model 62DS) equipped with a thermoelectric temperature controller. Samples were prepared at 10 μM rHA2-X117 in 50 mM sodium phosphate pH 7.3, 150 mM sodium chloride or in the same buffer at the indicated concentration of urea. Samples were incubated 15 minutes at the indicated temperature or urea concentration, and the molar ellipticity was recorded. Transition midpoints do not represent equilibrium unfolding midpoints.

### **Equilibrium Sedimentation and Gel Filtration**

The molecular weight of rHA2-X117 was determined at 20 °C by analytical ultracentrifugation (reviewed in Laue et al., 1992) using a Beckman XL-A Optima Analytical Ultracentrifuge equipped with absorbance optics. An An-60Ti rotor was used at 17,000 and 20,000 rpm. Experiments were performed over a 10-fold concentration range (12.5-125 μM), at pH 7.3. The molecular weight was determined by a simultaneous fit of each data set using a non-linear least squares fit algorithm, HID-4000 (Johnson et al., 1981). In 50 mM sodium phosphate, 150 mM sodium chloride the mass was 34,091 Da (expected 33,330 Da), consistent with an oligomerization state of 3 monomers per folded unit.

### **Proteolysis with Proteinase K**

In the native conformation, HA is resistant to proteolysis. In the fusogenic conformation, HA1 is digested, and extensive proteolysis causes removal of the fusion peptide from HA2 (Skehel et al., 1982; Ruigrok et al., 1988). Concentrated samples of 100 μl Influenza A/Beijing/32/92/X-117 (1mg/ml HA) were prepared on ice and incubated for 15 min under experimental conditions. The pH of each sample was then raised to ~ pH 7.5 with 0.5 M Tris pH 8. All samples were then diluted to 1.2 ml and digested with proteinase K under identical, aqueous conditions at neutral pH (1:1 ratio of proteinase K to HA, by weight) for 1-4 hours at room temperature.

Virus proteins and protein fragments were separated by 12% and 14% acrylamide gels under non-reducing and reducing conditions, respectively. Proteins were either stained with coomassie brilliant blue R250, or transferred to nitrocellulose for detection by immunoblot analysis, using polyclonal antisera 238-3, which was raised against a keyhole-limpet hemocyanin-conjugate (Pierce, Rockford, IL) of LOOP-36 (a fragment of the HA2 domain, Carr and Kim, 1993). Identity of the HA2 fragment was confirmed by 5 cycles of



N-terminal amino acid sequence analysis: intact HA2 begins at residue 1 (GIFGA) and the HA2 proteolytic fragments generated by the conformational change with acid, heat, or urea, begin at residue 28 (NSEGT). We find that proteolysis with stoichiometric concentrations of trypsin for 30 minutes does not lead to a conformationally specific HA2 fragment (not shown), whereas proteinase K results in complete removal of the fusion peptide region of HA2, allowing for detection of a conformationally specific HA2 fragment after 30 minutes. In previous studies of HA at neutral pH (Ruigrok et al., 1986), the BHA samples were incubated at 70 °C and digested with low concentrations of trypsin; thus, the BHA2 fragment detected at low concentrations of trypsin could not have been distinguished from HA2 in the heat-induced, fusogenic conformation.

### **Preparation of Fluorescent-labelled Vesicles**

Phospholipids were combusted, acid hydrolyzed and quantitated using an inorganic phosphate assay (Fiske and SuBarrow, Sigma, St. Louis, MO). Ganglioside GD1a was prepared in a 1:1 mixture of methanol:chloroform as a 1mg/ml solution (by weight) and allowed to dissolve at least 4 hours at 4 °C. The fluorophores NBD-PE (fluorescent donor, "NBD") and lisammine-rhodamine-PE (fluorescent acceptor, "Rho") were incorporated into DOPC/GD1a lipid bilayers at molar ratios of 95.3 : 4.5 : 0.6 : 0.6 (DOPC:GD1a:NBD:Rho), and dried, first under a slow stream of nitrogen, and then by vacuum desiccation for 2hr with a liquid nitrogen trap. Large, unilamellar vesicles (LUVs) were prepared by freeze-thaw hydration (Mayer et al., 1986) of the lipids in "fusion buffer" (50 mM HEPES pH 7.4, 150 mM sodium chloride, 0.1 mM EDTA) followed by extrusion, 10 times, through two 0.1mm polycarbonate filters (Poretics, Livermore, CA). Lipid bilayer integrity was confirmed for DOPC/GD1a LUVs prepared in fusion buffer and for LUVs prepared in fusion buffer with 3.75 M urea, by entrapping self-quenching concentrations of carboxyfluorescein. Dilution of LUVs, purified by PD-10 (Pharmacia-Biotech, Uppsala, Sweden) G-25 gel-filtration did not lead to dequenching of 5(6)-carboxyfluorescein. However, addition of ~5µl, 25% Triton-X 100 to dilute LUVs lead to rapid release of carboxyfluorescein and a color change to fluorescent green, indicating a breach of lipid bilayer integrity. Viral lipids were extracted (Bligh and Dyer, 1959) and quantitated using the phosphate assay.

### **Membrane fusion Assay**

Membrane fusion is monitored using the resonance energy transfer (RET) assay of Struck (Struck et al., 1981). Samples excited at 465 nm are monitored at 535 nm for NBD emission as a function of time, using an ISS-PC Greg fluorimeter (with 16 nm excitation and emission slits) equipped with a thermostatted cuvette holder. NBD emission is minimal for the intact LUVs due to efficient energy transfer from NBD to Rho (which emits at 590 nm). Upon fusion of virus with LUVs, NBD energy transfer is uncoupled from Rho by dilution with the viral lipids, resulting in an increase in emission at 535 nm. The lipid mixing endpoint is defined as the fluorescence emission observed at 535 nm upon addition of Triton-X 100 to 1%. Percent fusion is calculated as the observed signal minus the initial baseline at pH 7.4, divided by the lipid-mixing endpoint minus the initial baseline.

### ***Sample Preparation for Acid-Induced Fusion***

Virus is mixed with 0.6% NBD and Rho-labelled LUVs, at a 1:1 phospholipid ratio and incubated at 4 °C for 30 minutes at pH 7.4 (Alford et al., 1994) to allow for maximal binding of virus to the GD1a on the LUVs. The effect of prebinding on the extent of fusion is temperature-dependent: no difference is seen at 20 °C, but at 37 °C, fusion is less

efficient in the absence of prebinding (not shown). These “prebound” virus and LUVs are added to a fluorescence cuvette containing fusion buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 0.1 mM EDTA) to a final concentration of 50  $\mu$ M phospholipids and stirred at the final temperature indicated.

For low pH experiments, the pH of the sample is lowered to the indicated pH upon addition of a small volume (10-30  $\mu$ l) of 0.1 M citric acid. The initial baseline at pH 7.4 is defined as 0% fusion, and the endpoint, 100% lipid mixing, is defined by the signal seen upon addition of Triton X-100 to 1%.

For acid-inactivation experiments, virus is incubated with citrate at the indicated pH (experiments have been done for pH 4.5, 5.0 and 5.8) for 30 minutes at 37 °C, reneutralized and cooled to room temperature, mixed 1:1 with LUVs and prebound, as described above. Inactivated virus and LUVs are added with stirring to heated buffer in the fluorescence cuvette, and the signal at 535 nm (constant) is recorded. Citrate is added to lower the pH (experiments have been done for pH 4.6, 5.3, 5.6 and 5.8), and the signal (constant and unchanged, sometimes slightly lower) is recorded.

#### *Sample Preparation for Temperature-Induced Fusion*

For temperature experiments, prebound virus is added at pH 7.4 directly to heated buffer; thus there is no initial baseline. For these experiments, the initial baseline is estimated from the inactivated control (see below). As a check for complete fusion at the end of the temperature experiment, citrate is added to pH 5.8, and no change in signal is detected. The endpoint, 100% lipid mixing, is defined by the signal seen upon addition of Triton X-100 to 1%.

For heat-induced inactivation experiments, virus is incubated at the final temperature (experiments have been done for 58 °C and 60 °C inactivation) for 30 minutes, cooled to room temperature, mixed 1:1 with LUVs and prebound, as described above. Virus and LUVs are added with stirring to heated buffer in the fluorescence cuvette, and the fluorescence signal (constant) is recorded. The constant signal measured in the inactivation experiments is estimated as the initial baseline for the heat-induced fusion experiments. Inactivation at 58 °C was used to determine the initial baseline for the 58 °C fusion experiment, and inactivation at 60 °C was used to determine the initial baseline for the 60 °C experiment. Shown are the results from the 58 °C inactivation experiment (Fig. 4b)

### *Sample Preparation for Urea-Induced Fusion*

In the urea experiments prebound virus and LUVs prepared in urea at the indicated concentration are added directly to urea plus buffer at pH 7.4, with stirring at 37 °C. Because fusion is initiated immediately upon addition of virus and LUVs to the warm urea, there is no initial baseline. For these experiments, the initial baseline is estimated from the inactivated control (see below). As a check for fusion-competence of urea LUVs, fusion was initiated by lowering the pH at room temperature (as above), which did induce fusion between virus and LUVs, both in aqueous buffer and in urea. Prebinding virus and LUVs yields more efficient fusion, though fusion does proceed in the absence of prebinding (not shown).

For inactivation experiments, virus is incubated at 3.75 M urea for ~1 hour at 37 °C, cooled to room temperature, mixed with LUVs and prebound, as described above. Virus and LUVs are added to stirring urea plus buffer at 37 °C, in the fluorescence cuvette and the signal (constant) is recorded. The average of the constant signal observed for the inactivation experiment is estimated as the initial baseline for the fusion experiments. Shown are the results from the inactivation experiment measured at 3.0 M urea (Fig. 4c).

## References

- Baker, D., and Agard, D. A. (1994). Kinetics versus thermodynamics in protein folding. *Biochemistry* 33, 7565-9.
- Bligh, E. G., and Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37, 911-917.
- Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994). Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* 371, 37-43.
- Carr, C. M., and Kim, P. S. (1993). A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* 73, 823-32.
- Doering, D. (1992). Functional and Structural studies of a small F-actin binding domain. (USA: Massachusetts Institute of Technology).
- Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* 6, 1948-1954.
- Grosjean, H., and Fiers, W. (1982). Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene* 18, 199-209.
- Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* 36, 575-588.
- Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U S A* 82, 488-92.
- Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. E. Harding, Rowe, A. J., and Horton, J. C., eds. (Cambridge: The Royal Society of Chemistry).
- Mayer, L. D., Hope, M. J., and Cullis, P. R. (1986). Vesicles of variable sizes produced by a rapid extrusion procedure. *Biochim. Biophys. Acta.* 858, 161-168.
- Ramalho-Santos, J., Nir, S., Duzgunes, N., de Carvalho, A. P., and de Lima M da, C. (1993). A common mechanism for influenza virus fusion activity and inactivation. *Biochemistry* 32, 2771-9.
- Ruigrok, R. W., Aitken, A., Calder, L. J., Martin, S. R., Skehel, J. J., Wharton, S. A., Weis, W., and Wiley, D. C. (1988). Studies on the structure of the influenza virus haemagglutinin at the pH of membrane fusion. *J. Gen. Virol.* 69, 2785-95.
- Ruigrok, R. W., Martin, S. R., Wharton, S. A., Skehel, J. J., Bayley, P. M., and Wiley, D. C. (1986). Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. *Virology* 155, 484-97.

- Ruigrok, R. W., Wrigley, N. G., Calder, L. J., Cusack, S., Wharton, S. A., Brown, E. B., and Skehel, J. J. (1986). Electron microscopy of the low pH structure of influenza virus haemagglutinin. *EMBO J.* *5*, 41-9.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). (Cold Spring Harbor, New York: Cold Spring Harbor Press).
- Schoch, C., Blumenthal, R., and Clague, M. J. (1992). A long-lived state for influenza virus-erythrocyte complexes committed to fusion at neutral pH. *FEBS Lett.* *311*, 221-5.
- Skehel, J. J., Bayley, P. M., Brown, E. B., Martin, S. R., Waterfield, M. D., White, J. M., Wilson, I. A., and Wiley, D. C. (1982). Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc. Natl. Acad. Sci. U S A* *79*, 968-72.
- Struck, D. K., Hoekstra, D., and Pagano, R. E. (1981). Use of resonance energy transfer to monitor membrane fusion. *Biochemistry* *20*, 4093-9.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* *185*, 60-89.
- Weber, T., Paesold, G., Galli, C., Mischler, R., Semenza, G., and Brunner, J. (1994). Evidence for H(+)-induced insertion of influenza hemagglutinin HA2 N-terminal segment into viral membrane. *J. Biol. Chem.* *269*, 18353-8.
- Weissman, J. S. and Kim, P. S. (1991). Re-examination of the Folding of BPTI: Predominance of Native Intermediates. *Science* *253*, 1386-93.
- Wharton, S. A., Calder, L. J., Ruigrok, R. H. R., Skehel, J. J., Steinhauer, D. A., and Wiley, D. C. (1995). Electron microscopy of antibody complexes of influenza virus haemagglutinin in the fusion pH conformation. *EMBO J.* *14*, 240-246.
- Wharton, S. A., Skehel, J. J., and Wiley, D. C. (1986). Studies of influenza haemagglutinin-mediated membrane fusion. *Virology* *149*, 27-35.
- White, J., Kartenbeck, J., and Helenius, A. (1982). Membrane fusion activity of influenza virus. *EMBO J.* *1*, 217-22.
- Wiley, D. C., and Skehel, J. J. (1987). The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu. Rev. Biochem.* *56*, 365-94.
- Wiley, D. C., and Skehel, J. J. (1990). *Viral Membranes*. (New York: Raven Press, Ltd.).
- Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* *289*, 366-73.

## Figure 1

A protease assay for the conformation of HA. Virus samples are incubated under various conditions (pH 7.3, 37 °C; pH 5.0, 37°C; pH 7.3, 65°C; or pH 7.3, 37°C, 4.5M urea), and the conformation of HA is assayed by its sensitivity to proteinase K digestion at neutral pH, room temperature, and dilute aqueous buffer. Viral proteins and proteolytic fragments are separated by SDS-PAGE and visualized by coomassie-blue staining (a, c) or immunoblot analysis (b, d).

(a) HA and HA fragments, nucleoprotein (NP) and matrix protein (M1) are detected by coomassie-blue staining of viral proteins separated under non-reducing conditions.

(b) Immunoblot analysis with anti-HA2 antibodies reveals HA and HA fragments separated under non-reducing conditions.

(c) HA1, NP, M1 and HA2 fragments are separated under reducing conditions and stained with coomassie blue.

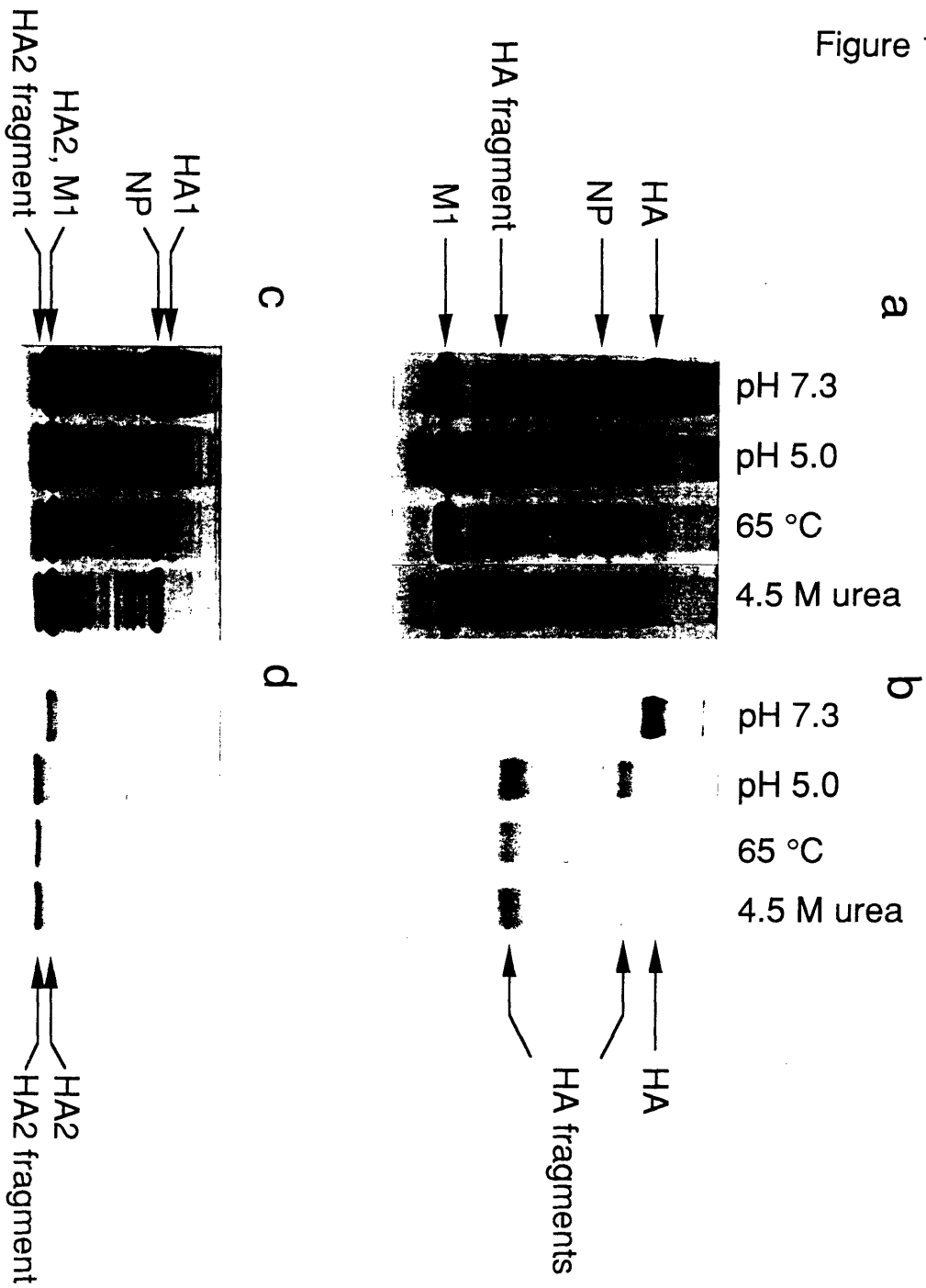
(d) Immunoblot analysis using anti-HA2 antibodies reveals HA2 and HA2 fragments, separated under reducing conditions. Identity of the HA2 fragment, was confirmed for all three samples (pH 5.0, 65 °C and 4.5 M urea) by 5 cycles of N-terminal amino acid sequence analysis. The N-terminus of the HA2 fragments begins at residue 28, with the sequence, NSEGT, for all three conditions.

HA = one HA1 subunit disulfide-bonded to one HA2 subunit. HA fragments = disulfide-bonded HA1 and HA2 proteolytic fragments. HA2 = one monomer HA2 subunit. HA2 fragment = proteolytic fragment of HA2. M1 = matrix protein. NP= nucleoprotein.

N = asparagine. S = serine. E = glutamic acid.  
G = glycine. T = threonine.

# HA Conformational Change

Figure 1

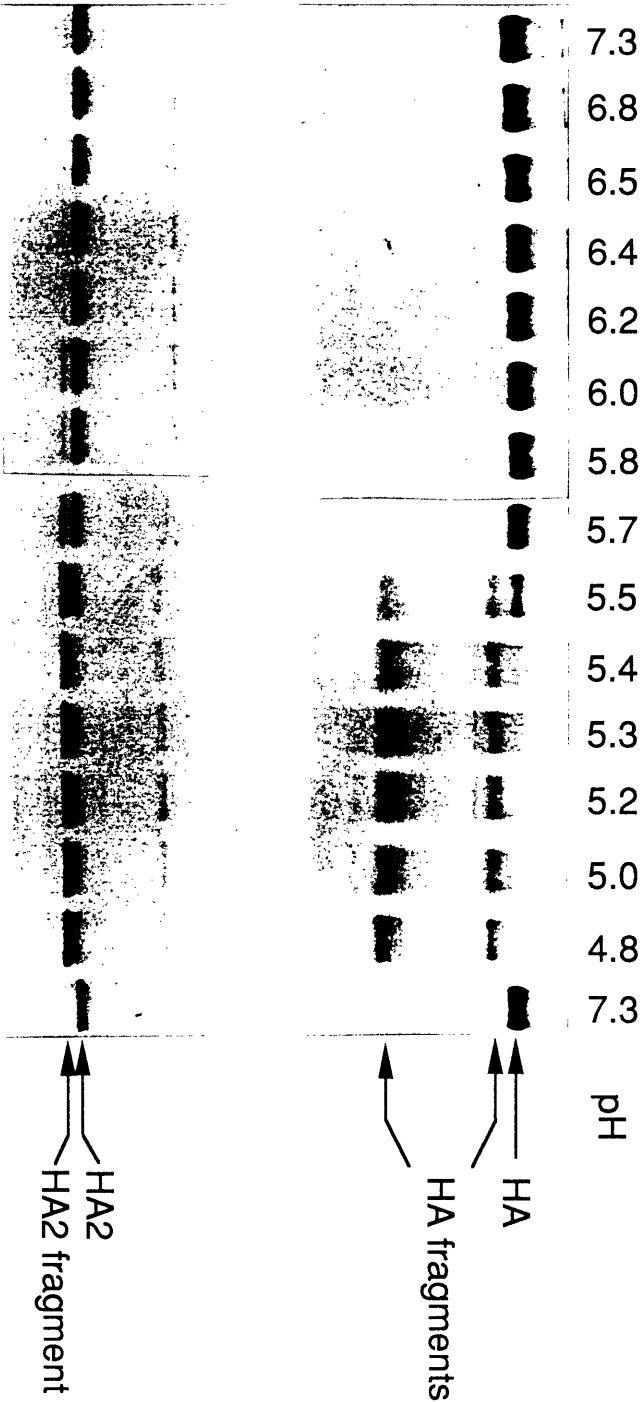


## Figure 2

Immunoblot-detection of HA and HA fragments generated by the protease assay described in Fig. 1. Virus samples were incubated for 15 minutes under conditions of variable pH, temperature, or urea concentration. Samples were then cooled, reneutralized, diluted and assayed for conformation by proteolysis with proteinase K. Incubation conditions are varied as a function of (a) pH at 37°C; (b) temperature, at pH 7.3 and (c) urea concentration at 37°C, pH 7.3.



Figure 2a



Acid - Induced HA Conformational Change

Figure 2b

Heat - Induced HA Conformational Change, pH 7.3

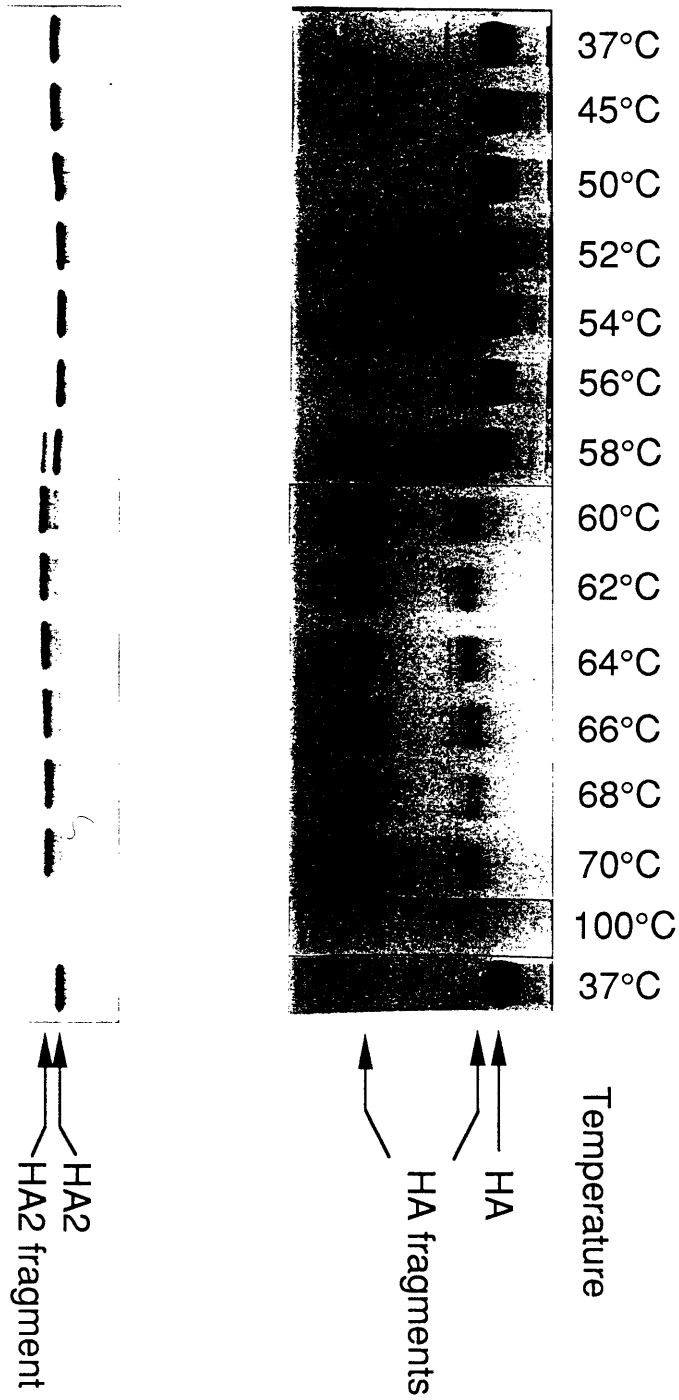
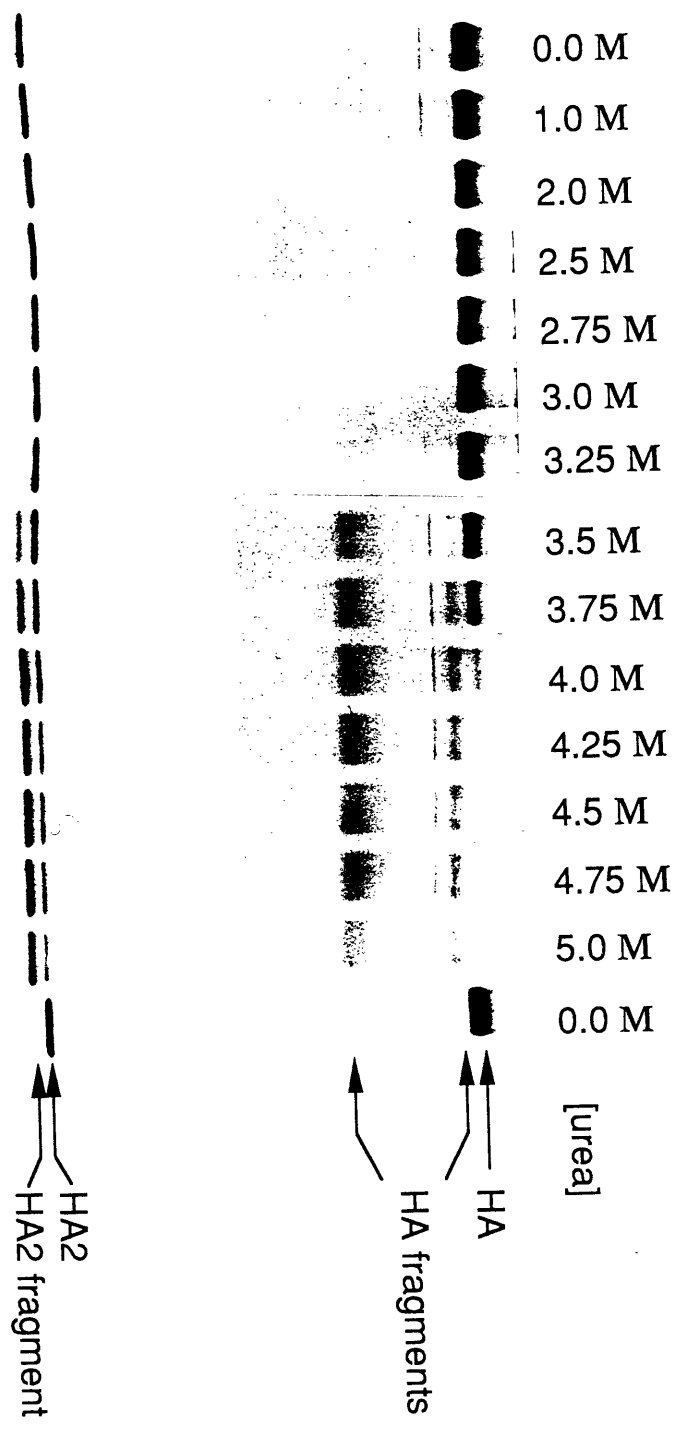


Figure 2c

Urea - Induced HA Conformational Change



### Figure 3

Stability of the rHA2-X117 coiled coil peptide. Samples of 10 $\mu$ M peptide were incubated at pH 7.3 for 15 minutes at the indicated temperature or urea concentration. Denaturation of the helical structure is indicated by an increase in molar ellipticity ( $[\theta]_{222}$ ) measured by circular dichroism spectroscopy at 222nm. (a) Helicity of rHA2-X117 as a function of temperature, at pH 7.3. (b) Helicity of rHA2-X117 as a function of urea concentration at pH 7.3 and 37 °C.

Figure 3a

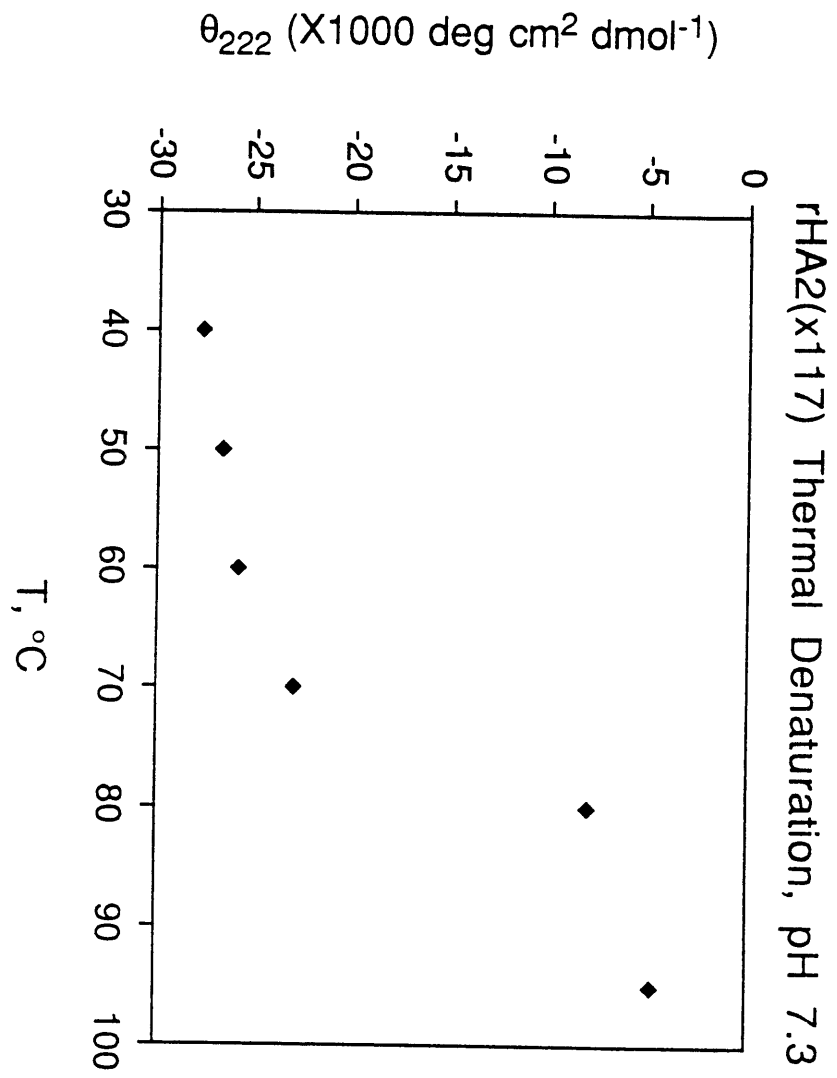
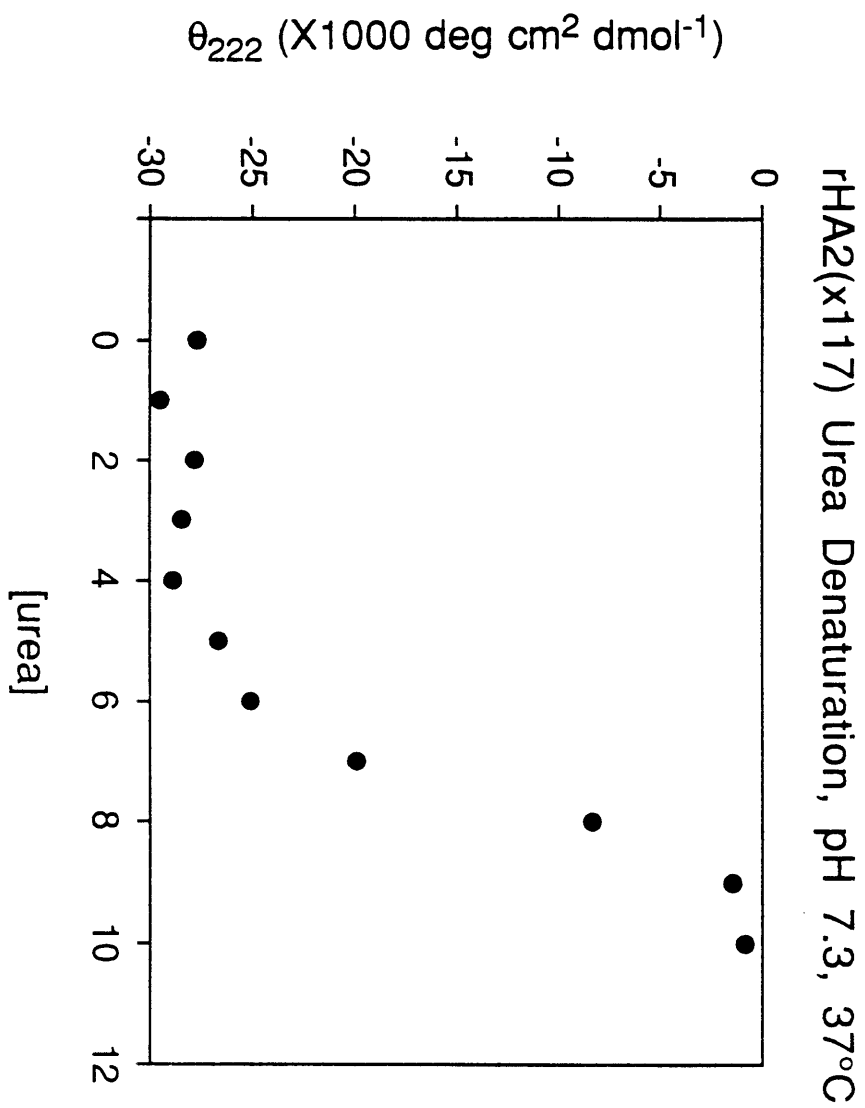


Figure 3b



#### Figure 4

Membrane fusion between influenza virus and synthetic vesicles. Membrane fusion is induced by (a) acid at 37 °C, initiating membrane fusion at pH 6.0 (filled circles), with optimal fusion at pH 5.8 (filled squares); (b) heat at pH 7.3, initiating membrane fusion at 58 °C (filled circles), with optimal fusion at 60 °C (filled squares); or (c) urea, at 37 °C, pH 7.3, initiating membrane fusion at 3.0 M urea (filled circles), with optimal fusion at 3.5 M urea (filled squares). Conditions that induce membrane fusion in the presence of vesicles also result in inactivation of the virus, when incubated under the same conditions in the absence of vesicles (open circles) (a) acid-inactivated; (b) heat-inactivated, and (c) urea-inactivated virus assayed for membrane fusion). See methods for a detailed description of the membrane fusion assay.

Figure 4a

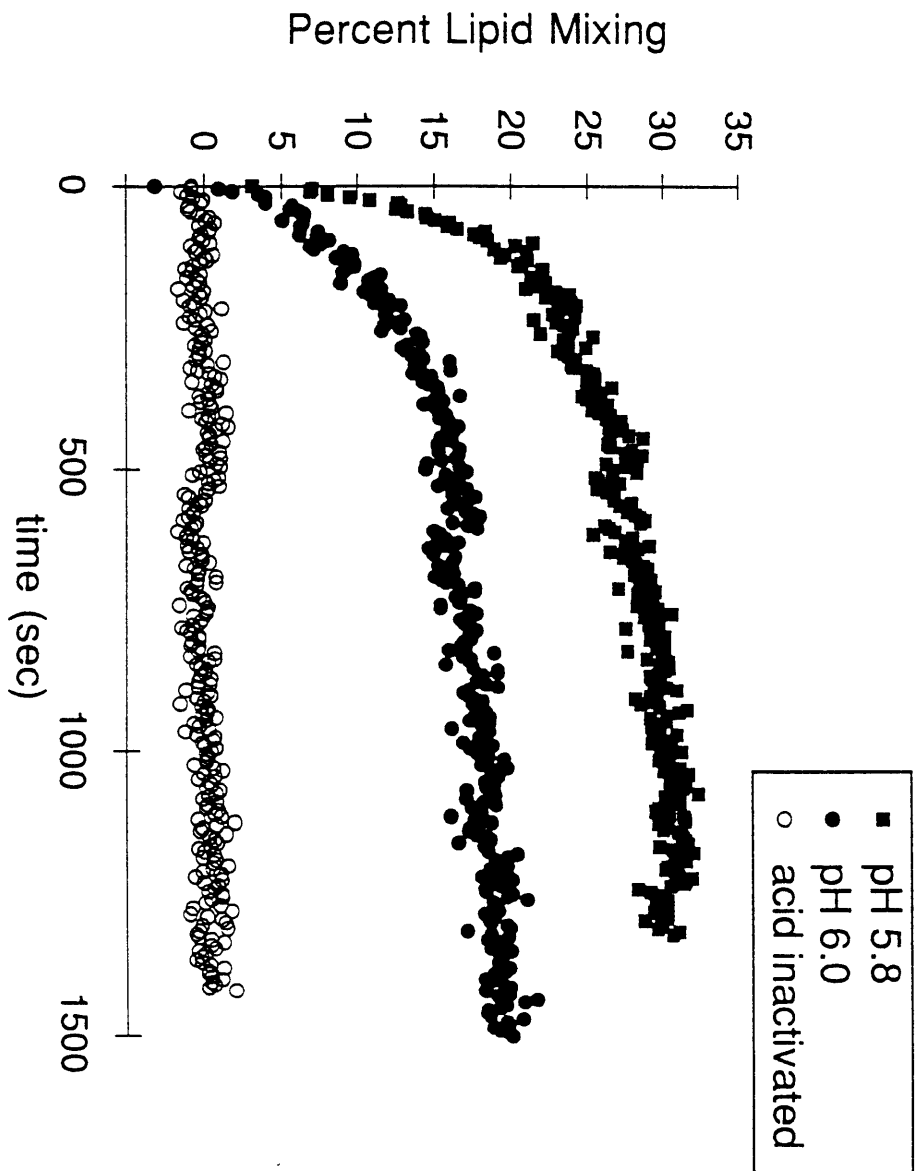




Figure 4b

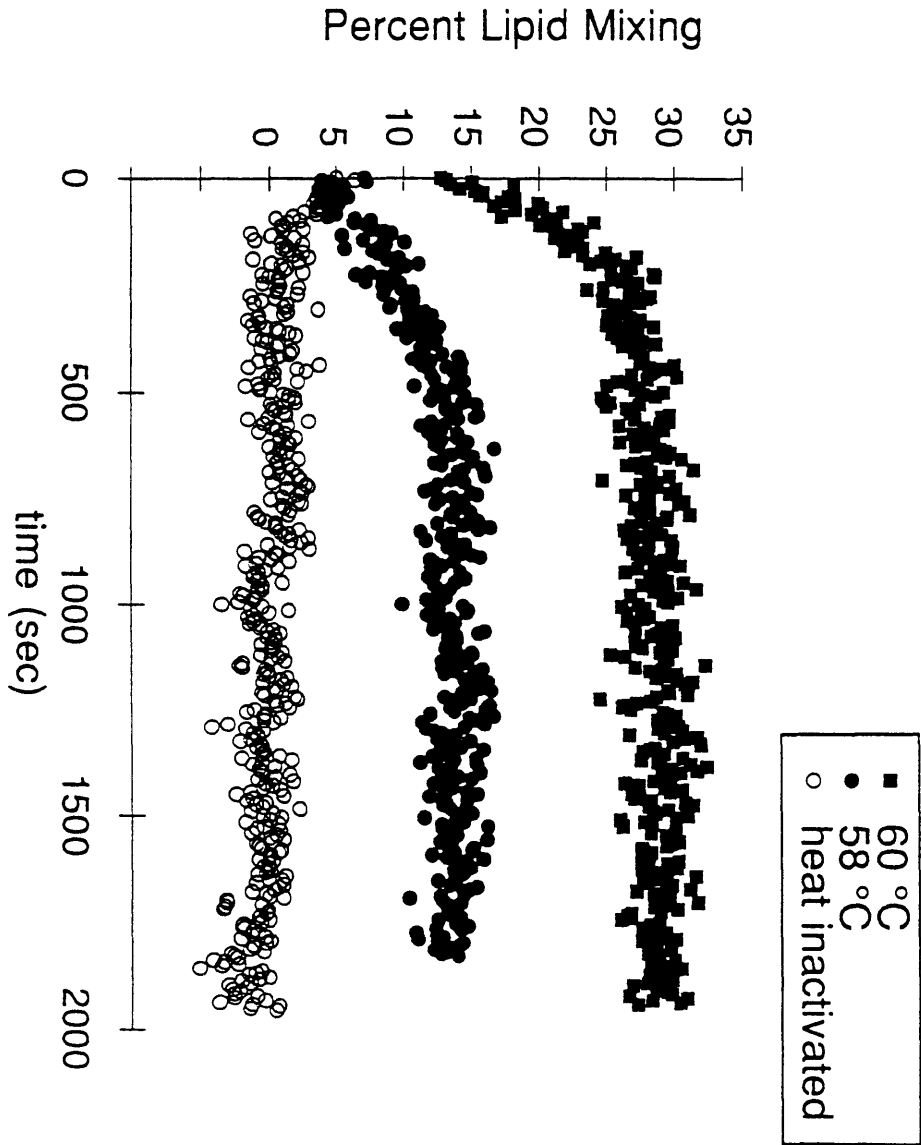
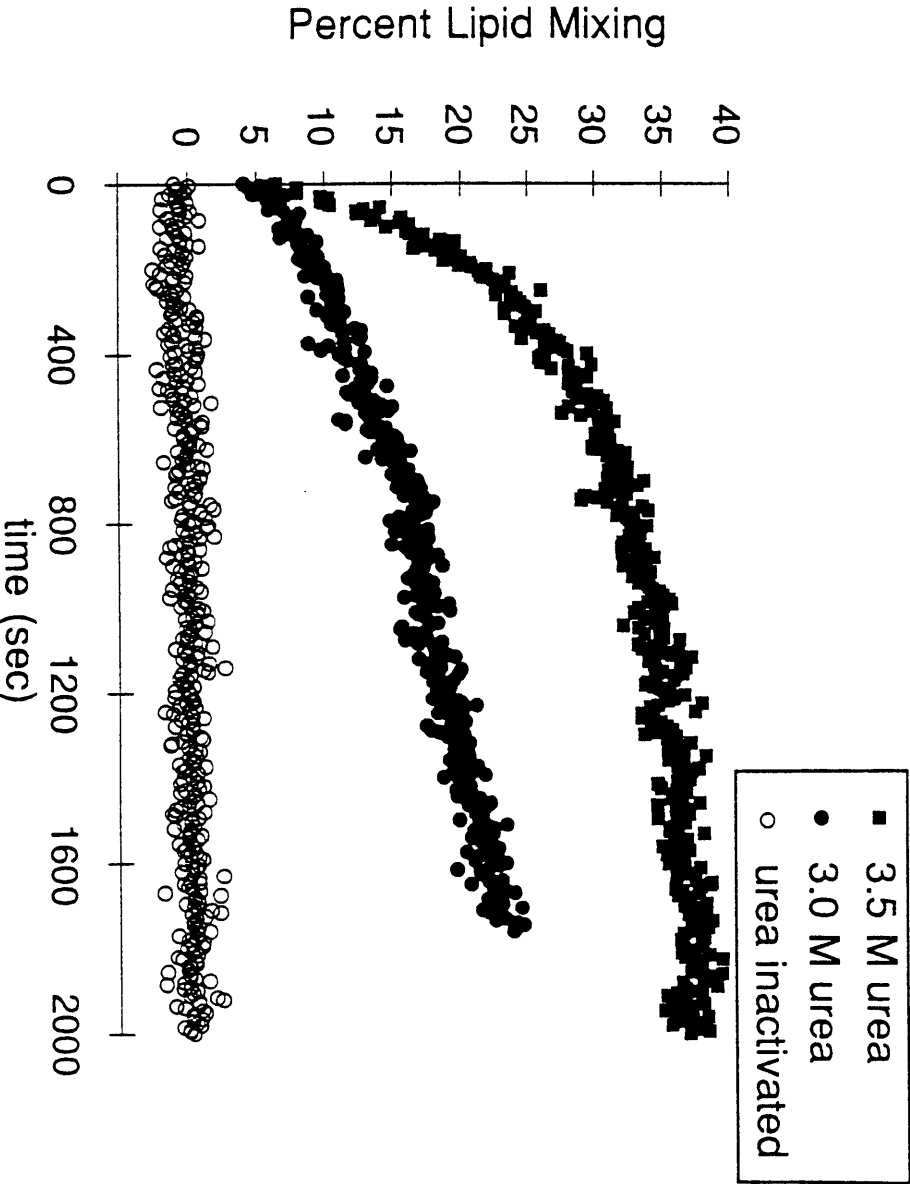


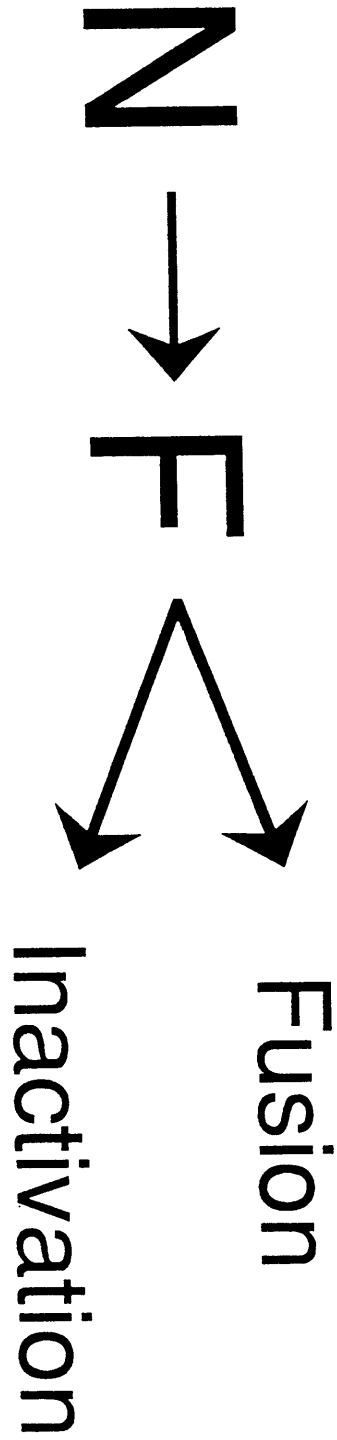
Figure 4c



## **Figure 5**

A model for the mechanism of the HA conformational change. Depicted is the irreversible transition from the metastable native state (N) to the stable fusogenic state (F). The fusogenic state induces membrane fusion in the presence of target membranes and causes viral inactivation in the absence of target membranes.

Figure 5



## Chapter 4

## Flu Virus Invasion: Halfway There

Fertilization of an egg cell by a sperm is perhaps the most dramatic consequence of membrane fusion. Enveloped viruses, such as HIV, also utilize membrane fusion to infect animal cells. The fusion of two distinct lipid bilayers is an exceedingly difficult process in the absence of specialized proteins, such as the sperm protein, PH-30 $\alpha$ / $\beta$  (1), or the HIV envelope protein, gp120/gp41 (2).

The best characterized membrane-fusion protein is the hemagglutinin (HA) protein found on the surface of influenza (“flu”) virus. In spite of decades of research, the mechanism by which HA induces fusion remains elusive. Recent developments, however, including one on page 274 of this issue of Science (3), suggest that we are getting closer to an understanding of this protein-mediated membrane fusion process.

Flu onset begins with the binding of influenza virus to nasal epithelial cells. In a function that is distinct from membrane fusion, HA mediates viral attachment to sugar groups (sialic acid) found on the surface of the host cell. Binding does not lead directly to membrane fusion, since HA in the “native” state is not fusion-active. Instead, the cell internalizes the bound virus and surrounds it in an endosome. Only in the mildly acidic conditions of the mature endosome does the HA protein switch from an inactive to a fusion-active (“fusogenic”) state. In this fusogenic conformation, HA promotes fusion of the viral and cellular membranes, leading to release of the nucleocapsid into the cytoplasm, thereby initiating replication and proliferation of the virus.

But how does HA induce membrane fusion? Thirteen years ago, the x-ray crystal structure of HA in the native state was reported by Don Wiley and coworkers (4). HA is a trimeric protein with three identical subunits that span the viral membrane. Each subunit is synthesized as a precursor that subsequently is proteolytically cleaved (5) in a process required for fusion activity (6). The newly created amino terminus contains a hydrophobic sequence of ~25 amino-acid residues that are known to be essential for membrane fusion

(7); this region is therefore called the “fusion peptide.” The baffling aspect of the structure was that the fusion-peptide region is buried deep inside the native HA protein.

It seemed that the fusion-peptide regions must be exposed in the fusogenic state. Indeed, it was demonstrated that when HA is subjected to acid, the fusion-peptide region becomes accessible to antibodies, proteases, and chemical reagents (8). In addition, experiments using lipid-soluble chemical modification reagents showed that the fusion-peptide region of HA inserts into the target membrane before fusion (9, 10). Thus, a key intermediate in the fusion process consists of HA in two different membranes at the same time: the transmembrane region of HA spans the viral membrane and the fusion-peptide region inserts into the target membrane.

Nonetheless, the structural constraints associated with formation of this crucial fusion intermediate appeared formidable. A conformational change that releases the fusion-peptide region from the native-HA core would still leave these exposed regions more than 100 Å away from the target membrane. How do the fusion-peptide regions reach the cellular membrane? Imaginative models were proposed to answer this question. For example, one model (11) proposed that HA tilts 90°, lying against the viral membrane to facilitate interaction between the fusion-peptide region and the target membrane.

Recently we proposed a different model (12) which evolved out of our interest in the coiled coil, a structure of interwound  $\alpha$ -helices. A sequence of 28 amino-acid residues with a marked propensity for adopting the coiled-coil structure was identified within HA. In contrast to the prediction, however, this sequence does not form a coiled coil in the native conformation of HA, but forms an extended “loop” structure. The significance of this loop region is that it joins the fusion-peptide region to the three-stranded, coiled-coil core of the native state (see panel 1). A synthetic peptide corresponding to this loop region was found to form a trimeric coiled coil at the pH of membrane fusion. The fact that this loop region could form two completely different structures suggested a “spring-loaded” mechanism for the conformational change of HA (see panel 2): at acidic pH the loop region forms a coiled coil, projecting the fusion-peptide regions to the top of the molecule where they can then interact easily with the target membrane.

Last month, Wiley and coworkers reported the crystal structure of a large fragment of the acid-induced state of HA (13). The structure confirms the salient feature of the “spring-loaded” model: the loop region of the native state is transformed into a three-stranded coiled coil in the acid-induced state (Fig. 1). In addition, the new crystal structure reveals a second major change at the opposite end of HA, near the viral membrane: the long  $\alpha$  helix of the coiled-coil core stops near the bottom and reverses directions, bending upward to pack against itself (see panel 2). The transition is remarkable: fewer than half of the residues in the new crystal structure are in the same conformation as in the native state.

Although the mechanism by which the fusion-peptide region reaches the target membrane may now be understood, this structure presents a new problem. HA in this conformation (see panel 2) holds the two membranes 100 Å apart! How then do the membranes come together for fusion? Using electron paramagnetic resonance (EPR) methods, Shin and coworkers provide a clue to the next step: membrane apposition (3). Acid-induced interactions between lipid membranes and a peptide that corresponds to the loop region of native HA were detected by monitoring changes in the spectral properties of a spin-labelled probe attached at different positions in the peptide. In their model, the top of the acid-induced coiled coil splays apart and “melts” into the lipid bilayer of the cellular membrane, thereby bringing the viral and cellular membranes closer together (see panel 4). Although verification using intact HA protein is required, this model provides a tantalizing solution to the problem of how the two membranes are juxtaposed.

Even so, apposition of the two membranes is not the only requirement for fusion. By replacing the viral-membrane spanning region of HA with a glycosyl-phosphatidylinositol (GPI) lipid anchor (14), White and coworkers identified a new, putative fusion intermediate. Upon acidification, lipid mixing occurs between cells expressing the altered HA and target red-blood cells, as monitored with a lipid-soluble fluorescent label. Remarkably, however, the contents of these cells do not mix, as determined by following a water-soluble fluorescent dye.

While questions remain, any solution to the membrane-fusion problem must now contend with an acid-induced structure of HA that is remarkably different from the native state, biochemical evidence that more than just the fusion peptide can interact with the target membrane, and the observation that lipid mixing induced by GPI-anchored HA is not



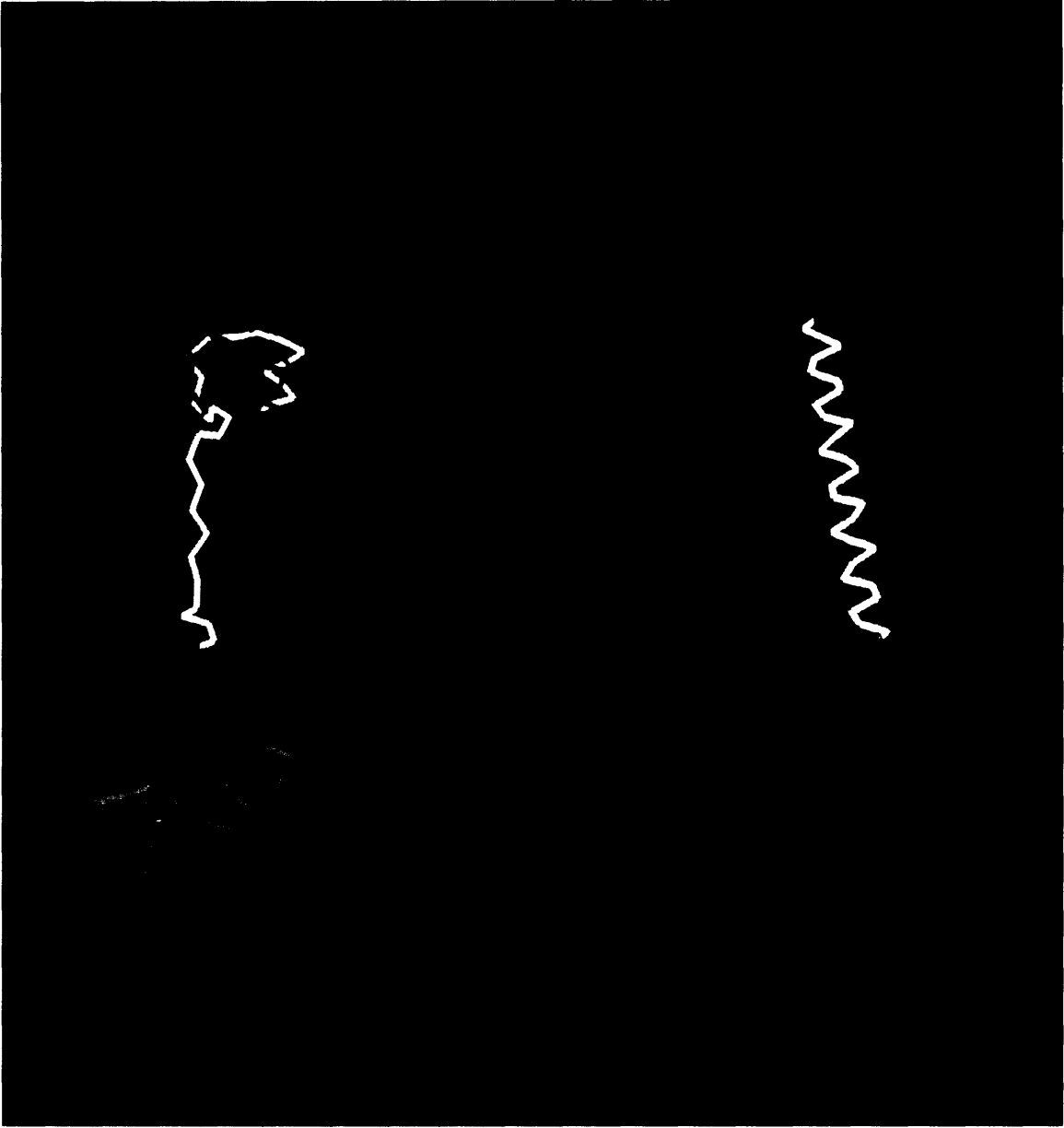
sufficient for membrane fusion. A major challenge ahead is to reveal the architecture of the “fusion pore” that forms as the viral and cellular membranes fuse (15). The puzzle is not complete, but we may be halfway to understanding how influenza virus invades our cells.

## References and Notes

1. C. P. Blobel, T. G. Wolfsberg, C. W. Turck, D. G. Myles, P. Primakoff, J. M. White, *Nature* **356**, 248-52 (1992).
2. For a review see: Y. N. Vaishnav and F. Wong-Staal, *Ann. Rev. Biochem.* **60**, 577-630 (1991).
3. Y. G. Yu, D. S. King, Y. -K. Shin, *Science* vol, pages (1994).
4. I. A. Wilson, J. J. Skehel, D. C. Wiley, *Nature* **289**, 366-73 (1981). Each monomer unit of the trimer is composed of two, disulfide-bonded subunits, HA1 and HA2, which are cleaved from a single precursor, HA0. HA1 is the globular, sialic-acid binding domain at the top of the molecule (green spheres in the figure; the most distal part with respect to the center of the virus). The HA2 subunit spans the viral membrane once and forms a three-stranded,  $\alpha$ -helical coiled-coil structure that serves as the central core of the HA trimer.
5. S. G. Lazarowitz, R. W. Compans, P. W. Choppin, *Virology* **46**, 830-43 (1971); J. J. Skehel and M. D. Waterfield, *Proc Natl Acad Sci USA* **72**, 93-7 (1975).
6. H. D. Klenk, R. Rott, M. Orlich, J. Blödorn, *Virology* **68**, 426-39 (1975); S. G. Lazarowitz and P. W. Choppin, *Virology* **68**, 440-54 (1975).
7. R. S. Daniels, J. C. Downie, A. J. Hay, M. Knossow, J. J. Skehel, M. L. Wang, D. C. Wiley, *Cell* **40**, 431-9 (1985); M. -J. Gething, R. W. Doms, D. York, J. White, *J Cell Biol* **102**, 11-23 (1986).
8. For reviews see: D. C. Wiley and J. J. Skehel, *Annu Rev Biochem* **56**, 365-94 (1987); J. M. White, *Science* **258**, 917-24 (1992).
9. T. Stegmann, J. M. Delfino, F. M. Richards, A. Helenius, *J Biol Chem* **266**, 18404-10 (1991).
10. M. Tsurudome, R. Glück, R. Graf, R. Falchetto, U. Schaller, J. Brunner, *J Biol Chem* **267**, 20225-32 (1992). Similar methods have also shown that the fusion peptide can insert into the viral membrane (T. Weber, G. Paesold, C. Galli, R. Mischler, G. Semenza, J. Brunner, *J Biol Chem* **269**, 18353-58 [1994]). This is likely to be a cause of viral inactivation (see discussion, ref. 13).
11. T. Stegmann, J. M. White, A. Helenius, *EMBO J* **9**, 4231-41 (1990).
12. C. M. Carr and P. S. Kim, *Cell* **73**, 823-32 (1993).
13. P. A. Bullough, F. M. Hughson, J. J. Skehel, D. C. Wiley, *Nature* **371**, 37-43 (1994). In addition to being cleaved from the viral membrane, it was necessary to subject the low-pH form of HA to further proteolysis to permit crystallization, because the exposed fusion-peptide region leads to aggregation and precipitation of the protein. Aggregation of fusion peptides within or between hemagglutinin proteins is likely to be another cause of inactivation of hemagglutinin-mediated membrane fusion (ref. 10; P. R. Junankar and R. J. Cherry, *Biochim Biophys Acta* **854**, 198-206 [1986]). Although proteolysis removes the sialic-acid binding domains of HA, it is unlikely that these domains contribute significantly to the structure of the acid-induced conformation (see discussion in Bullough et al., 1994).
14. G. W. Kemble, T. Danieli, J. M. White, *Cell* **76**, 383-91 (1994).
15. The appearance of pores opening and closing during membrane fusion has been detected using electrophysiologic methods adapted from neurobiology (A. E. Spruce, A. Iwata, J. M. White, W. Almers, *Nature* **342**, 555-8 [1989]).

## Figure 1

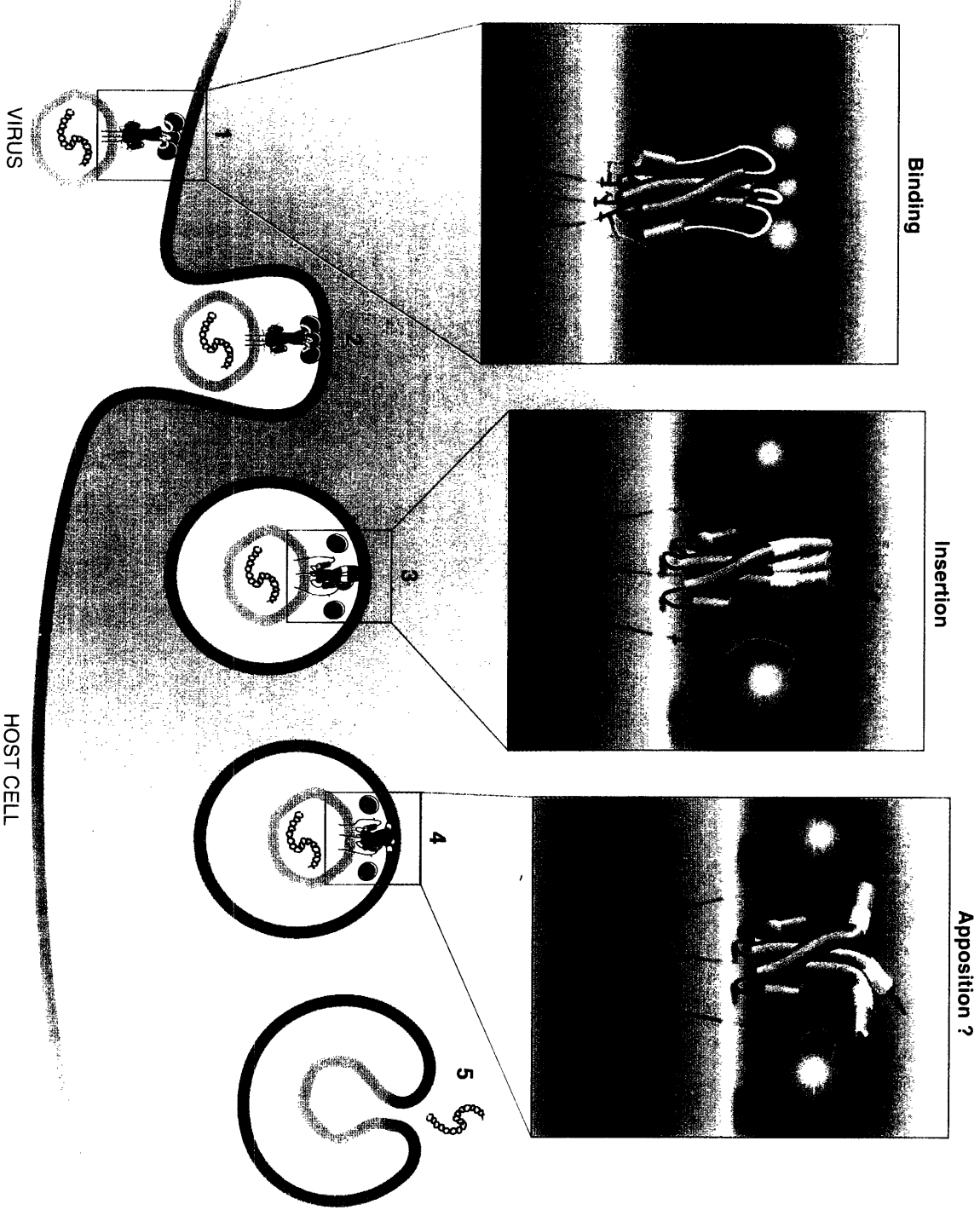
The structure of HA in the native (left) and the acid-induced (right) conformation. One monomer of the trimer is highlighted in multi-color, while the other two are purple. Two major conformational changes are observed: the loop region (yellow, left) of the native conformation is transformed into a coiled coil in the acid-induced conformation (yellow, right). In addition, the end of the long helix (pink, left) of the native conformation bends back, to pack against itself (pink, right) in the acid-induced conformation. HA1 (blue) and the fusion-peptide region (light blue) are proteolytically removed to facilitate crystallographic analysis of the acid-induced conformation.



## Figure 2

Flu virus invasion. **Step 1.** Hemagglutinin in the native state mediates binding of flu virus to the host cell. The domains at the top of HA (green balls) bind to sialic acid residues on the cell surface. **Step 2.** The virus is internalized in an endosome. **Step 3.** In the acidic conditions of the mature endosome, HA switches to a fusion-active state. The loop region forms a coiled coil (yellow), which includes the  $\alpha$  helix (orange). As a result, the fusion-peptide regions (black) insert into the cellular membrane at the top of the molecule. At the opposite end, each long  $\alpha$  helix (pink) bends up to pack against the core of the trimer. Hemagglutinin is now attached to both the viral and the cellular membranes. **Step 4.** In a model for membrane apposition, the long coiled coil separates and the top of the coiled coil (orange and yellow) “melts” into the cellular membrane, bringing the two membranes closer together. **Step 5.** Ultimately, in a process that is poorly understood, the bilayers fuse, releasing the nucleocapsid into the cell.

ILLUSTRATION K SUTLIFF



## Appendix

## Structural and Functional Studies of the Coiled Coil

First described as the highly stable structure of fibrous proteins such as keratin, myosin and tropomyosin, the coiled coil has gained renewed interest as a modular oligomerization motif used by a variety of proteins (for reviews see Oas and Endow, 1994; Cohen and Parry, 1994) and for its regulatory role in the dynamics of partner specificity between a large number of proteins from the bZIP and bHLH-ZIP families of transcription factors (for reviews, see Hu and Sauer, 1992; Baxevanis and Vinson, 1993).

The structure of the coiled coil can be described as a rope, typically consisting of 2 or 3  $\alpha$ -helices wrapped together either in a parallel or antiparallel fashion with a left-handed, superhelical twist. The details of the coiled-coil structure were deduced from an x-ray diffraction pattern of keratin (Crick, 1953; Pauling and Corey, 1953), and proven correct by the first atomic structure of a coiled coil: the crystal structure of a peptide corresponding to the 33 amino-acid dimerization domain from the yeast transcription factor, GCN4 (GCN4-p1; O'Shea et al., 1991).

The coiled coil can be identified as a pattern in the amino-acid sequence of a protein. As a consequence of the regular, repeating nature of the structure, coiled-coil sequences reveal a preference for hydrophobic amino-acid residues at the first position, (denoted **a**) and the fourth position (denoted **d**) of a repeated, seven amino-acid sequence pattern. Hydrophilic residues predominate at the fifth (**e**) and seventh (**g**) positions. The hydrophobic residues at positions **a** and **d** of this heptad repeat form the interface between helices. This feature is known as the "4-3 hydrophobic repeat" (Hodges et al., 1972; McLachlan and Stewart, 1975).

A subset of coiled coils is the leucine-zipper structural motif. The leucine-zipper (Landshultz et al., 1988) is a dimerization motif found in two families of transcription factors known as bZIP (Landschulz et al., 1988) or bHLH-ZIP (Blackwood and Eisenman, 1991). Leucine-zipper sequences are characterized by invariant leucines every seventh residue, or the **d** position in the heptad repeat of the coiled coil.



Superimposed on the coiled-coil heptad repeat pattern are the amino-acid determinants that account for variations in stability and specificity between different leucine-zipper proteins. Mutagenesis of leucine-zipper sequences have revealed the strong preference for leucines at the **d** position for stability of the leucine zipper (for a review, see Hu and Sauer, 1992), the role of charged amino acids at the **e** and **g** position for specificity of heterodimer formation between Fos and Jun (O'Shea et al., 1992) and the role of branched versus unbranched hydrophobic amino acids at the **a** and **d** position for oligomerization order (Harbury et al., 1993).

The leucine-zipper dimerization motif faces the challenge of forming a stable, short coiled coil, while retaining specificity for dimerization with the correct partner. In this appendix we examine the factors that contribute to coiled-coil structure and function, using as a model the well-described leucine-zipper from GCN4. Part A discusses the structural and functional effects of mutating a conserved asparagine buried in the hydrophobic core of the leucine zipper coiled coil. In part B, non-natural amino acids are used to measure the contribution to stability of each methylene group added to the sidechain at position 16 of the GCN4 leucine-zipper interface. The identification of a subdomain in the GCN4 coiled coil has implications for protein folding, as discussed in Appendix II.

## Appendix I

## **A. A Conserved Asparagine Destabilizes the Leucine Zipper of GCN4**

### **Introduction**

An asparagine at a central position of leucine-zipper sequences forms a hydrogen bond, buried in the hydrophobic interface of the GCN4-p1 dimer (O'Shea et al., 1991). While it is relatively rare for asparagines to occupy an a position in coiled coils (Asparagine is found at 3.7% of a positions in coiled coils from fibrous proteins; Parry, 1982), asparagine is highly conserved at the central a position in the leucine-zipper dimerization motif from the bZIP and bHLH-ZIP families of transcription factors (Hu and Sauer, 1992).

In an attempt to determine the role of this conserved, buried asparagine, we changed asparagine 16 of the GCN4 leucine zipper to valine (N16V; Fig. 1) and compared the stability, structure and DNA-binding function of wild-type and N16V peptides. We found a 45-degree increase in the stability of the N16V peptides with no change in DNA-binding specificity or affinity. A high-resolution crystal structure of the dimer reveals symmetric valines in place of the asymmetric, hydrogen-bonded asparagines of the wild-type structure. Subsequently, the N16V leucine zipper has been observed as a mixture of dimer and trimer species (Harbury et al., 1993), and the trimer structure has recently been solved by x-ray crystallography (Lino Gonzales, personal communication). Thus, the buried asparagine destabilizes the leucine zipper, but provides specificity for the dimer structure. Although this dimer specificity does not appear to be essential for DNA binding *in vitro*, yeast containing the N16V mutation in their only GCN4 allele exhibit a GCN4-defective phenotype (Daniel Kornitzer, unpublished result).

### **The N16V Mutant Forms a Super-Stable Leucine Zipper**

The concentration dependence of  $T_m$  suggests that the N16V peptide forms an oligomeric species (Fig. 2). Reversible thermal denaturation of N16V was monitored by circular dichroism (CD), at 222 nm, which is a spectroscopic probe for helical structure (Woody, 1985). N16V in 2M guanidine hydrochloride (GuHCl) exhibits a thermal stability

which is concentration dependent, with  $T_{ms}$  of  $46\text{ }^{\circ}\text{C} \pm 2\text{ }^{\circ}$  and  $60\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}$  at  $1\text{ }\mu\text{M}$  and  $5\text{ }\mu\text{M}$  monomer concentrations, respectively (Fig. 2a).

The proposal of a dimeric conformation for N16V predicts that a disulfide-linked dimer species should exhibit the concentration-independent stability of a unimolecular system. If N16V forms higher-order oligomers under these conditions, we would predict that the disulfide-linked dimer would exhibit a concentration dependence of  $T_m$ . In order to test whether the system is dimeric or multimeric, the N16V peptide was synthesized with an N-terminal linker: Acetyl-Cys-Gly-Gly. By air-oxidizing the cysteines in  $0.2\text{ M}$  Tris, pH 8.7, at room temperature, the hypothetical dimer is made unimolecular. Concentration-independent thermal stability in  $5\text{ M}$  GuHCl was confirmed by superimposing the thermal denaturation curves of  $1\text{ }\mu\text{M}$  and  $5\text{ }\mu\text{M}$  disulfide-bonded peptide monomers (Fig. 2b). In this concentration range, where the  $T_m$  is sensitive to peptide concentration, the  $T_{ms}$  of two repetitions at each concentration are the same:  $63\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}$ . Though by these methods it appears to form a dimer in GuHCl, the N16V peptide has been subsequently shown to form a mixture of dimers and trimers in the absence of GuHCl, though the disulfide-bonded N16V is a dimer, even at higher concentrations (Harbury et al., 1993).

The N16V mutant is more stable than the wild-type sequence (Fig. 3). Comparison of the relative thermal stability of N16V and wild-type GCN4 leucine zipper peptides reveals a difference of  $\sim 43 - 45\text{ }^{\circ}$  between  $T_{ms}$  of the two species in the same conditions. At pH 7.0 a thermal denaturation curve of  $34\text{ }\mu\text{M}$  wild-type peptide has a  $T_m$  of  $57\text{ }^{\circ}\text{C}$ . At identical conditions,  $34\text{ }\mu\text{M}$  N16V does not unfold completely, though approximately one-half of a transition can be seen for a melt from  $0-94\text{ }^{\circ}\text{C}$  (data not shown). In order to observe a full unfolding transition for N16V, the denaturant guanidine hydrochloride (GuHCl) was introduced to  $2\text{ M}$ . The difference between  $T_{ms}$  under these conditions is  $45\text{ }^{\circ} \pm 2\text{ }^{\circ}$  (Fig. 3a).

The effect of the N16V substitution on the free energy of dimerization was measured by monitoring GuHCl denaturation at  $0\text{ }^{\circ}\text{C}$  (Fig. 3b). A  $2.8\text{ kcal/mol}$  difference between the free energy of unfolding the wild-type and N16V coiled coils (see Table I) was calculated from GuHCl denaturation experiments for each of the peptides, by fitting the

unfolding transitions to an equation for 2-state bimolecular unfolding reactions (see Appendix III).

To evaluate the conformational effects of the N16V replacement, the x-ray crystal structure of this peptide was determined at 1.8 Å resolution (Fig. 4; Table II; see also Fig. 7). The structure of the N16V dimer is virtually identical to the structure of the wild-type peptide (GCN4-p1; O'Shea et al., 1991). The hydrophobic valine sidechains form symmetric packing interactions at the center of the coiled coil interface. By comparison, the wild-type asparagine at this position packs asymmetrically, forming a buried hydrogen bond with the asparagine from the other monomer.

### **DNA-binding by bZIP Peptides is Unaffected by the N16V Mutation**

The function of the intact GCN4 protein is to activate the expression of amino-acid genes in response to starvation via a general control mechanism, by binding to a GCN4 recognition element (GRE) as a dimer (Arndt and Fink, 1986; Hope and Struhl, 1987). Yeast containing the N16V mutation in their only GCN4 allele exhibit a poorly understood, defective phenotype, assayed as a 10-fold reduction in the activation of transcription of *his4*. The level of GCN4 protein expression is, however, unaltered (Daniel Kornitzer, unpublished result).

In spite of the defective phenotype *in vivo*, in the context of the bZIP DNA-binding domain, the N16V mutation has no effect on DNA-binding stability or specificity, as judged by DNase I footprinting experiments (Fig. 5). Peptides of the wild-type and N16V mutant bZIP domain were made by recombinant DNA methodologies with and without a disulfide linker, to examine the effects, if any, of forming a unimolecular bZIP domain (see methods, chapter 3; basic region plus leucine zipper: GG60 = wild-type; GG63 = disulfide-bonded wild type; GG60NV = N16V and GG63NV = disulfide-bonded N16V). DNase I footprint analysis was performed using trace amounts of the radioactively labelled DNA of the GRE from *His 3*, one of the genes activated by GCN4-binding, *in vivo*. The apparent DNA-binding affinity for GG60, GG63, GG60NV and GG63NV is between 0.1 and 0.3 μM. By comparison, the equilibrium constant for dimerization of GCN4-p1 is ~ 0.006 μM (Table 1), and the other three peptides are even more stable (not shown). Thus, the relatively weak DNA-affinities and strong dimerization affinities preclude measurement

of DNA binding at concentrations where the leucine-zipper monomer is populated significantly.

### **A Role for the Conserved Asparagine**

Sequence homology within the bZIP family suggests an important role for N16 in the leucine-zipper structure, yet the substitution of valine at this position results in a large increase in the stability of peptides corresponding to the coiled-coil region of GCN4. Selection of functional GCN4 leucine-zipper/lambda-repressor chimeras which bear substitutions at the GCN4 coiled-coil interface suggest that N16 is tolerant to substitution by a wide variety of amino acids, leading to a prediction that its role may not be structural (Hu et al., 1990). Alternatively, there may be a structural role, such as maintenance of register, oligomerization order, or orientation of the monomer helices in the dimer, which survives the stringency level of the selection *in vivo*.

In an effort to reveal the role of the conserved asparagine buried in coiled-coil dimer, the two N16 residues were removed, eliminating a hydrogen bond buried in the hydrophobic interface. This non-conservative substitution replaces two buried polar sidechains with two hydrophobic ones and is worth ~2.8 kcal/mol. Therefore, we conclude that the conserved N16 residue is destabilizing. A similar, but more dramatic effect is seen for a mutant of *E. coli* thioredoxin: the replacement of a buried aspartic acid residue with the non-polar amino acid, alanine, has been shown to increase the stability by 4.6 kcal/mol (Langsetmo et al., 1991).

The increased stability of N16V may be due to a combination of factors. First, the increased hydrophobicity of the dimer interface may be the cause of the increase in the stability of this structure. Testing this hypothesis requires stability measurements of increasingly hydrophobic sidechains substituted at the dimer interface (see Appendix I B).

Second, dehydrating the buried polar groups in the hydrophobic core of the coiled coil may be more energetically costly than the stability gained upon formation of a buried hydrogen bond (Bruce Tidor, personal communication). The decreased stability of the N16A mutant argues against this, although this mutant is the least well characterized of the series; thus, the instability of N16A may be due to other factors.

Third, solution studies of the wild-type structure by NMR revealed an elevated rate of exchange of the N16 amide, suggesting strongly that a polar sidechain at this position is destabilizing (Goodman and Kim, 1991). Solution studies of the N16V variant are complicated by the heterogeneity of conformations (mixture of dimer and trimer) in the absence of GuHCl, precluding accurate measurements of the exchange rate of the amide of N16 and comparison with the wild-type exchange rates (Kevin Lumb, unpublished results).

Finally, comparison of the crystal structures suggests that burial of the polar sidechains may also destabilize the coiled-coil structure by the disruption of a potential salt bridge between the dimers. The disruption of a potential salt bridge by N16 is observed in the structure of GCN4p1; however, this potential salt bridge is also absent in the atomic structure of N16V, though in this case the putative salt bridge is prevented by intermolecular crystal contacts (Alber, personal communication). Whether or not these sidechains can form a salt bridge awaits another crystal form which removes the crystal contacts.

Subsequent analysis of the N16V peptide revealed a mixture of dimer and trimer species at peptide concentrations higher than 5  $\mu$ M, in the absence of GuHCl (Harbury et al., 1993). In another study, the same N16V mutation was reported to form a trimeric species, 40 ° more stable than the wild-type sequence (Potekhin et al., 1994). Recent crystallographic studies have revealed the trimeric structure of N16V (Alber, personal communication). Thus, although the conserved asparagine destabilizes the coiled coil structure, it provides specificity for the dimeric conformation.

The conserved asparagine appears to be required for dimer specificity, yet removal of this specificity determinant apparently does not affect the DNA-binding specificity or stability, *in vitro* (note, however, that in the DNA-binding studies peptide concentrations were 3 $\mu$ M and less, where N16V is predominantly dimer; Harbury et al., 1993). What is the selective advantage of a conserved, destabilizing, buried polar sidechain for dimer specificity, if a stable trimer works just as well? Indeed, the heat-shock transcription factor (HSF) is a trimeric coiled coil in its active state (Sorger and Nelson, 1989). *In vivo*, the conserved asparagine of bZIP and bHLH-ZIP proteins may be essential for partner

specificity; introduction of valine may cause non-specific pairing with other transcription factors or other coiled coils in the cell. In addition, in the presence of the acidic activation domain of intact GCN4, the trimer may not interact productively with either the DNA site or the transcription machinery. A third possibility is that a trimeric bZIP protein may have a higher affinity for non-specific DNA; thus, transcription activation would be reduced, *in vivo*, because the transcription factor is not available for site-specific DNA binding.

## **B. Effects of Increased Hydrophobicity at the Coiled-Coil Interface**

### **Introduction**

It is now generally accepted that the major driving force for stabilizing proteins in the folded conformation comes from the burial of hydrophobic amino-acid sidechains, which interact unfavorably with water in the unfolded state (Kauzmann, 1959; Tanford, 1962; Dill, 1990; Lim and Sauer, 1989). Since 1962, protein chemists have been trying to quantify the contribution of hydrophobicity to protein stability. Early efforts toward this goal focused on measuring the free-energy of transfer of amino acids and other model compounds from apolar to polar solvents (Tanford, 1962; Nozaki and Tanford, 1971; Chothia, 1976; Wolfenden et al., 1981; Eisenberg and McLachlan, 1986).

With the advent of molecular biology, protein chemists have been able to exploit mutagenesis to measure the effects of amino-acid substitution on protein stability (reviewed in Matthews, 1993). In mutagenesis experiments, the buried hydrophobic surface area is changed by replacement of hydrophobic sidechains either with larger or smaller apolar sidechains. In each case, the effect on the stability of the protein is measured, and compared with the change in hydrophobicity (see Pace, 1992 for a synthesis and analysis of many studies). In some cases, the correlation between the stability and the hydrophobicity is in good agreement with the predicted values for model compounds in the non-polar to polar transfer experiments (Matsumura et al., 1988; Eriksson et al., 1992). Nonetheless, in other cases there is little correlation between stability and the hydrophobicity (Shortle et al., 1990; Serrano et al., 1992), or the change in free energy does not agree well with the transfer experiments (Karpusas et al., 1989; Kellis et al., 1988; Yutani et al., 1987; Sandberg and Terwilliger, 1991).



Disagreement between the results obtained from separate mutagenesis studies are likely to be due to a combination of three factors: differences between the shape of the hydrophobic sidechains, fraction of sidechain accessibility to solvent and compensatory changes in the core of the protein (reviewed in Matthews, 1993). Here, we use non-natural amino acids to measure the contribution of methylene groups substituted in the center of the GCN4 coiled-coil core. By measuring the free energy of unfolding the coiled coils, we observed ~1.2 kcal/mol for each methylene group substituted in the coiled-coil core, with the exception of the norleucine variant. X-ray crystallography of three of the mutants revealed that the structures are unchanged from the wild-type except for the symmetrically oriented, non-natural sidechains at position 16. The N16NLeu mutant is less stable than N16NVal, and the structure reveals that the norleucine sidechain is too large for complete burial and optimal packing at this position in the coiled-coil core.

We sought to measure experimentally the effect of discrete methylene additions to the hydrophobic core of a protein. The 33 amino-acid, GCN4 leucine-zipper peptide was chosen for the analysis because it is a simple, helical structure, amenable to studies of stability and structure and because substitutions can be made by peptide synthesis, increasing the repertoire of chemical changes that can be introduced. Non-natural amino acids were substituted into the hydrophobic core, to avoid introducing additional steric constraints associated with the branched sidechains of naturally occurring amino acids. The straight-chain aliphatic sidechains of alanine (-CH<sub>3</sub> ; N16A), amino-butyric acid (-CH<sub>2</sub>CH<sub>3</sub> ; N16Abu), norvaline (-CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub> ; N16NVal) and norleucine (-CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub> ; N16NLeu) all have approximately the same helical propensity as alanine (Padmanabhan and Baldwin, 1991), and increase in hydrophobicity without branching or forming a ring structure, either of which might disrupt helical or coiled-coil interactions. Position 16 in the GCN4 leucine zipper was chosen as the site of amino-acid substitution. Asparagine 16 (N16) forms a hydrogen bond with its symmetry mate in the center of the coiled coil dimer (O'Shea et al., 1991). This position is conserved among leucine zippers in nature, but highly tolerant to amino-acid substitution, as determined from a screen for mutants of the hydrophobic interface (Hu et al., 1990). Furthermore, it was known that N16 to valine dramatically increases the stability of this peptide (see section A).

## Coiled-coil Stability is Correlated with Increased Hydrophobicity

The stability of each coiled-coil variant was assessed from GuHCl denaturation experiments, by monitoring coiled-coil structure by CD at pH 7.0 and 0 °C (Fig. 6). The change in free energy of unfolding (extrapolated to 0 M GuHCl; see Appendix III) was determined for each of the variants (Table I) using a bimolecular-unfolding equilibrium model applied to the denaturation curves.

The peptide variants at position 16 were compared with the wild-type structure (GCN4-p1). All but one of the variants were much more stable than GCN4-p1. N16Abu, N16NVal and N16NLeu crystallized in the same space group as GCN4-p1 and revealed structures nearly identical to the wild-type structure (Fig. 7 and Table II). N16A, which is less stable than the wild-type leucine zipper, crystallized in a different spacegroup, and that structure has yet to be determined (Lino Gonzales, personal communication).

Gel filtration chromatography of the variants revealed a dimer species for each of the variants at concentrations of GuHCl lower than those required for the unfolding transitions observed by circular dichroism (Fig. 8a-e). Nonetheless, the peptides form a mixture of dimer and trimer species in the absence of GuHCl (Pehr Harbury, unpublished result; Fig. 8f). N16A forms a species slightly heavier than the dimer standard at room temperature (see Fig. 8f), yet measurements in the presence of GuHCl are not possible, because N16A is significantly unfolded in the presence of even low concentrations of GuHCl. Because the peptides are dimeric in the folded baseline of the guanidine HCl denaturation curves, we argue that the free energy measurements from the GuHCl denaturation curves represents the free energy of unfolding the coiled-coil dimers. Thus fitting the denaturation curves using a bimolecular unfolding model is probably a reasonable method to determine the unfolding free energy of these coiled coil variants.

A linear correlation is observed between the increase in the free-energy of folding and an increase in the number of buried methylene groups (Fig. 9), with one exception: N16NLeu is less stable than the less hydrophobic N16NVal. The decreased stability of the norleucine variant may be explained by its structure. The packing of the norleucine sidechains is not optimal in the x-ray structure, resulting in incomplete burial of the terminal

methyl group (see Fig 7e). The structures of N16Abu and N16NVal, by contrast, show complete burial and optimal packing of the sidechains at position 16 (see Fig. 7c, d; Tom Alber and Lino Gonzales, personal communication).

### **Contribution of the Buried Methylene Group to Stability**

If the major source of protein stability is burial of hydrophobic sidechains, then the stability of a protein should increase linearly with incremental additions of non-polar groups to the protein core, until the increase in sidechain size can no longer be accommodated by the protein without adjustments. This hypothesis was tested with one natural and three non-natural variants of the GCN4 leucine zipper. We found that, until it is too large to be fully buried, each methylene group added to position 16 of the GCN4 leucine zipper increases the stability by  $\sim 1.2$  kcal/mol (see Fig. 9b). This number is in good agreement with  $1.3 \pm 0.5$  kcal/mol for the burial of one methylene group, as reported by Pace (Pace, 1992), in a comprehensive synthesis and re-evaluation of the existing mutants available in the literature.

Subsequent to these studies, a similar analysis was performed with T4 lysosyme, in which a core position was substituted by aliphatic, non-natural amino acids; however, the increase in stability was not linearly correlated with the increase in hydrophobicity of the buried sidechain (Ellman et al., 1992). Those authors concluded that in addition to the hydrophobic contribution, the effects of these substitutions on core packing were likely to contribute significantly to the protein stability. We observe a linear correlation between stability and buried hydrophobicity, except for the norleucine sidechain, which adopts a partially exposed conformation in the x-ray structure, as a result of the steric constraints associated with packing the entire sidechain into the hydrophobic interface.

Although hydrophobicity does appear to play a major role in stabilizing the leucine-zipper coiled coil, we now know that other considerations may be more important for maintaining the leucine-zipper structure. Obviously, there is a limit to the size of the sidechain that can be accommodated by the coiled-coil core, as seen for the over-large norleucine substitution, which destabilizes the coiled coil relative to the less hydrophobic norvaline variant. As discussed in Appendix I A, removal of the buried asparagine at position 16 increases the stability of the coiled coil, but at the price of losing specificity for

a dimeric structure: N16V forms a mixture of dimers and trimers (see also Harbury et al., 1993). Likewise, the non-natural variants of GCN4 also have a tendency to form higher-order oligomers, reinforcing the importance of the buried polar group to form a unique dimeric species. In addition, the shape of the hydrophobic amino acids in the coiled coil core is critical, as the occurrence of  $\beta$ -branched sidechains either at position **a** or **d**, or both, determines the oligomerization order of the coiled coil (Harbury et al., 1993).

## Experimental Procedures

### Chemicals

Na<sub>2</sub>HPO<sub>4</sub>, NaH<sub>2</sub>PO<sub>4</sub> are from Fisher, NaCl from Mallinckrodt, CH<sub>3</sub>CN from EM Science, TFA from Applied Biosystems, CH<sub>3</sub>COOH from Baker and EM Science, HCl from Baker, NaOH from Sigma, guanidine HCl, (GuHCl), from Schwarz/Mann.

### Solutions

CD Buffer: 50 mM Na Phosphate, 150 mM NaCl, pH 7.0  
Guanidine Stocks: in CD buffer, 8M GuHCl, as determined by refractive index (Nozaki, 1972). Alternatively, stocks were made from 0-8 M in 1 M intervals in the CD buffer, adjusted to pH 7.0 and frozen at -20 °C (by Pehr Harbury).

### Peptide Synthesis and Purification

Peptides were synthesized on an Applied Biosystems model 430A peptide synthesizer using Fmoc chemistry with Fastmoc reaction cycles modified to include acetic anhydride capping (Fields and Noble, 1990). In each peptide, the N-terminus is acetylated and the C-terminus is amidated. The peptides were cleaved using standard Fmoc protocols and desalted on a Sephadex G-10 or G-25 column (Pharmacia) in 5% acetic acid. Final purification was by reverse-phase, high-performance liquid chromatography (HPLC, Waters, Inc.) at 25° C using a Vydac preparative or semipreparative C18 column (stock# 218TP1022 or 218TP510, respectively). A linear acetonitrile-H<sub>2</sub>O gradient of 0.1% buffer B increase per minute was used with a flow rate of 10 ml/min (5 ml/min for the semipreparative column). Buffer A is 0.1% trifluoroacetic acid (TFA) in water; buffer B is 90% acetonitrile and 0.1% TFA in water. The identity of the peptides was confirmed  $\pm$  1 mass unit, by laser desorption mass spectrometry (Mass Search, Modesto, CA).

### Peptide Stocks

Stocks were prepared from HPLC purified, lyophilized peptides in H<sub>2</sub>O or 1 mM HCl. The stocks were filtered and purity was assessed by analytical HPLC on a VYDAC C18 column and a Waters HPLC system, monitored at 229 nm with a Waters Multiwavelength detector. Concentration of the stocks was determined by absorbance spectroscopy on an AVIV uv/vis 14 DS spectrometer at 275.5 nm, as described (Edelhoch, 1967). Each peptide monomer has one tyrosine. 5 repetitions were taken from 320-240 nm at 0.5 nm intervals and 0.3 second averaging time. The spectra were averaged and the blank repetitions, averaged, were subtracted. The remaining, sloped baseline offset was subtracted by fitting a line to data points from 320 to 310 nm (where tyrosine absorbance should be zero) and subtracting it from the averaged spectrum. This gave us the most reproducible results.

## Circular Dichroism Spectroscopy

Helical structure was monitored by circular dichroism spectroscopy. Pure, concentrated peptide was added in fixed amounts to 34  $\mu\text{M}$  in 1 X CD buffer or GuHCl stocks (2 mls, final volume) and the pH adjusted with HCl and NaOH. Each point from 0-6.5 M GuHCl was made independently for the initial studies. Later studies utilized flanking samples, i. e. 1.5 M GuHCl sample was made from 1 M and 2 M samples mixed 1:1, yielding results at least as consistent as the longer method, with less peptide. Blanks were made identically for every point and their signal was subtracted. Ellipticity was monitored at 222 nm, a minimum in the spectrum of  $\alpha$ -helices, as a function of guanidine concentration or temperature. The instruments used are AVIV spectrometers models 60 DS and 62 DS. Spectra were taken from 300 nm to 208 nm, with a 0.2 second averaging time, at 0  $^{\circ}\text{C}$  to verify helical structure. The signal at 222 nm was monitored in kinetics mode at 0  $^{\circ}\text{C}$  for 500-600 seconds with a 1 second averaging time. Data from 250-500 or 600 seconds (after temperature equilibration), was averaged and blanks monitored identically were averaged and subtracted. Molar ellipticity was calculated as described.

## Crystallography

Crystallographic studies were carried out in the laboratory of Tom Alber. The quality of the data are described by the parameters in Table II. Figures were generated by the program Insight II (Biosym) from the atomic coordinates of the refined structures, provided by the Alber lab.

## Free Energy Calculations

As a quantitative measure of stability, the free energy of dimerization and the dissociation constants were determined for GCN4-p1, N16A, N16Abu, N16NVal, and N16NLeu. For free energy measurements, we assumed a two-state model for the unfolding equilibrium of the coiled-coil dimer, according to the equation:  $P_2 \rightarrow P + P$  (see Appendix III). The peptides are fully dimeric at low concentrations of GuHCl (i.e. in the folded baseline of the denaturation curves; see gel filtration), although the peptides are a mixture of dimer and higher-order oligomers in the absence of GuHCl, as judged by gel filtration using known GCN4 dimer, trimer and tetramers as standards (Harbury et al., 1993; see gel filtration). The dimerization free energies were determined by fitting the GuHCl denaturation data to a sigmoidal curve which describes this two-state bimolecular transition. The equation expresses the observable parameter, molar ellipticity, as the fraction of folded peptide and corrects the resulting free energy for the effects of GuHCl on unfolding using a linear approximation for this effect (Pace, 1986). This equation was fit to GuHCl denaturation curves of 34  $\mu\text{M}$  peptide concentrations at pH 7.0, 0 $^{\circ}\text{C}$ . The baselines were fit to lines with fixed slopes. The y-intercept of the unfolded baseline was also fixed, while that of the folded baseline was allowed to vary.

The  $\Delta G^{\circ}$  reported is for an equilibrium between species at a concentration of 1M for each species (i.e. 1M dimer dissociates to 1M monomer plus 1M monomer). Standard state is defined here as 0  $^{\circ}\text{C}$ , 50 mM  $\text{NaPO}_4$ , 150 mM  $\text{NaCl}$ , pH 7.00  $\pm$  0.01 and 1 M in all peptide species. The standard-state free energies for mutant and wild-type peptides is reported in Table I. The equilibrium dissociation constants were determined directly from the standard-state free energies by the equation:

$$K_d = e^{-\Delta G^{\circ}/RT}$$

## DNase I Footprinting

DNA-binding affinity and specificity were assessed by DNase I protection experiments (Garner and Revzin, 1981), as described for GCN4-bZIP1 (Talanian et al., 1990) using GCN4 bZIP peptides and the His 3 GCN4-recognition element (GRE) as the DNA-recognition site: GCGGA TGACTC TTTTTTTT (this is the same DNA as used in Talanian et al., 1990, except a cryptic binding site was omitted by single-stranded mutagenesis of the plasmid to yield the plasmid pCGRE). GG60 (O' Shea, unpublished result) is identical to GCN4-bZIP1 except it was made by recombinant DNA methodologies. GG63 is GG60 with a C-terminal Gly-Gly-Cys linker, which stabilizes the dimer by covalent attachment upon oxidation. The N16V variants of GG60 and GG63 were made from pGG60 by single stranded mutagenesis (Kunkel, 1985), and expressed using the method of Studier (Studier et al., 1990). Peptide concentration ranged from 3  $\mu$ M to 0.01 $\mu$ M. The apparent DNA-binding affinity for all peptides is between 0.1 and 0.3  $\mu$ M, whereas the equilibrium constant for dimerization of GCN4-p1 is  $\sim$ 0.01 $\mu$ M (see Table I).

## References

Arndt, K., and Fink, G. R. (1986). GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5' TGACTC 3' sequences. *Proc. Natl. Acad. Sci. U S A* *83*, 8516-20.

Baxevanis, A. D., and Vinson, C. R. (1993). Interactions of coiled coils in transcription factors: where is the specificity? *Curr. Opin. Genet. Dev.* *3*, 278-85.

Blackwood, E. M., and Eisenman, R. N. (1991). Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* *251*, 1211-7.

Chothia, C. (1976). The nature of the accessible and buried surfaces in proteins. *Journal of Molecular Biology* *105*, 1-14.

Cohen, C., and Parry, D. A. (1994). Alpha-helical coiled coils: more facts and better predictions [published erratum appears in *Science* 1994 May 20; 264(5162):1068]. *Science* *263*, 488-9.

Crick, F. H. C. (1953). The packing of alpha-helices: Simple coiled coils. *Acta. Cryst.* *6*, 689-697.

Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry* *29*, 7133-7155.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* *6*, 1948-1954.

Eisenberg, D., and McLachlan, A. D. (1986). Solvation energy in protein folding and binding. *Nature* *319*, 199-203.

Ellman, J. A., Mendel, D., and Schultz, P. G. (1992). Site-specific incorporation of novel backbone structures into proteins. *Science* *255*, 197-200.

Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Mathews, B. W. (1992). Response of a Protein Structure to Cavity-Creating Mutations and Its Relation to the Hydrophobic Effect. *Science* *255*, 178-83.

Fields, G. B., and Noble, R. L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *International Journal of Peptide and Protein Research* *35*, 161-214.

Garner, M. M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* *9*, 3047-60.

Goodman, E. M., and Kim, P. S. (1991). Periodicity of amide proton exchange rates in a coiled-coil leucine zipper peptide. *Biochemistry* *30*, 11615-11620.

- Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between two-, three-, and four stranded coiled coils in GCN4 leucine zipper mutants. *Science* 262, 1401-1407.
- Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* 262, 1401-1407.
- Hodges, R. S., Sodek, J., Smillie, L. B., and Jurasek, L. (1972). Tropomyosin: Amino acid sequence and coiled-coil structure. *Cold Spring Harbor* 37, 299-310.
- Hope, I. A., and Struhl, K. (1987). GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.* 6, 2781-4.
- Hu, J. C., O'Shea, E. K., Kim, P. S., and Sauer, R. T. (1990). Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* 250, 1400-3.
- Hu, J. C., and Sauer, R. T. (1992). The basic-region leucine-zipper family of DNA binding proteins. *Nucleic Acids and Molecular Biology* 6, 82-101.
- Karpusas, M., Baase, W. A., Matsumura, M., and Matthews, B. W. (1989). Hydrophobic packing in T4 lysozyme probed by cavity-filling mutants. *Proc. Natl. Acad. Sci. U S A* 86, 8237-41.
- Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. *Advances in Protein Chemistry* 14, 1-63.
- Kellis, J., Jr., Nyberg, K., Sali, D., and Fersht, A. R. (1988). Contribution of hydrophobic interactions to protein stability. *Nature* 333, 784-6.
- Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U S A* 82, 488-92.
- Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759-1764.
- Langsetmo, K., Fuchs, J. A., and Woodward, C. (1991). The conserved, buried aspartic acid in oxidized Escherichia coli thioredoxin has a pKa of 7.5. Its titration produces a related shift in global stability. *Biochemistry* 30, 7603-9.
- Lim, W. A., and Sauer, R. T. (1989). Alternative packing arrangements in the hydrophobic core of lambda repressor. *Nature* 339, 31-6.
- Matsumura, M., Becktel, W. J., and Matthews, B. W. (1988). Hydrophobic stabilization in T4 lysozyme determined directly by multiple substitutions of Ile 3. *Nature* 334, 406-10.
- Matthews, B. W. (1993). Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* 62, 139-60.



- McLachlan, A. D., and Stewart, M. (1975). Tropomyosin coiled-coil interactions: Evidence for an unstaggered structure. *Journal of Molecular Biology* 98, 293-304.
- Nozaki, Y. (1972). The preparation of guanidine hydrochloride. *Methods in Enzymology* 26, 43-50.
- Nozaki, Y., and Tanford, C. (1971). The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J. Biol. Chem.* 246, 2211-2217.
- O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. A. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* 254, 539-544.
- O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992). Mechanism of Specificity in the Fos-Jun Oncoprotein Heterodimer. *Cell* 68, 699-708.
- Oas, T. G., and Endow, S. A. (1994). Springs and hinges: dynamic coiled coils and discontinuities. *Trends Biochem. Sci.* 19, 51-4.
- Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* 131, 266-280.
- Pace, C. N. (1992). Contribution of the hydrophobic effect to globular protein stability. *Journal of Molecular Biology* 226, 29-35.
- Parry, D. A. D. (1982). Coiled-coils in alpha-helix-containing proteins: analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. *Bioscience Reports* 2, 1017-1024.
- Pauling, L., and Corey, R. B. (1953). Compound helical configurations of polypeptide chains: Structure of proteins of the  $\alpha$ -keratin type. *Nature* 171, 59-61.
- Potekhin, S. A., Medvedkin, V. N., Kashparov, I. A., and Venyaminov, S. U. (1994). Synthesis and properties of the peptide corresponding to the mutant form of the leucine zipper of the transcriptional activator GCN4 from yeast. *Protein Engineering* 7, 1097-1101.
- Sandberg, W. S., and Terwilliger, T. C. (1991). Energetics of repacking a protein interior. *Proc. Natl. Acad. Sci. U S A* 88, 1706-10.
- Serrano, L., Kellis, J., Jr., Cann, P., Matouschek, A., and Fersht, A. R. (1992). The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* 224, 783-804.
- Shortle, D., Stites, W. E., and Meeker, A. K. (1990). Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. *Biochemistry* 29, 8033-41.
- Sorger, P. K., and Nelson, H. C. (1989). Trimerization of a yeast transcriptional activator via a coiled-coil motif. *Cell* 59, 807-13.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* *185*, 60-89.

Talanian, R. V., McKnight, C. J., and Kim, P. S. (1990). Sequence-specific DNA binding by a short peptide dimer. *Science* *249*, 769-71.

Tanford, C. (1962). Contribution of Hydrophobic Interactions to the Stability of the Globular Conformation of Proteins. *J. A. C. S.* *84*, 4240-7.

Wolfenden, R., Andersson, L., Cullis, P. M., and Southgate, C. C. B. (1981). Affinities of amino acid side chains for solvent water. *Biochemistry* *20*, 849-855.

Woody, R. W. (1985). Circular Dichroism of Peptides. *The Peptides*. vol. 7 Udenfriend, S., and Meienhofer, J. and Hruby, V. J., eds (Academic Press, Inc: Orlando, Florida) 15-114.

Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. (1987). Dependence of conformational stability on hydrophobicity of the amino acid residue in a series of variant proteins substituted at a unique position of tryptophan synthase alpha subunit. *Proc. Natl. Acad. Sci. U S A* *84*, 4441-4.

## Table I      Stability of GCN4 Leucine Zipper Peptides

Denaturation curves were fit to a 2-state model and the change in free energy of unfolding ( $\Delta G^\circ$ , kcal/mol) was extrapolated to 0 M GuHCl (see Appendix III). Molar ellipticity  $[(\theta)_N]$ ; at 222 nm;  $\text{deg cm}^2 \text{dmol}^{-1}$ ) represents the helicity of the fully folded species at 0 °C, in the absence of GuHCl. The constant “m” corrects the resulting free energy for the effects of GuHCl on unfolding using a linear approximation for this effect:  $\Delta G = \Delta G_{\text{H}_2\text{O}} - m [\text{GuHCl}]$  (Pace, CN, [1986] Meth. Enz. 131, 266, 279 ). The equilibrium constant ( $K_d$ ; M) is calculated directly from  $\Delta G^\circ$  by the equation:  $K_d = e^{-\Delta G^\circ/RT}$ .

Table 1: Stability of GCN4 Leucine Zipper Peptides

Peptide	$\Delta G^\circ$	$\theta_N$	m	$K_d$
Wild-type	10.3	-32,500	1.8	$5.8 \times 10^{-9}$ M
N16A	9.7	-33,200	2.0	$1.7 \times 10^{-8}$ M
N16Abu	12.3	-33,100	1.9	$1.5 \times 10^{-10}$ M
N16NorV	14.5	-32,100	2.0	$2.5 \times 10^{-12}$ M
N16NorL	12.9	-31,400	1.8	$5.3 \times 10^{-11}$ M

**Table II      Crystal Parameters for GCN4 Peptides**

Parameters for the x-ray crystal structures of GCN4 variants (provided by Tom Alber and Lino Gonzales).

Peptide	N16V	N16Abu	N16NVal	N16NLeu
Space group	C2	C2	C2	C2
Unit Cell (a,b,c, beta)	101.99, 30.49, 21.79,beta=95.48	101.87, 30.35, 21.80,beta=95.19	102.54, 30.22, 21.99, beta=96.47	102.85, 30.01, 22.14, beta=96.82
Resolution	1.8 Å	1.8 Å	1.7 Å	1.8 Å
Measured Reflections	5178/6308	14869	5096/7314	Not determined
Data completeness	Not determined	5897	Not determined	Not determined
Unique reflections	Not determined	94%	Not determined	Not determined
R-merge	0.32	0.034	0.049	Not determined
R-crystallographic	0.173	0.182	0.184	0.181
rmsd bond length	0.020	0.020	0.016	0.017
rmsd bond angle	2.754	2.9	2.685	2.835

## Figure 1

The leucine zipper of GCN4. (a) Sequences of the GCN4 leucine zipper peptides: wild-type (GCN4-p1) and N16V, which contains valine in place of asparagine at position 16. Peptides correspond to the last 33 amino-acids of intact GCN4, and each is acetylated at its N-terminus. (b) Helical wheel projection of the GCN4 leucine-zipper coiled coil, GCN4-p1 (from O'Shea et al, 1989). The view is down the superhelical axis, beginning at the N-terminus, with position **a**, methionine (M; residue number 2 of the GCN4 leucine-zipper sequence). The dimer interface is composed of predominantly hydrophobic amino acids, with the exception of asparagine (N) at position **a**, which is a conserved feature of leucine zippers. The residues at positions **b**, **c**, **e**, **f** and **g** are predominantly hydrophilic. Arginine 1 (R) at the N-terminus is omitted from this view, but would fall at position **g** in the helical wheel.

Amino acid abbreviations:

A = alanine; D = aspartic acid; E = glutamic acid; F = phenylalanine;

G = glycine; H = histidine; I = isoleucine; K = lysine; L = leucine;

M = methionine; N = asparagine; Q = glutamine; R = arginine; S = serine; T = threonine; V = valine; Y = tyrosine.

FIGURE 1a

GCN4-p1 :

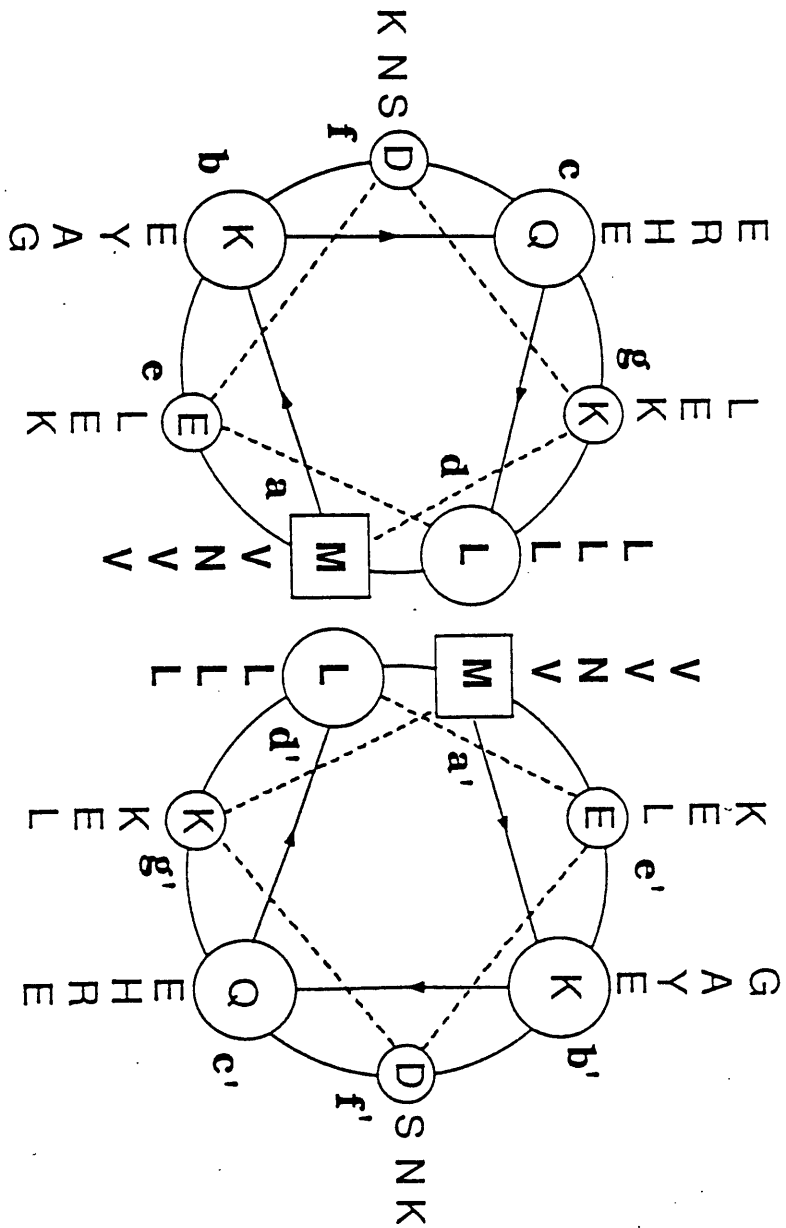
Ac-RMKQLEDKVEELL**SKN**YHLENEVARLKKLVGER

N16V :

Ac-RMKQLEDKVEELL**SKV**YHLENEVARLKKLVGER



FIGURE 1b



GCN4-p1

GCN4-p1

## Figure 2

Concentration dependence of the thermal stability of N16V and disulfide-linked N16V peptides. Helicity is monitored by circular dichroism spectroscopy (CD) as a function of temperature at pH 7.0. More negative values of molar ellipticity ( $[\theta]_{222}$ ; y-axis) represent the higher degree of helicity in the folded peptide. (a) N16V displays a concentration dependence of  $T_m$  between 1 and 5  $\mu\text{M}$  peptide. Peptide samples include 2 M GuHCl to facilitate thermal denaturation. Open circles = 1  $\mu\text{M}$  peptide with a thermal stability of  $\sim 48^\circ\text{C}$ ; open diamonds = 5  $\mu\text{M}$  peptide with a thermal stability of  $\sim 60^\circ\text{C}$ . (b) Disulfide-linked N16V displays a concentration independent stability, with a  $T_m$  of  $\sim 53^\circ\text{C}$  in 5 M GuHCl at both 1  $\mu\text{M}$  (open circles) and 5  $\mu\text{M}$  (filled circles).

FIGURE 2a

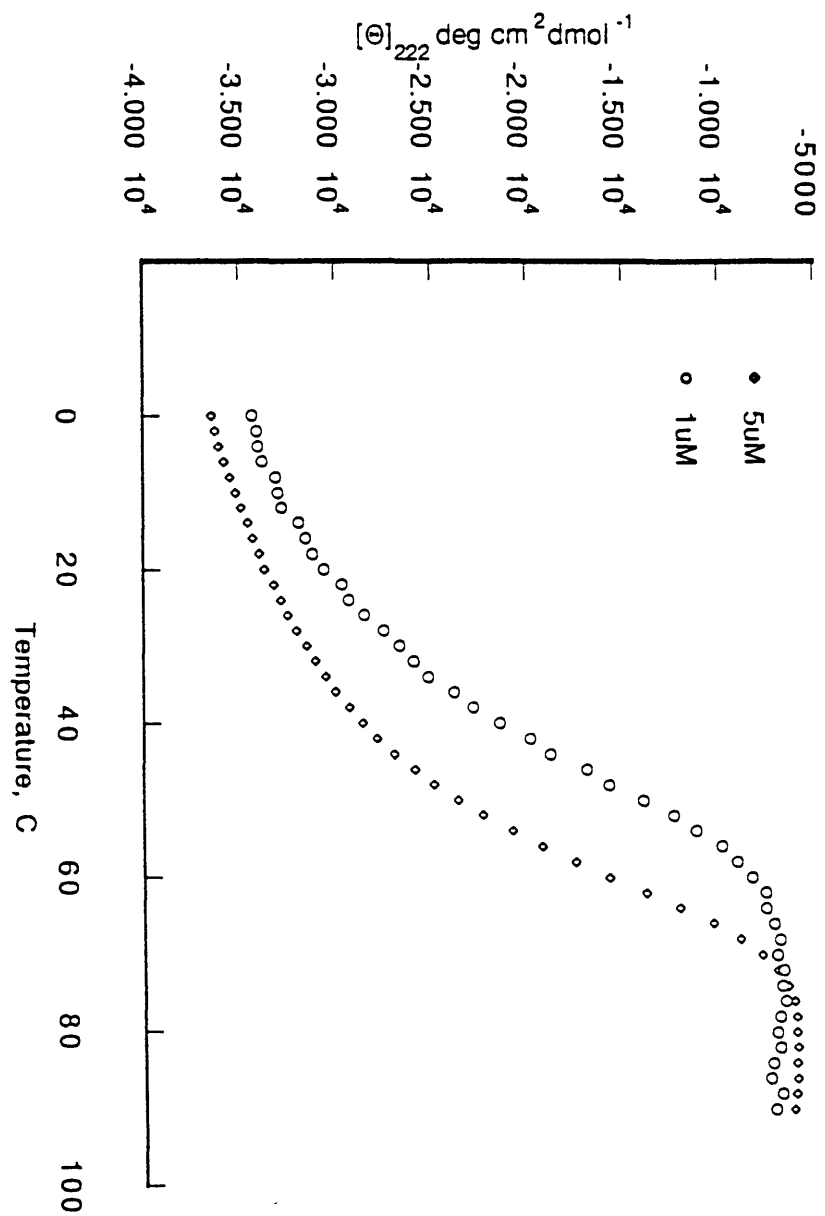
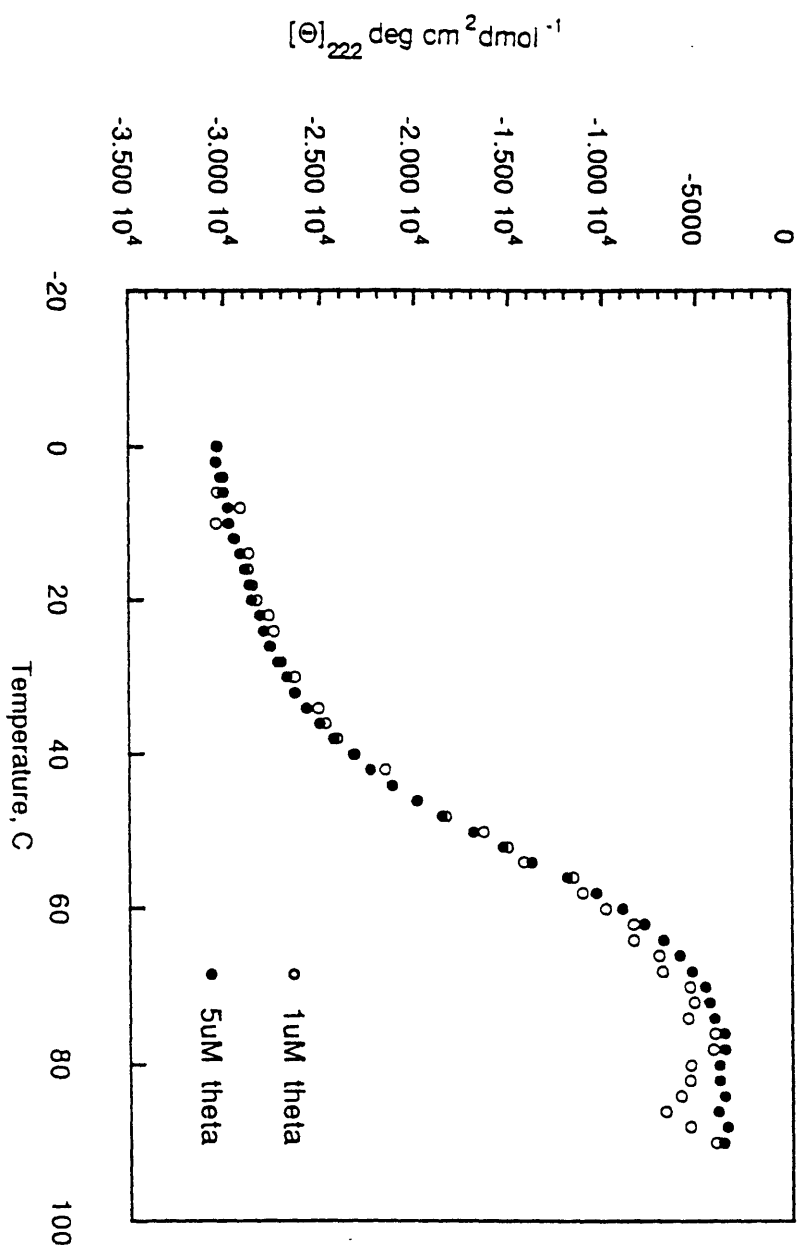


FIGURE 2b



### Figure 3

Comparison of the stabilities of wild-type and N16V peptides. Experiments are performed by CD, as described in Fig. 2. (a) N16V is ~ 45 degrees more stable than the wild-type peptide. In 2 M GuHCl, pH 7.0, the  $T_m$  of 34  $\mu$ M GCN4-p1 is ~ 29 °C (open diamonds), whereas the  $T_m$  of 34  $\mu$ M N16V is ~74 °C (filled diamonds). (b) N16V is more stable than the wild-type peptide to denaturation by GuHCl at pH 7.0 and 0 °C. 34  $\mu$ M GCN4-p1 unfolds with a denaturation midpoint of ~ 2.5 M GuHCl, whereas 34  $\mu$ M N16V unfolds with a denaturation midpoint of ~ 4.0 M GuHCl. Denaturation curves were fit to a 2-state model (see text) and the change in free energy of unfolding was extrapolated to 0 M GuHCl (see Table 1).

FIGURE 3a

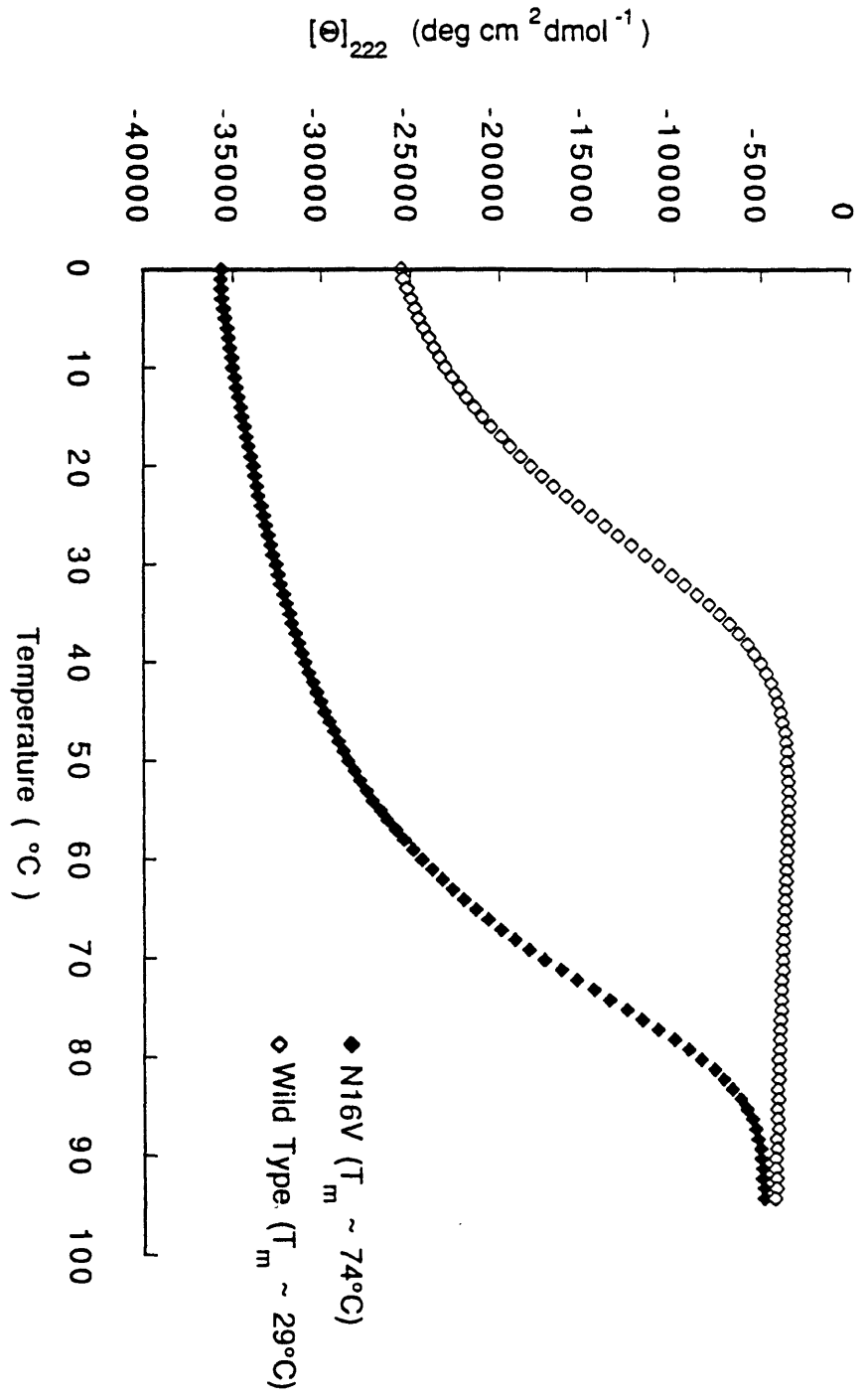
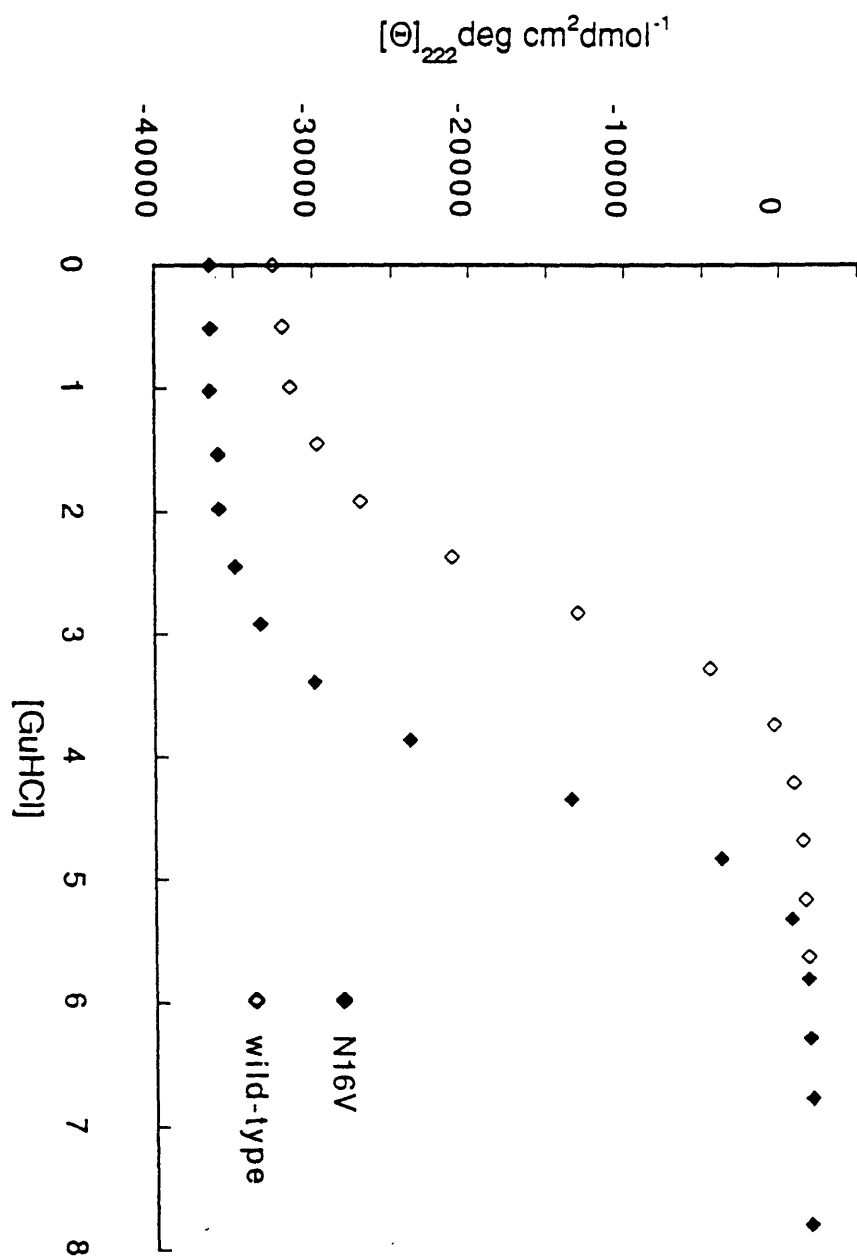


FIGURE 3b



#### **Figure 4**

Superposition of the refined, x-ray crystal structures of N16V (pink) and GCN4-p1 (multi-colored). One level (residues 15-17) of the coiled-coil dimer is shown from an axial view. The hydrogen-bonded, asymmetric asparagines of GCN4-p1 are shown colored by atom (red = oxygen; blue = nitrogen; white = hydrogen) while the symmetric, aliphatic valine sidechains of N16V are depicted in pink. Except for position 16, the structures are nearly identical (see, for example, the neighboring tyrosine ring).





## Figure 5

Specificity and stability of DNA-binding by bZIP peptides of GCN4 and N16V. DNase I protection experiments (see methods) were used to monitor the stability and specificity of GCN4 bZIP peptides for the His 3 DNA site. Peptide concentrations range from 10 nM to 3 $\mu$ M. (a) The His 3-protection pattern seen for wild-type (GG60) GCN4 bZIP peptides is unchanged by the addition of a disulfide-linkage (GG63 s-s). The same pattern is also seen for the N16V mutants (GG60 N16V and GG63 N16V); disulfide-crosslinking did not change the protection pattern. The DNase I protection experiments were quantitated by phosphorimagery and the intensity of the protected bands ( $[\text{Intensity of the His 3 site} - \text{Intensity of background of the same size}] / [\text{total Intensity in the lane} - \text{intensity of the same size region of background}]$ ) is plotted as a function of the logarithmic increase in peptide concentration. (b) Wild-type bZIP peptides bind with an apparent affinity of 0.3  $\mu$ M, whether or not disulfide-linked. (c) N16V mutant bZIP peptides bind with an apparent affinity of 0.3  $\mu$ M, whether or not disulfide-linked. Open circles represent the non-crosslinked species; filled circles represent the cross-linked species.

Wild-Type and N16V b-Zip Peptides Footprint the His 3 Site Equally Well

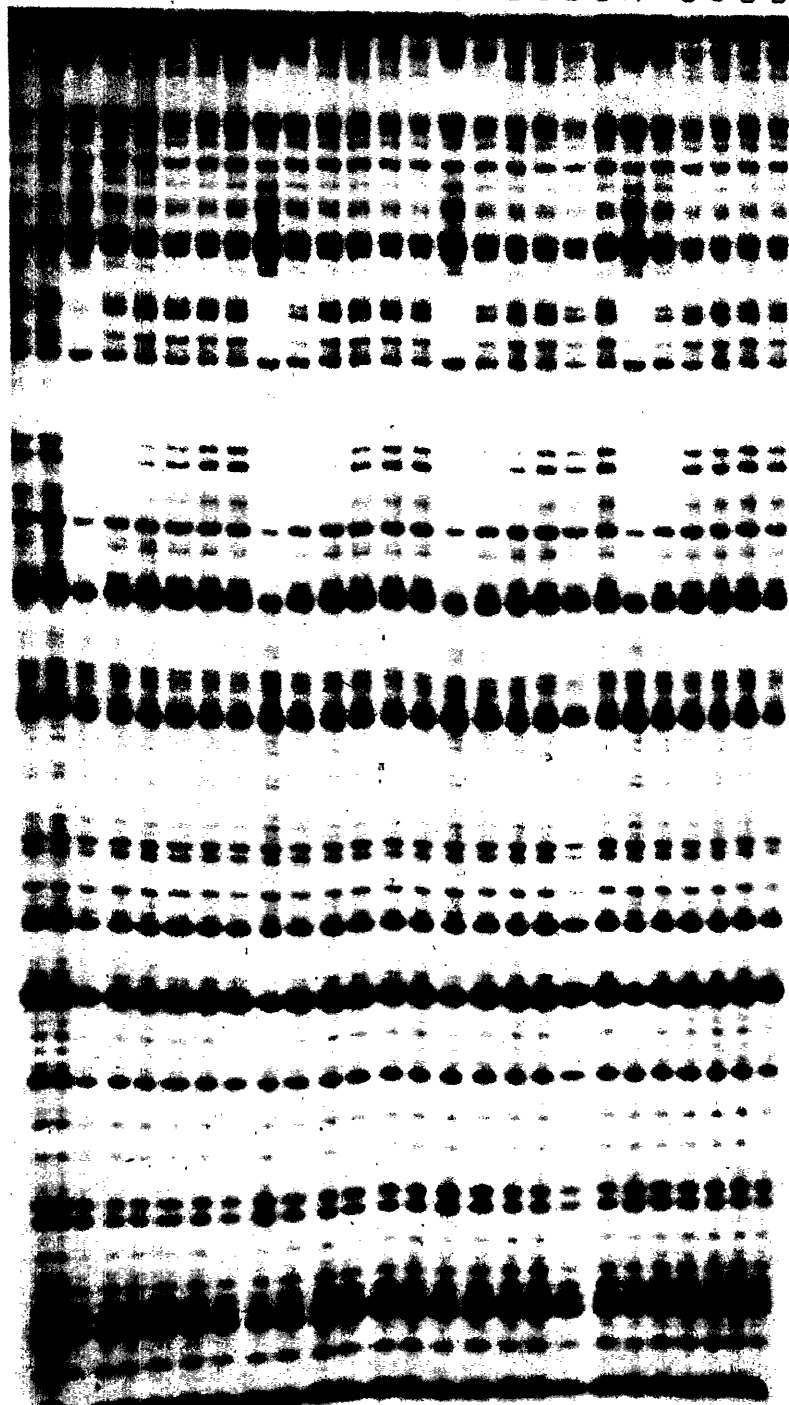
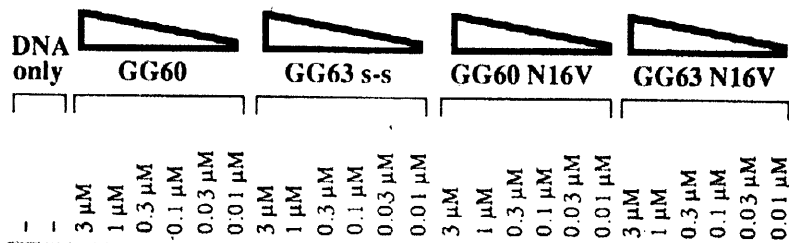


FIGURE 5b

His 3 Footprint Analysis: GG60, GG63

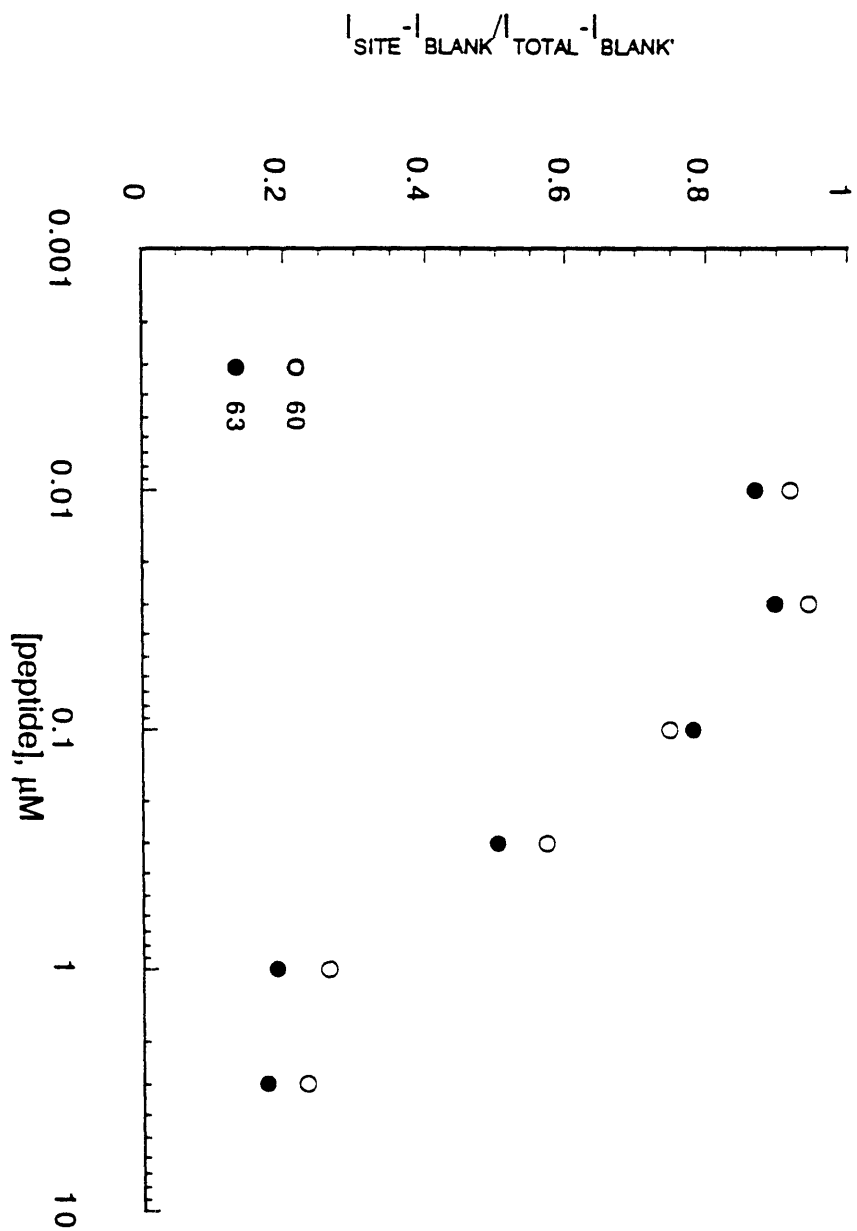
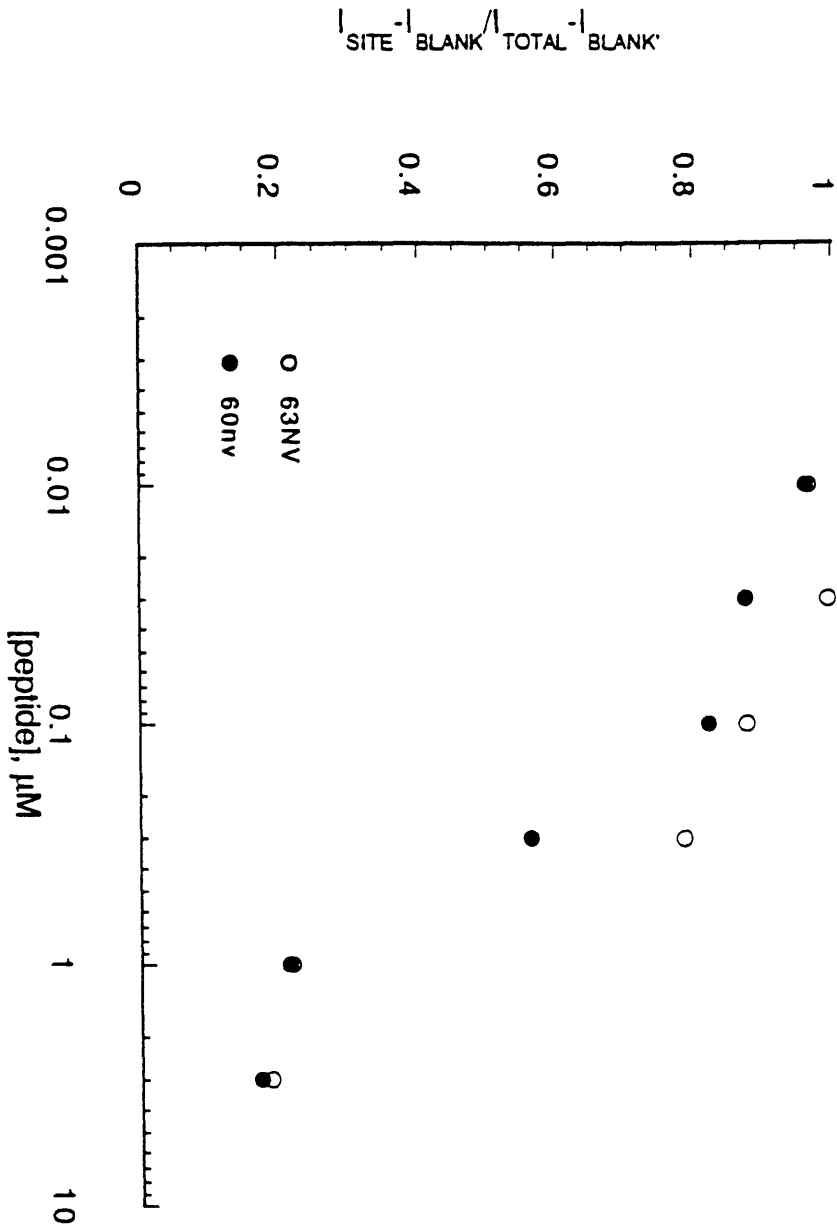


FIGURE 5c



## Figure 6

Guanidine HCl denaturation of GCN4 leucine-zipper variants. Helicity is monitored by CD as a function of [GuHCl] at pH 7.0, 0 °C for samples at 34  $\mu$ M peptide concentration. Curves were fit as described in the methods and the Appendix. (a) N16A; (b) N16Abu; (c) N16NVal; (d) N16NLeu.

FIGURE 6a

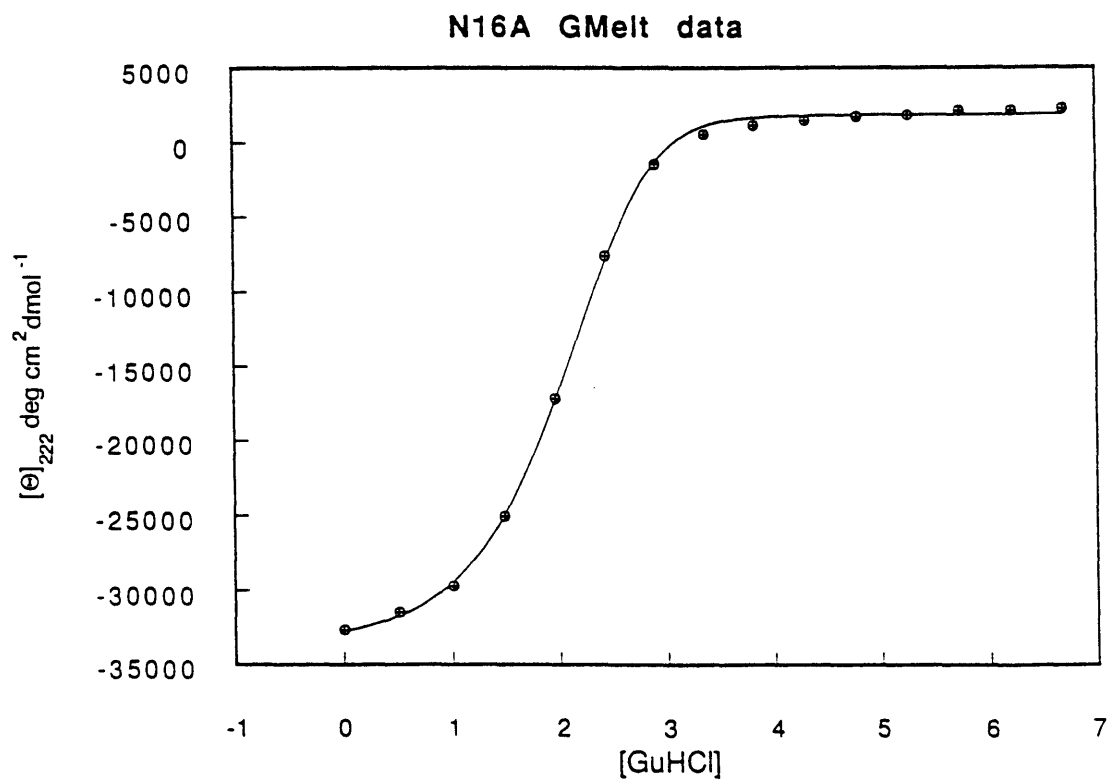


FIGURE 6b

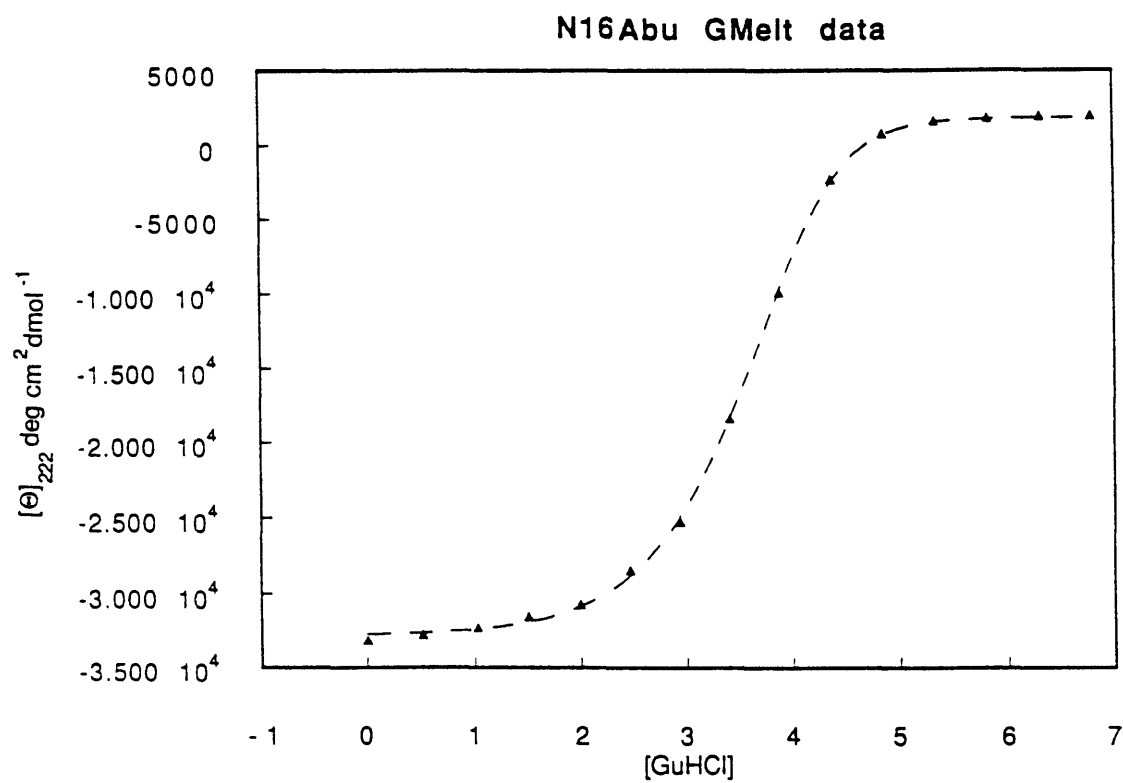




FIGURE 6c

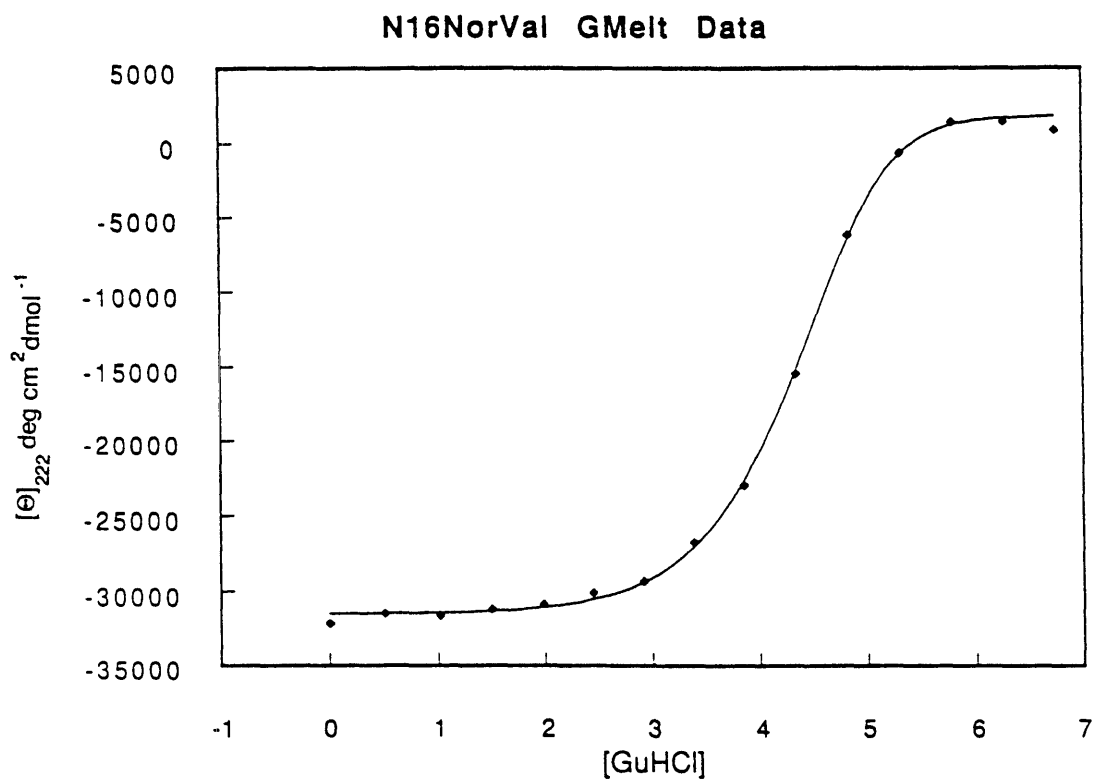
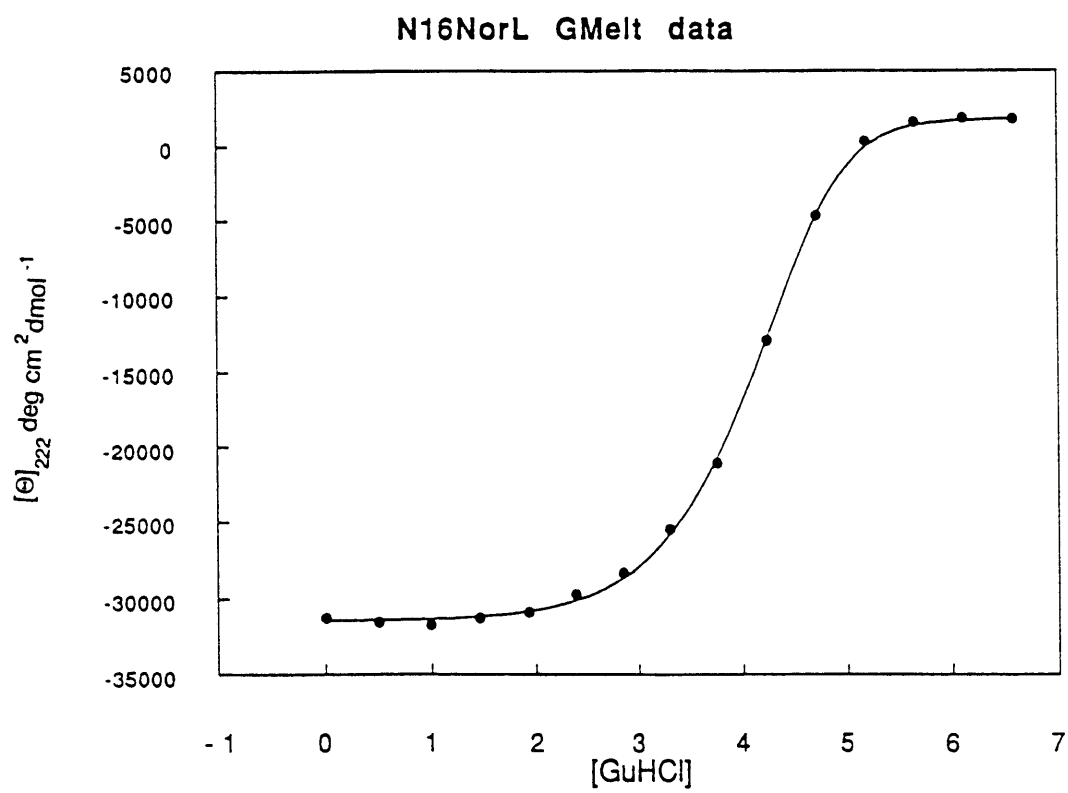
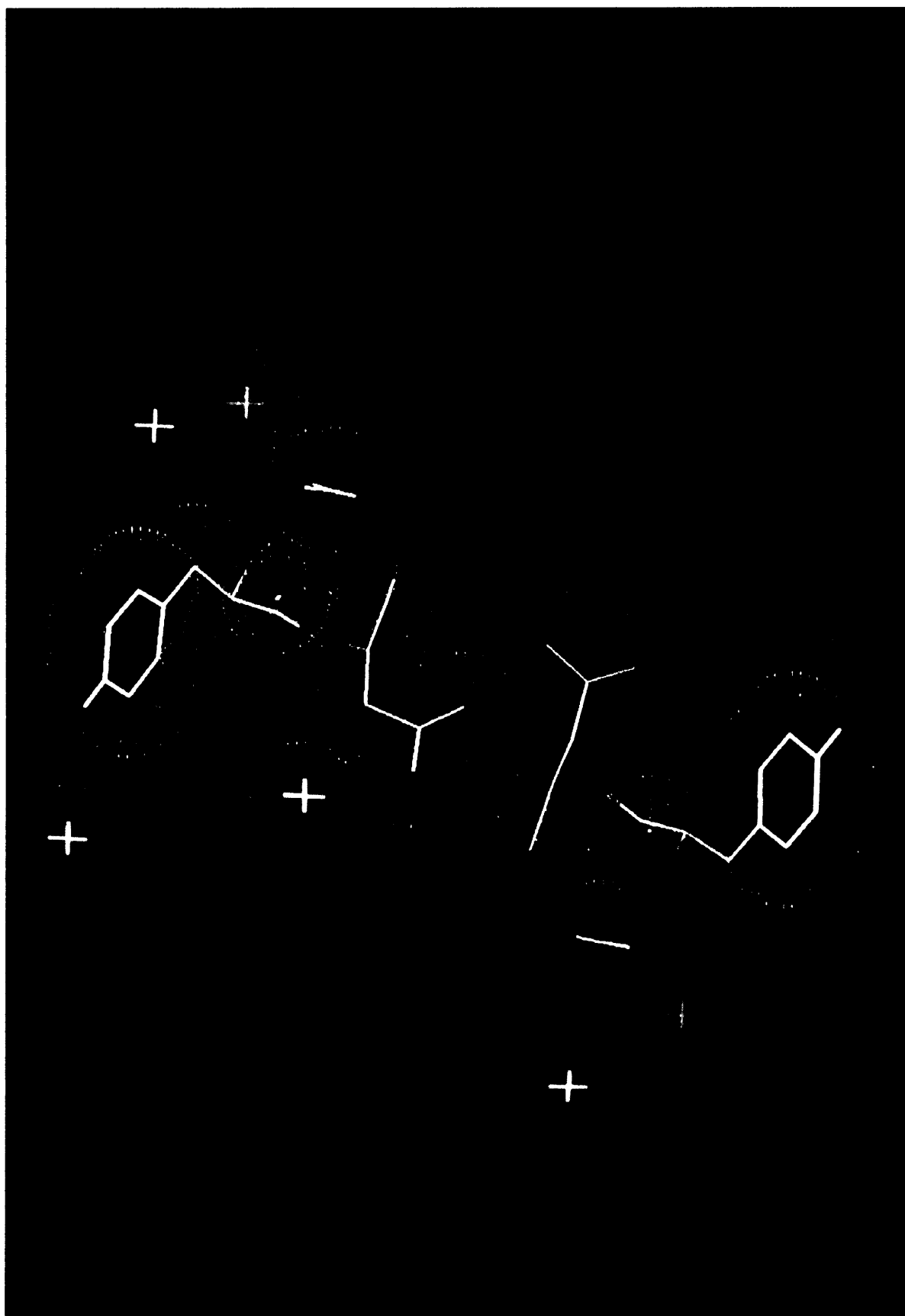


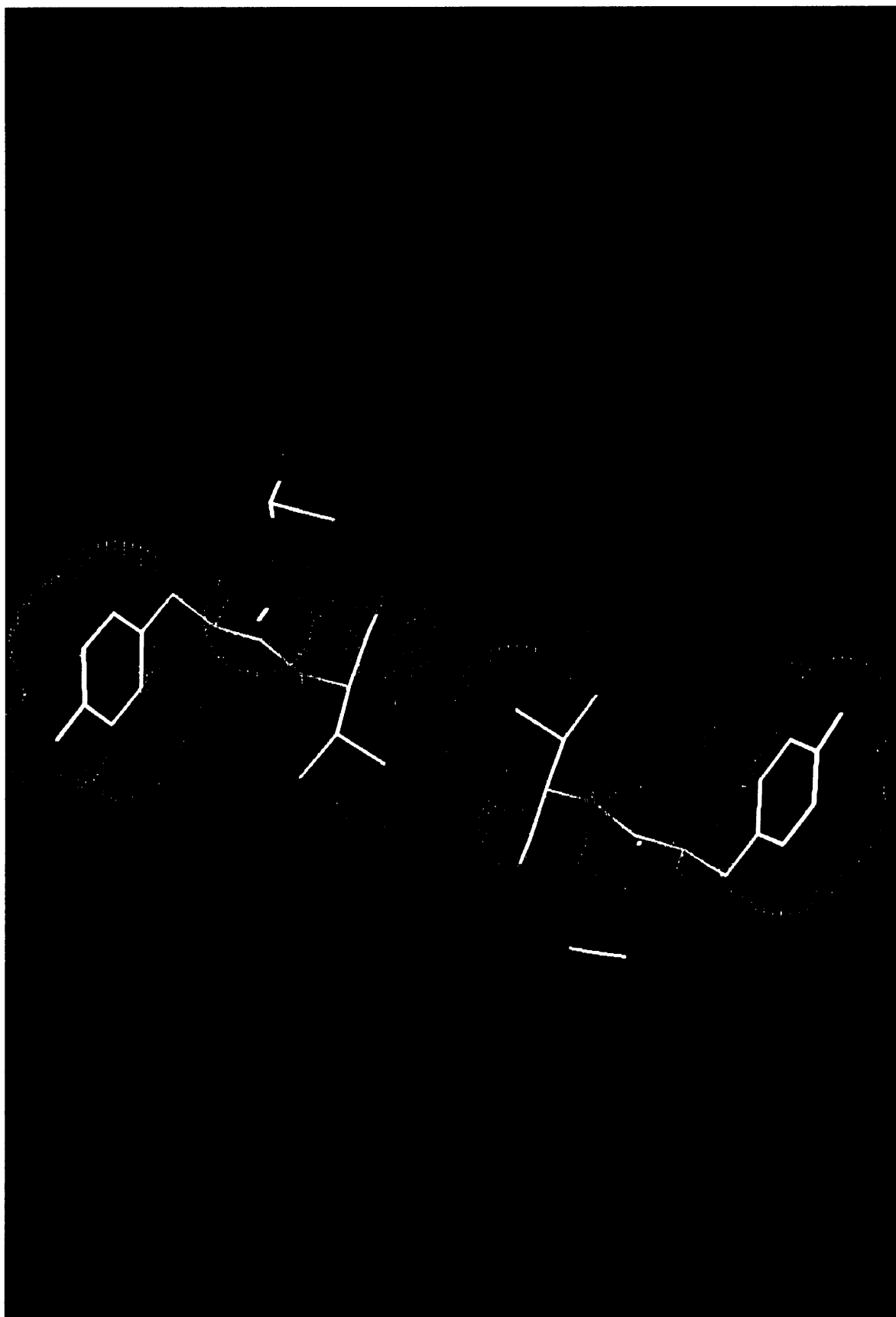
FIGURE 6d

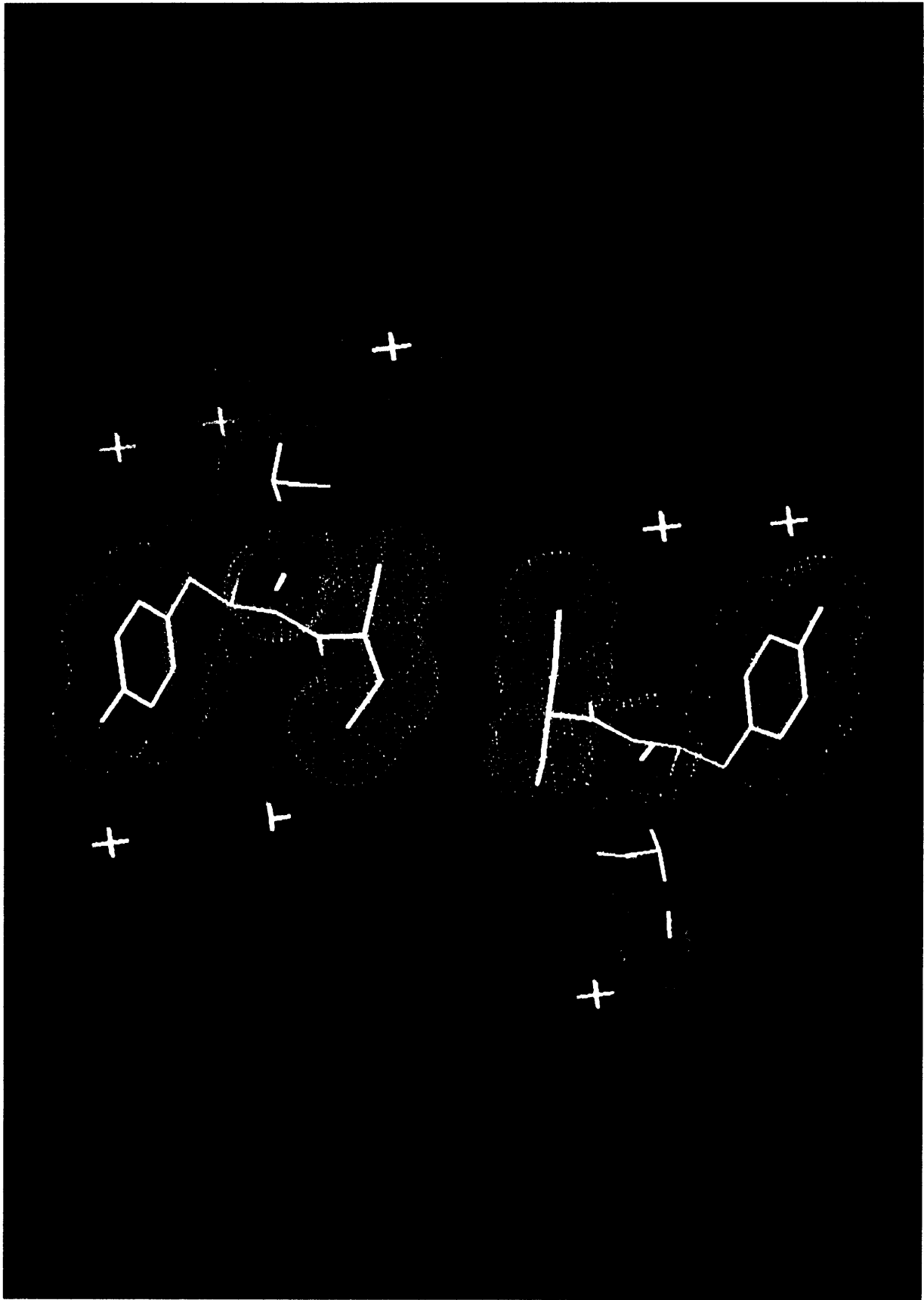


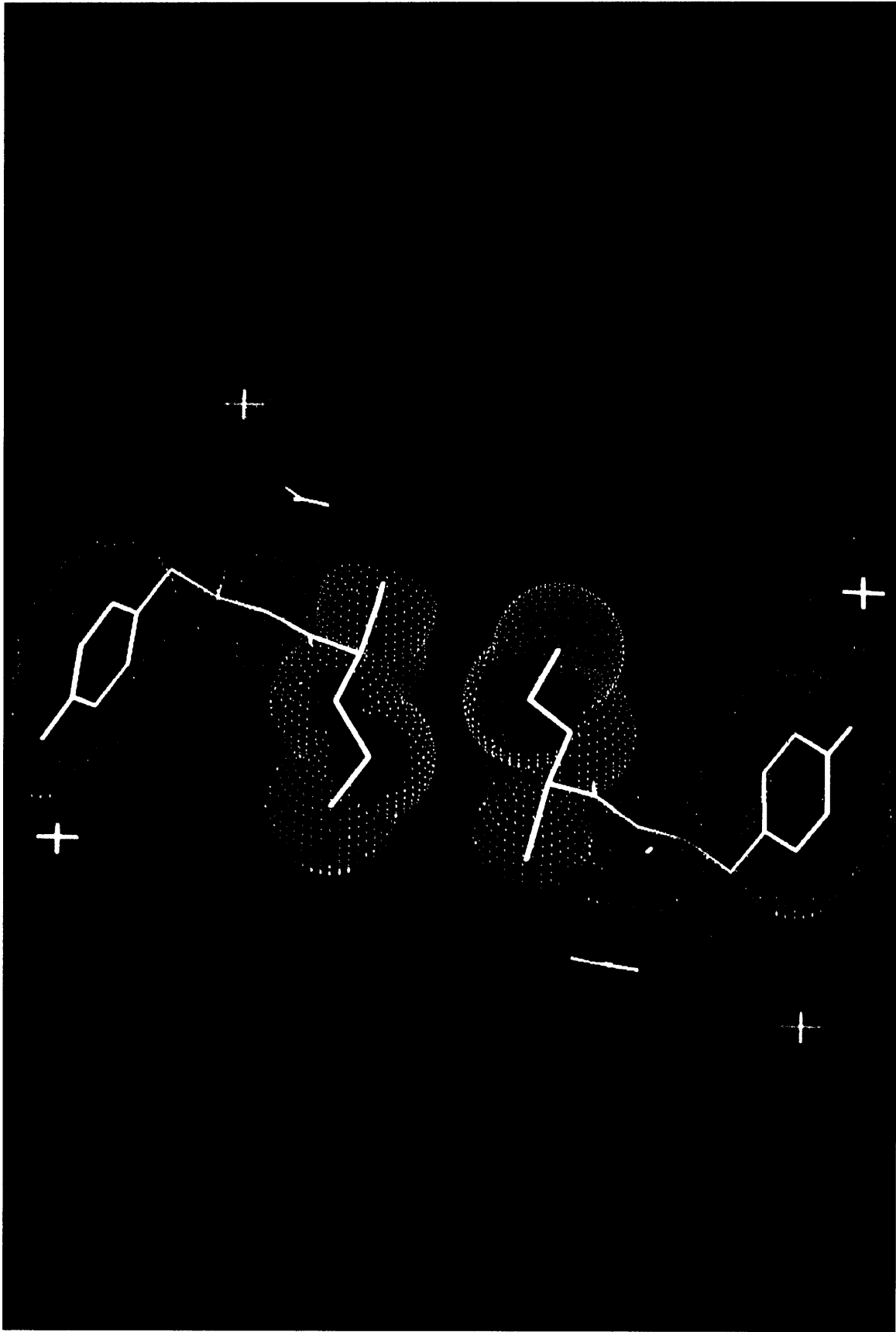
## Figure 7

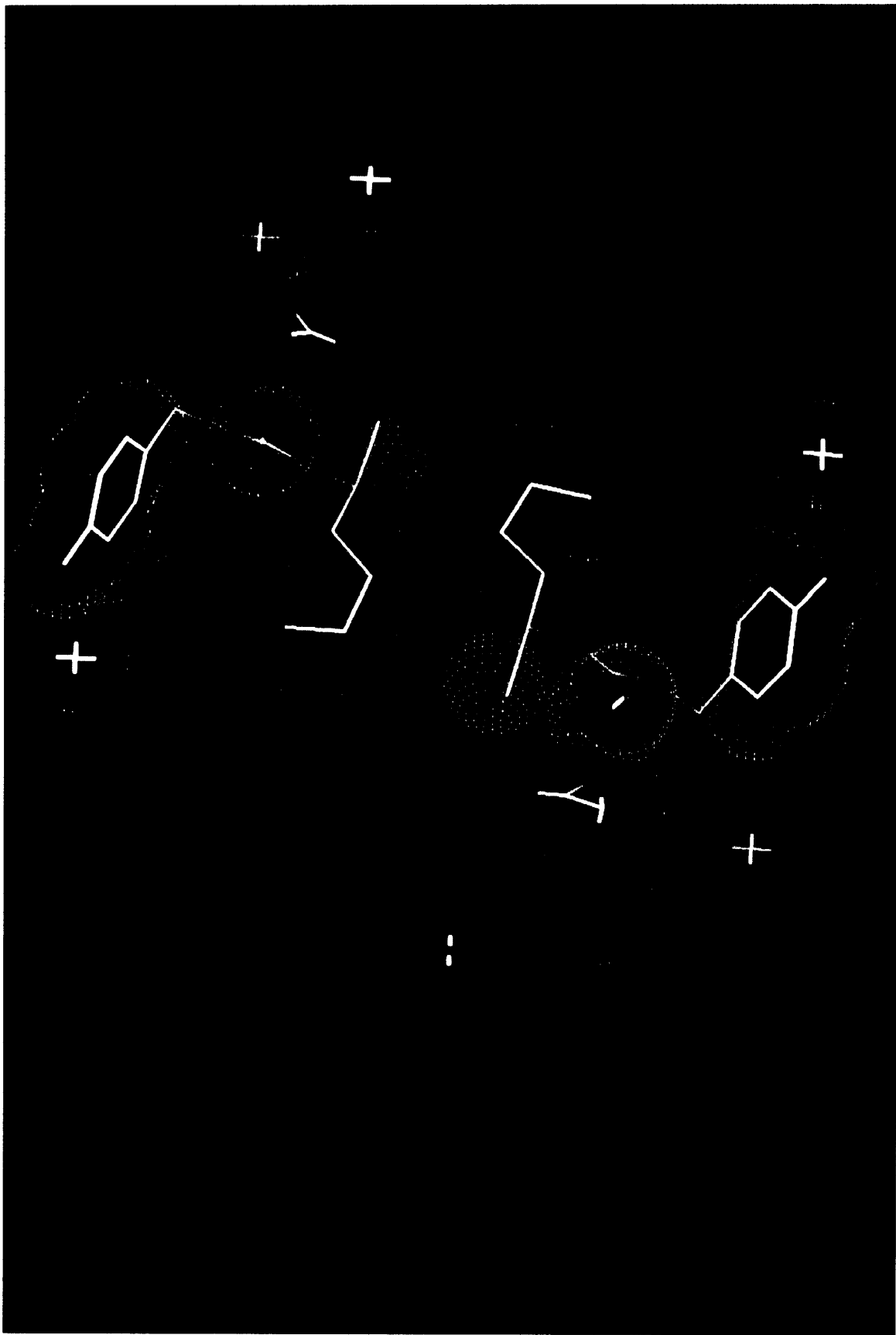
Refined, x-ray crystal structures of each of the GCN4 peptides. One level (residues Lys 15, Asn (or substitution) 16 and Tyr 17) of each coiled-coil dimer is shown. The view is looking down the superhelical axis. In each picture, position 16 is highlighted by a different color, while the other residues remain the same color. (a) N16 (wild-type; pink); (b) N16V (orange); (c) N16Abu (yellow); (d) N16NVal (white); (e) N16NLeu (green).













## Figure 8

Oligomerization state of GCN4 peptides assayed by gel filtration. Peptides were chromatographed at pH 7.0, and 0 °C (or 20 °C; f). Each trace represents a separate chromatography event, and multiple traces are plotted on each figure. The standards for oligomerization order were as follows (Harbury et al., 1993): dimer ("2") = pIL; trimer ("3") = pII; and tetramer ("4") = pLI.

(a) Standards at 1 M GuHCl, pH 7.0, and 0 °C. Note that the trimer and tetramer are not resolved under these conditions.

(b) 34 μM N16Abu plus standards at 1 M GuHCl, pH 7.0, and 0 °C. N16Abu is retained consistent with the dimer conformation.

(c) Standards at 2.5 M GuHCl, pH 7.0, and 0 °C.

(d) 34 μM N16NVal plus standards at 2.5 M GuHCl, pH 7.0, and 0 °C. N16NVal is retained consistent with the dimer conformation under these conditions.

(e) 34 μM N16NLeu plus standards at 2.5 M GuHCl, pH 7.0, and 0 °C. N16NLeu is retained with the dimer standard under these conditions.

(f) N16A, N16Abu, N16NVal and N16NLeu plus the dimer and tetramer standards at 0 M GuHCl, pH 7.0 and 20 °C. N16A is retained slightly less than the dimer, while the other peptides migrate as a mixture of dimer and higher-order oligomer species under these conditions (this experiment was performed by Pehr Harbury).

FIGURE 8a

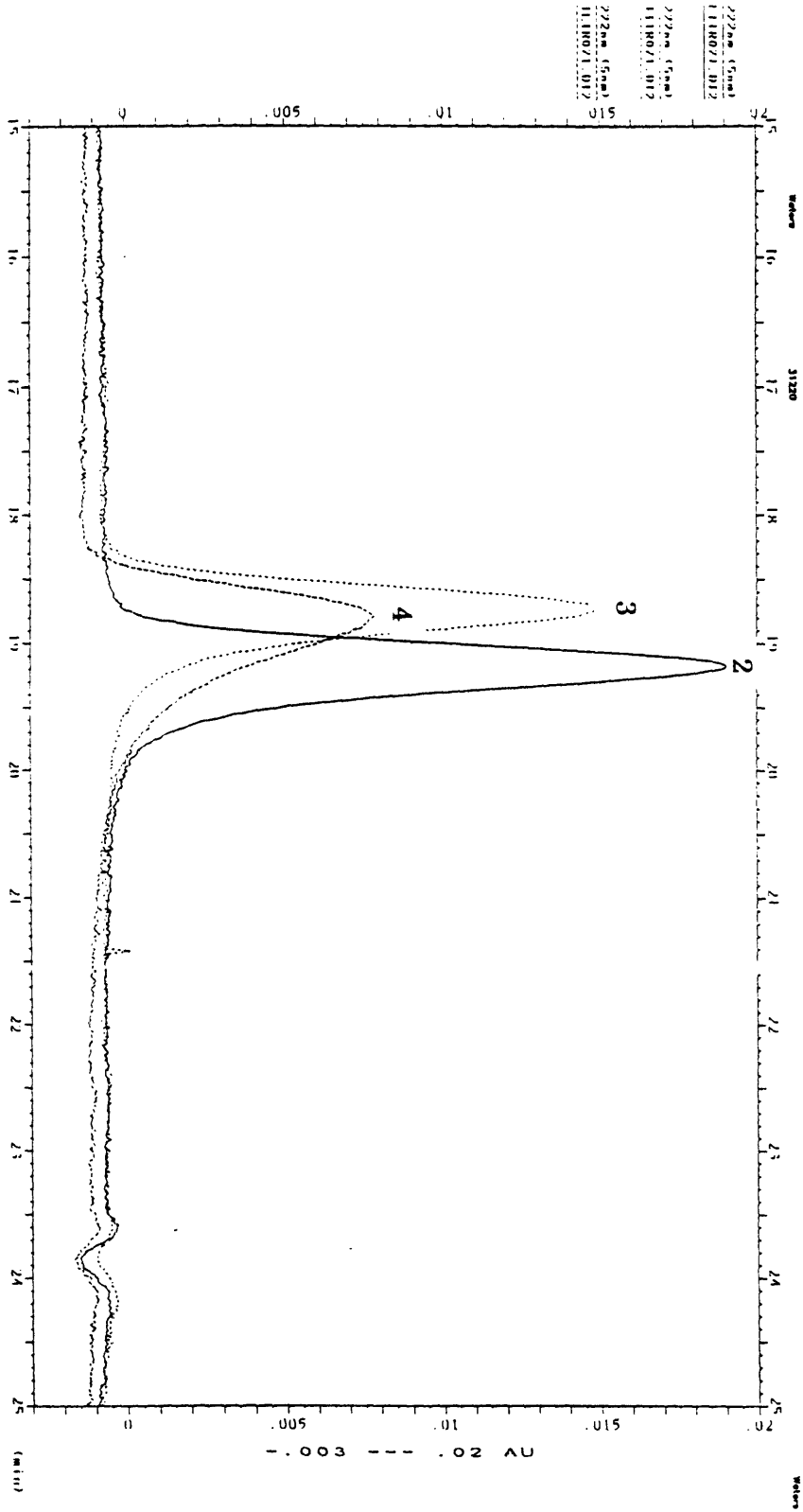


FIGURE 8b

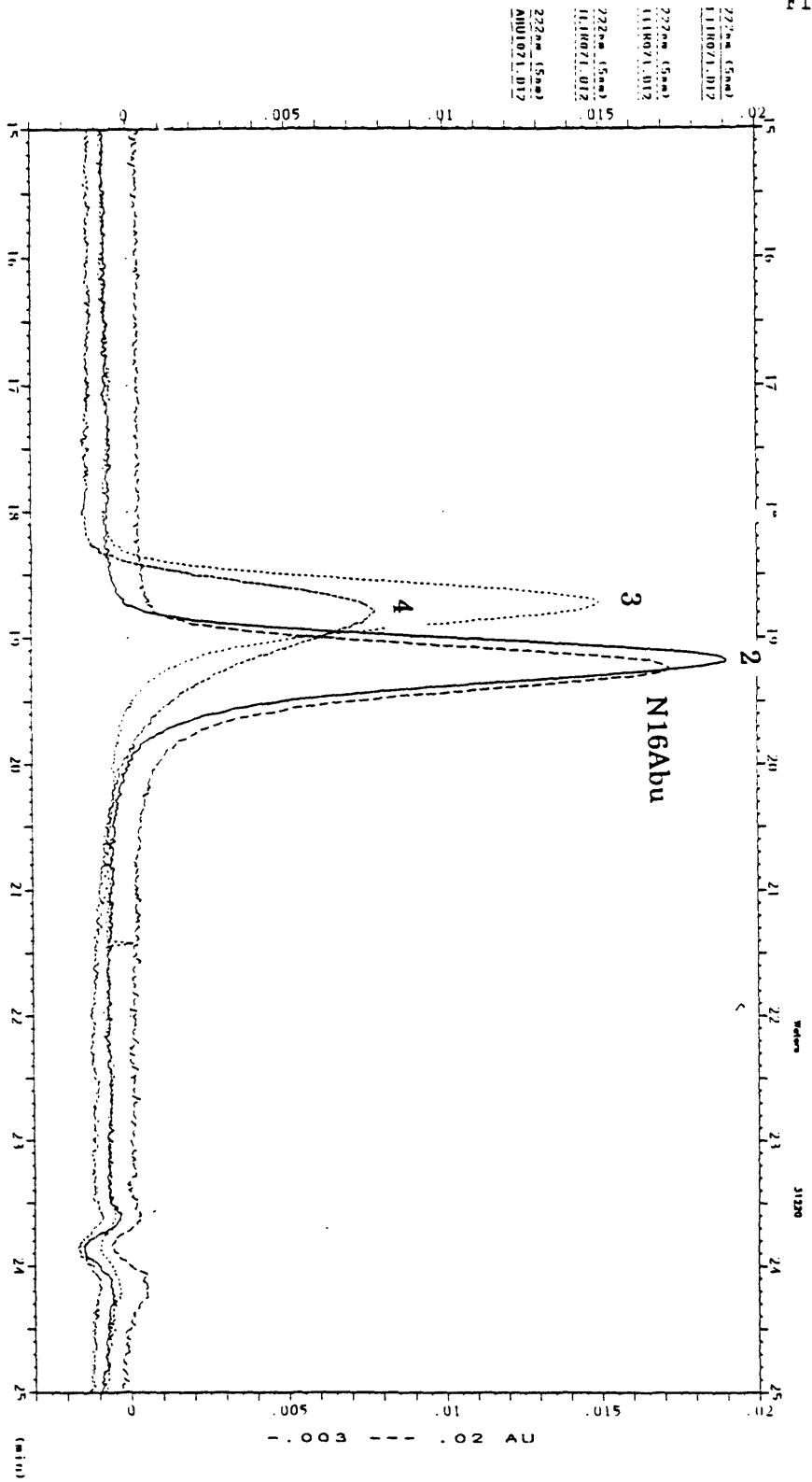


FIGURE 8c

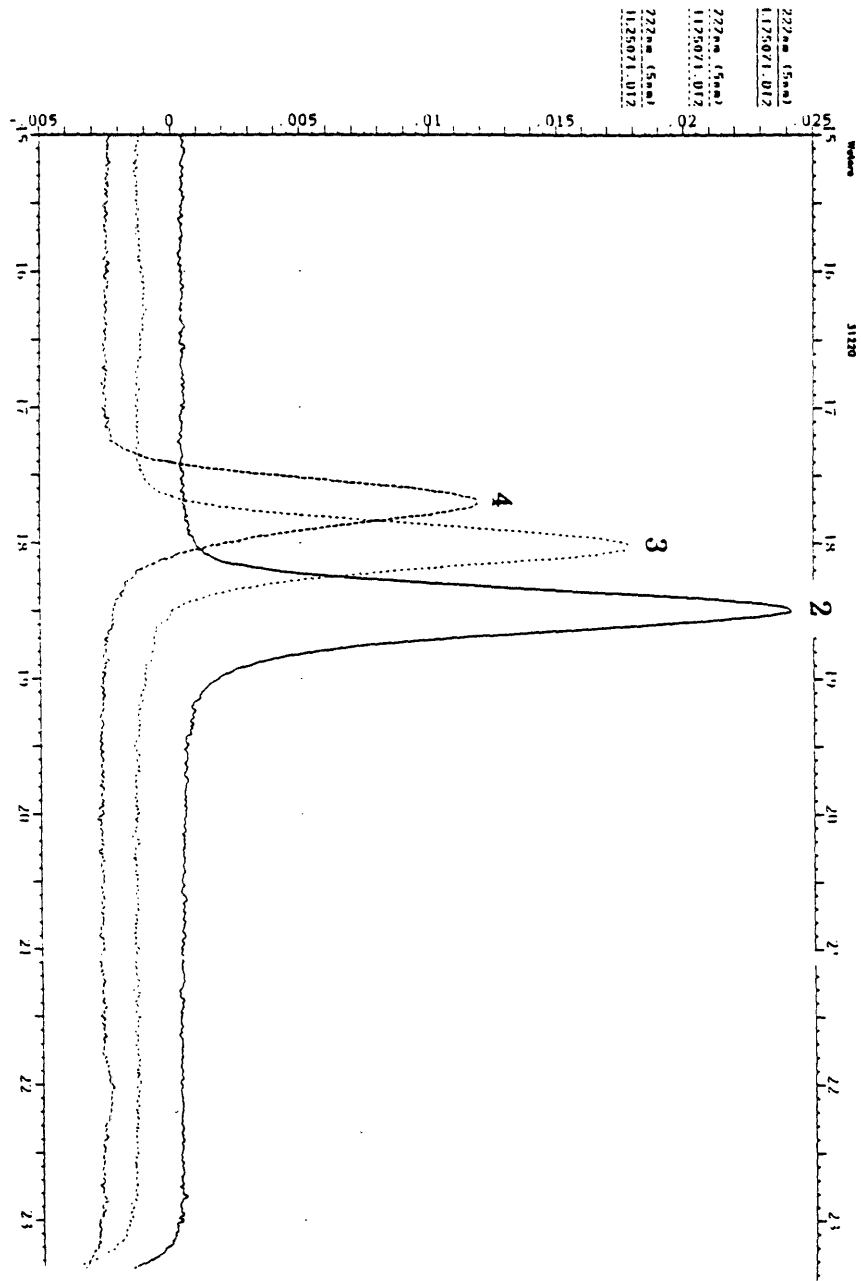


FIGURE 8d

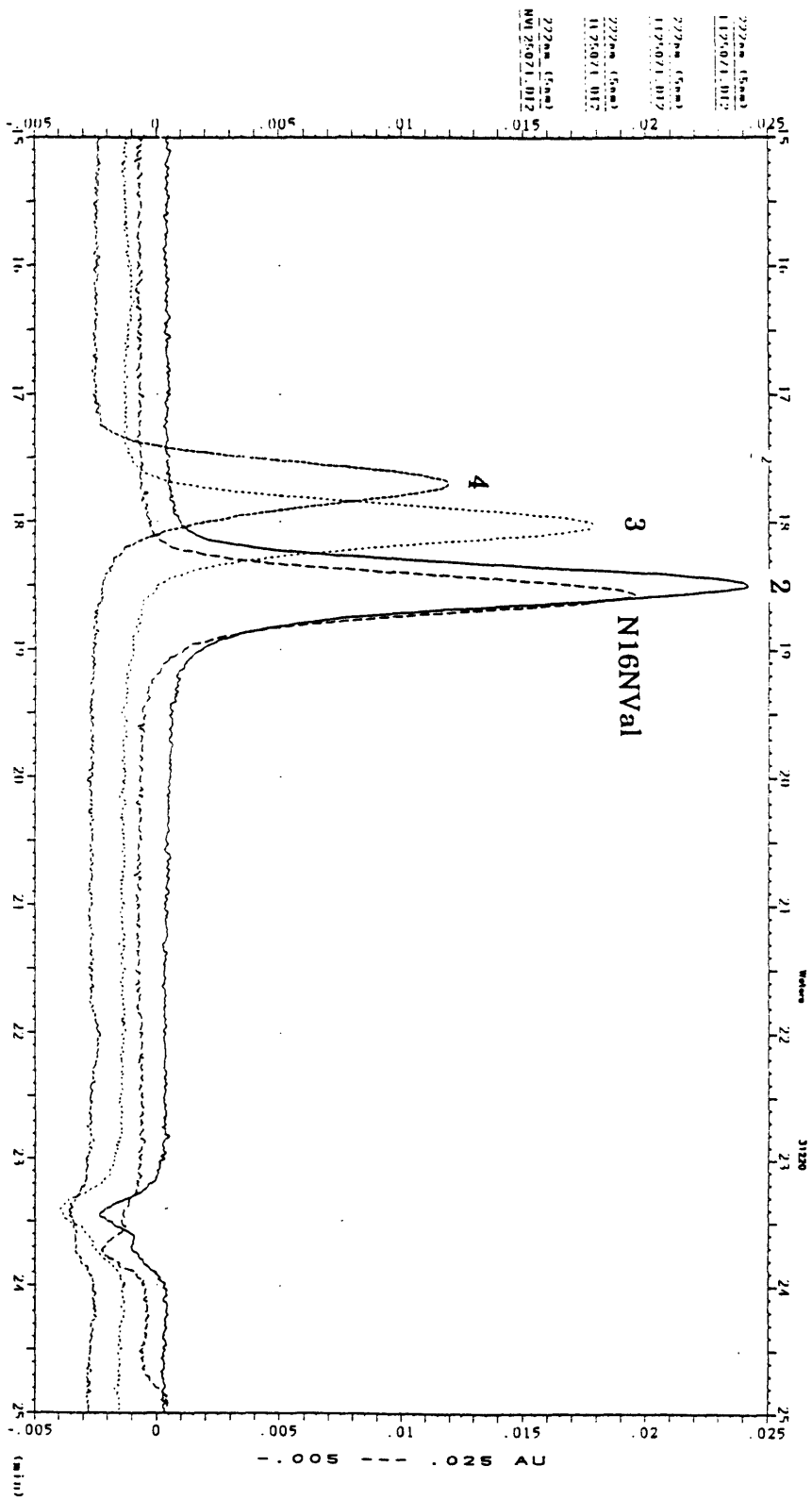


FIGURE 8e

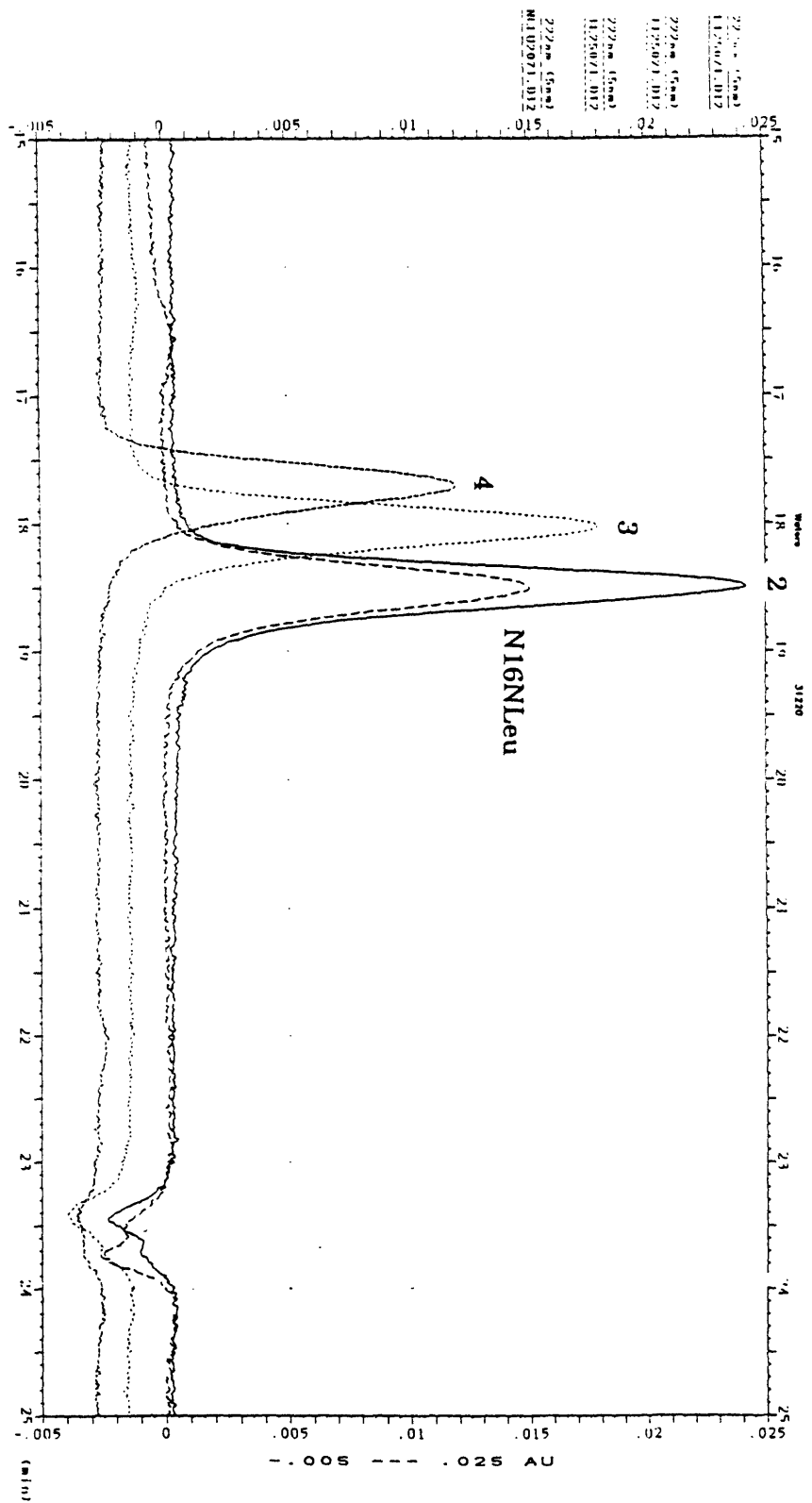
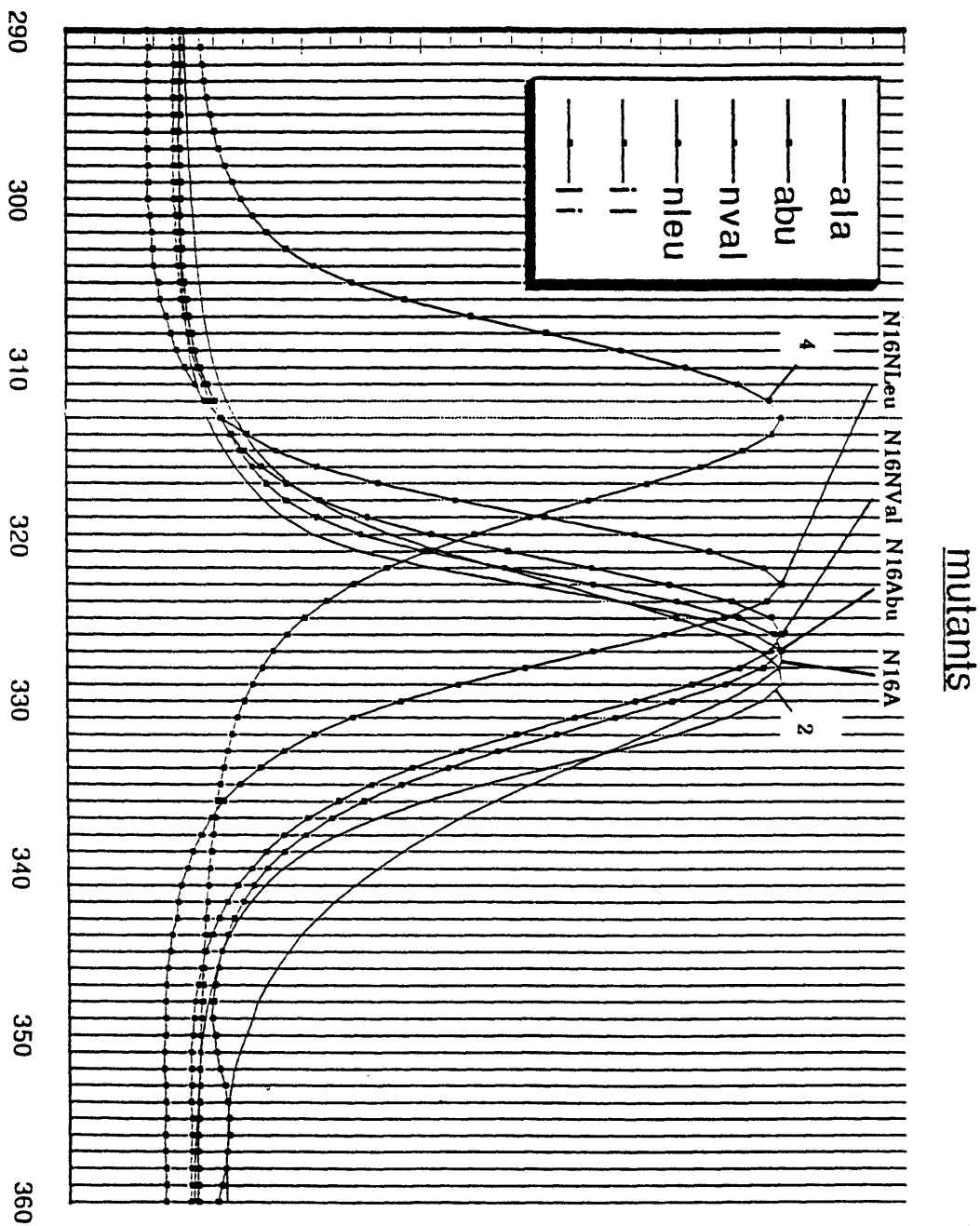


FIGURE 8f



## Figure 9

Correlation between  $\Delta G^\circ$  and the number of methylene groups substituted at position 16, per dimer, in the GCN4 leucine-zipper peptide. (a) A linear correlation is observed for N16A (0 methylene groups), N16Abu (2 methylene groups per dimer, 1 for each monomer) and N16NVal (4 methylene groups per dimer, 2 for each monomer). By contrast, N16NLeu shows no correlation. (b) Ignoring N16NLeu, the contribution of each methylene group can be determined from the slope of a line fit to the data plotted. Each methylene group contributes 1.2 kcal/mol.



FIGURE 9a

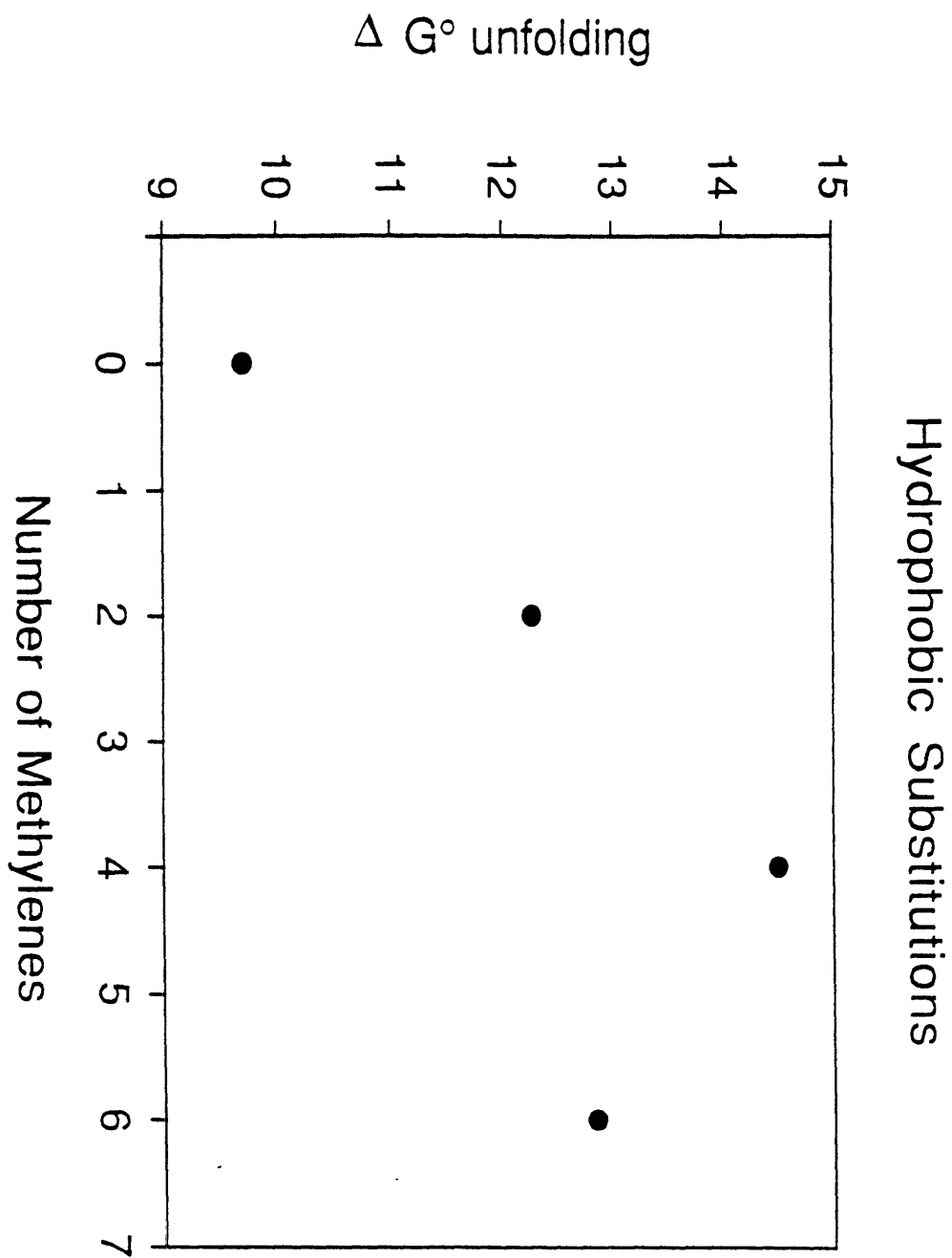
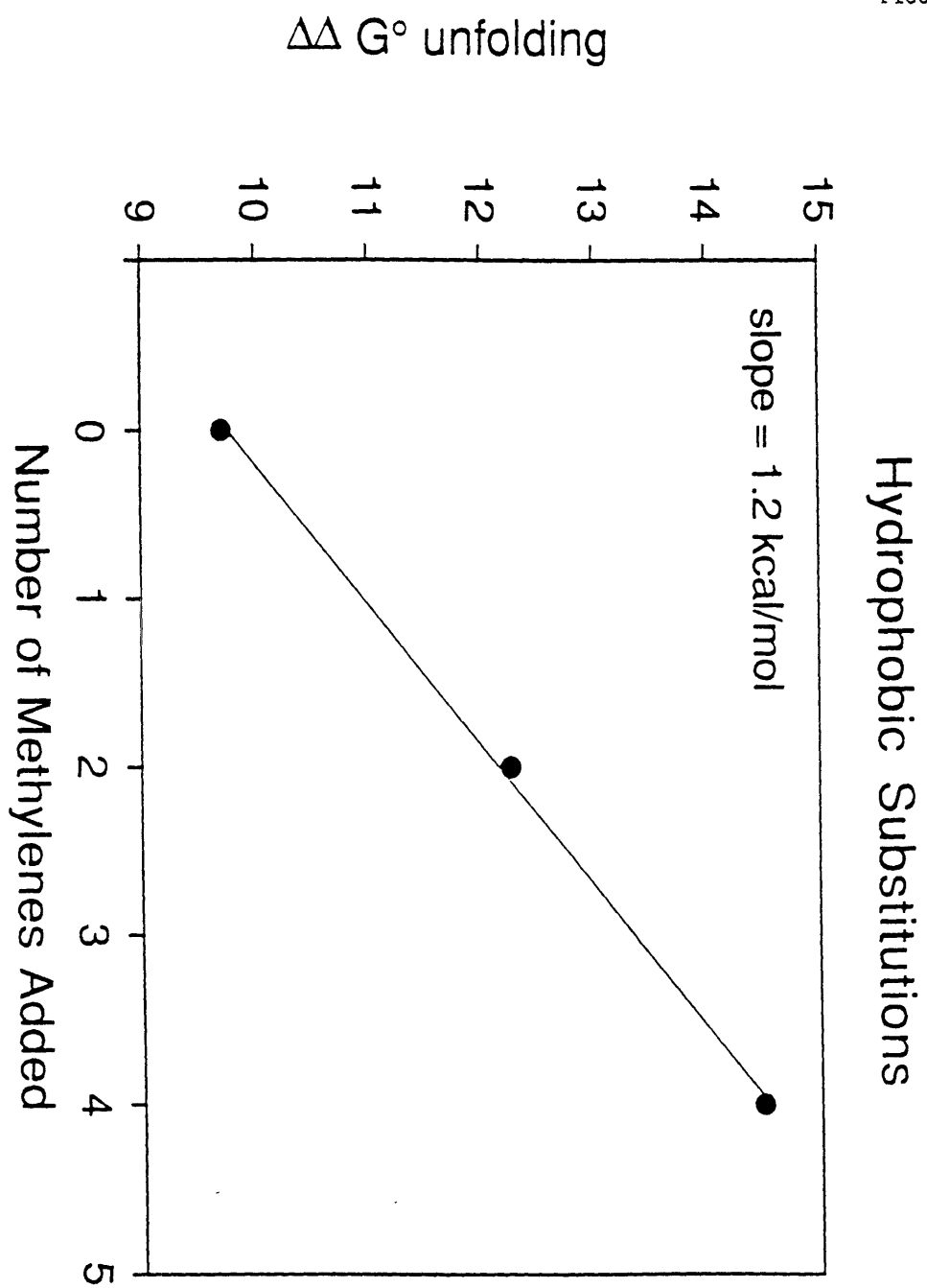


FIGURE 9b



## Appendix II

## Subdomain Folding of the Coiled Coil Leucine Zipper from the bZIP Transcriptional Activator GCN4

One popular model for protein folding, the framework model, postulates initial formation of secondary structure elements, which then assemble into the native conformation. However, short peptides that correspond to secondary structure elements in proteins are often only marginally stable in isolation. A 33-residue peptide (GCN4-p1) corresponding to the GCN4 leucine zipper folds as a parallel, two-stranded coiled coil (O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. A. (1991) *Science* 254, 539-544). Deletion of the first residue (Arg 1) results in local, N-terminal unfolding of the coiled coil, suggesting that a stable subdomain of GCN4-p1 can form. N- and C-terminal deletion studies result in a 23-residue peptide, corresponding to residues 8-30 of GCN4-p1, that folds as a parallel, two-stranded coiled coil with substantial stability (the melting temperature of a 1 mM solution is 43 °C at pH 7). In contrast, a closely related 23-residue peptide (residues 11-33 of GCN4-p1) is predominantly unfolded, even at 0 °C, as observed previously for many isolated peptides of similar length. Thus, specific tertiary packing interactions between two short units of secondary structure can be energetically more important in stabilizing folded structure than secondary structure propensities. These results provide strong support for the notion that stable, cooperatively folded subdomains are the important determinants of protein folding.

### Footnotes

<sup>1</sup> Abbreviations: 1D, one-dimensional; 2D, two-dimensional;  $[\theta]_{222}$ , molar ellipticity at 222 nm; Ac, acetyl; bZIP, basic-region leucine zipper; CD, circular dichroism; HPLC, high performance liquid chromatography; HSMQC, <sup>1</sup>H detected heteronuclear single and multiple quantum correlation spectroscopy; LB, Luria-Bertani; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; GCN4-p1<sub>x-y</sub>, synthetic

peptide corresponding to residues x to y of GCN4-p1; ppm, parts per million;  $T_m$ , midpoint of thermal denaturation.

<sup>2</sup> GCN4-p3 and GCN4-p1 $\Delta$ Ac are less stable at acidic than at neutral pH (the  $T_m$  of 35  $\mu$ M GCN4-p3 decreases from 47 °C at pH 7 to 25 °C at pH 3, and the  $T_m$  of 35  $\mu$ M GCN4-p1 $\Delta$ Ac decreases from 49 °C at pH 7 to 25 °C at pH 3). In contrast, the stability of the corresponding peptides with neutral (acetylated) N-termini are relatively insensitive to pH (the  $T_m$  of 35  $\mu$ M GCN4-p1 is 56 °C at pH 7 and 52 °C at pH 3, and the  $T_m$  of 35  $\mu$ M GCN4-p3Ac is 52 °C at pH 7 and 51 °C at pH 3). The  $pK_a$  of an  $\alpha$ -amino group lies between 7 and 8 (Creighton, 1993), so the N-terminus will be more charged at acidic than at neutral pH. This might result in a unfavorable interaction either between the two like-charged termini or between the charged N-terminus and the helix dipole (cf. Shoemaker et al., 1987; see, however, Greenfield et al., 1994). In addition, GCN4-p3 is only 76 % helical at pH 10, although GCN4-p1 is also destabilized at basic pH, such that at pH 10 the  $T_m$  of 35  $\mu$ M GCN4-p3 (47 °C) is comparable to that of 35  $\mu$ M GCN4-p1 (50 °C). This suggests that, in addition to the potential for destabilizing charge effects at low pH, loss of other favorable interactions (e.g., the g-a packing interaction or the Arg 1-Leu 5 mainchain hydrogen bond) may contribute to the relatively low helix content of GCN4-p3.

## INTRODUCTION

One popular model for protein folding, the framework model, postulates initial formation of secondary structure elements, which then assemble into the native conformation (reviewed in Kim & Baldwin, 1982). Isolated peptides that correspond to secondary structure elements, however, are often only marginally stable (e.g., Scholtz & Baldwin, 1992; Dyson & Wright, 1993). This suggests that stabilizing interactions between transiently formed elements of secondary structure are required for proceeding to the native state. To address how elements of secondary structure associate and stabilize each other, we have investigated the folding of a two-stranded coiled coil, which is a simple motif consisting of just two  $\alpha$ -helices (Figure 1A).

Coiled coils are formed from two, three or four amphipathic  $\alpha$ -helices that wrap around one another in a left handed supercoil (Crick, 1953; Pauling & Corey, 1953; O'Shea et al., 1991; Harbury et al., 1993; see also Banner et al., 1987; Lovejoy et al., 1993; Yan et al., 1993). This structural motif is found in many proteins (Cohen & Parry, 1990), including the leucine zipper or bZIP<sup>1</sup> class of transcription factors (Landschulz et al., 1988; O'Shea et al., 1989; for recent reviews, see Alber, 1992; Hu & Sauer, 1992; Ellenberger, 1994). The sequences of coiled coils are characterized by a heptad repeat of seven amino acid residues, denoted a to g, with a 4, 3 repeat of predominantly hydrophobic residues at positions a and d (Hodges et al., 1972; McLachlan & Stewart, 1975). The intercoil hydrophobic interface (Figure 1A) is formed by residues at positions a, d, e and g (O'Shea et al., 1991; Harbury et al., 1993).

Studies of the folding of coiled coils have focused predominantly on tropomyosin, a 400 Å long, two-stranded coiled coil. Tropomyosin unfolds thermally with a shallow pretransition, involving loss of about 20 % helicity, followed by a major unfolding step (Lehrer, 1978). The nature of the pretransition is unclear, having been attributed to non-cooperative helix fraying (Holtzer & Holtzer, 1992), unfolding in the vicinity of residues 130-190 (Ishii et al., 1992) and loss of coiled coil but not helical structure (Greenfield & Hitchcock-DeGregori, 1993).

Shorter coiled coils may exhibit simpler folding behavior than tropomyosin. A synthetic peptide, called GCN4-p1, corresponding to the leucine zipper of the yeast bZIP transcription factor GCN4, folds independently of the remainder of the GCN4 sequence as a parallel, two stranded coiled coil (O'Shea et al., 1989, 1991; Oas et al.,

1990). We define here peptides derived from the GCN4 coiled coil that form stable, folded structures. The results emphasize the importance of tertiary interactions in stabilizing marginally stable elements of structure during folding, and provide strong support for the notion that stable, cooperatively folded subdomains are the important determinants of protein folding.

## RESULTS AND DISCUSSION

GCN4-p3 (Figure 1B) corresponds to residues 2 to 33 of GCN4-p1 (Table 1) and has an unblocked N-terminus. It differs slightly from GCN4-p1 (O'Shea et al., 1989) which has a neutral (acetylated) N-terminus and begins at Arg 1, rather than Met 2. The X-ray structure of GCN4-p1 is helical for residues Arg 1 to Val 30, and is undefined for residues 31 to 33 (O'Shea et al., 1991). NMR studies indicate that GCN4-p1 in solution is helical from Arg 1 to at least Glu 32 (chemical shift degeneracy prevented the observation of NOEs to Arg 33; Oas et al., 1990). GCN4-p3 also differs slightly from GCN4-p2N (Goodman & Kim, 1991), which has the same sequence but with a neutral (amidated) C-terminus and a N-terminal Ac-Cys-Gly-Gly to allow interchain disulfide formation. N-terminally disulfide-bonded GCN4-p2N is helical as judged by NMR from Met 2 to Arg 33 and has a  $T_m$  of 83 °C (Goodman & Kim, 1991).

CD indicates that GCN4-p3 exists in a helical conformation (Figure 2A) which sedimentation equilibrium indicates is a dimer (Table 1). However, GCN4-p3 is only approximately two-thirds helical at 0 °C and pH 7.0 (Figure 2A and Table 1). In contrast, both GCN4-p1 (Figure 2A and Table 1; O'Shea et al., 1989) and the N-terminally disulfide-bonded GCN4-p2N (Goodman & Kim, 1991) are essentially 100 % helical under identical conditions.

Despite incomplete helix formation, GCN4-p3 exhibits a “folded” baseline during thermal unfolding i.e., prior to the major unfolding step, there is a region where little helical signal is lost with increasing temperature (Figure 2B). This suggests that the CD spectrum of GCN4-p3 at low temperatures does not result from a weighted average of completely folded and completely unfolded peptide molecules. Instead, the data suggest that GCN4-p3 contains both an unfolded region and a separate folded region that is stable to thermal denaturation.

The helices of the GCN4 coiled coil are parallel (O’Shea et al., 1989, 1991). To confirm that the helices of GCN4-p3 are also parallel, the concentration dependence of the  $T_m$  of a disulfide bonded variant of GCN4-p3 was studied. GCN4-p3C has the same sequence as GCN4-p3, except for a C-terminal Gly-Gly-Cys that allows formation of an interchain disulfide bond. If the two helices are disulfide bonded in the favorable orientation the  $T_m$  will be independent of concentration. Conversely, it has been shown that if the GCN4 helices are disulfide-bonded in an unfavorable

orientation (i.e., anti-parallel) then higher order oligomers form, reflected in a concentration dependence of  $T_m$  (O’Shea et al., 1989). The  $T_m$  of GCN4-p3C ( $77 \pm 1$  °C at pH 7) is independent of concentration over the range 50  $\mu$ M to 2 mM, confirming that the helices of GCN4-p3 are parallel.

NOEs that are characteristic of  $\alpha$ -helical structure are found from residues 5 to 29 in the NOESY spectrum of GCN4-p3 (Figure 2), and from residues 5 to 31 in the HSMQC-NOESY spectrum of disulfide-bonded GCN4-p3C (Figure 3), but not from residues 2 to 5. We conclude that GCN4-p3 is folded as a parallel, two-stranded coiled coil from Leu 5 to at least Leu 29, and that the N-terminal residues Met 2, Lys 3 and Gln 4 are predominantly unfolded.

These results suggest that N-terminal deletions of residues from the GCN4 coiled coil sequence may be possible without loss of the coiled coil structure. Truncations were made at residues one before the hydrophobic interface residues at positions a and d (i.e., the first residue in the peptide was either at



position **g** or **c**; Figure 1), since GCN4-p1ΔAc (which begins at Arg 1 and has a free N-terminus) is significantly more helical than GCN4-p3 (Table 1). It is likely that the loss of a favorable **g-a** packing interaction (Figure 1A; O'Shea et al., 1991; Greenfield et al., 1994) and a mainchain hydrogen bond (Arg 1-Leu 5) contribute to the lower helix content of GCN4-p3 compared to GCN4-p1ΔAc. The N-termini of the truncated peptides were synthetically neutralized (acetylated), since acetylation of the N-terminus of GCN4-p3 results in a fully folded, two-stranded coiled coil (GCN4-p3Ac in Table 1) <sup>2</sup>. A significant increase in coiled coil stability upon N-terminal acetylation has also been observed for peptides corresponding to the N-terminus of tropomyosin (Greenfield et al., 1994).

Peptides beginning at residues Gln 4 (GCN4-p14-33) and Lys 8 (GCN4-p18-33) fold as stable, two-stranded coiled coils (Table 1 and Figure 4), although the helix content does not approach 100 % by CD until the concentration of the peptide is increased to the mM level (Table 1). The next truncation was made at Glu 11. This 23-residue peptide (GCN4-p111-33) is largely unfolded by CD (Table 1) and NMR (Figure 4). Thus, N-terminal truncations of the GCN4 coiled coil beyond Glu 10 do not form stable structures.

C-terminal truncations were made from the sequence of GCN4-p18-33, which is the shortest of the N-terminal truncations that folds as a stable coiled coil. GCN4-p18-30 corresponds to a deletion of residues 31-33 and folds cooperatively as a two-stranded, coiled coil (Figure 5 and Table 1). Disulfide bonded GCN4-p18-30N (a variant of GCN4-p18-30, which has an N-terminal Ac-Cys-Gly-Gly to allow formation of an interchain disulfide bond) has a  $T_m$  ( $72 \pm 1$  °C at pH 7) that is independent of concentration over the range 35 μM to 2 mM, indicating that the helices are parallel (see above). A 20-residue peptide, GCN4-p18-27, corresponding to a further deletion of residues 28 to 30, is not folded by CD (Table 1) or NMR (Figure 4). Thus, the shortest peptide derived here from the GCN4 sequence that folds as a stable, two-stranded parallel coiled coil with a cooperative thermal unfolding transition is the 23-residue peptide GCN4-p18-30. This is the shortest isolated coiled coil to date

(cf. Lau et al., 1984), although shorter coiled coils can exist in the context of an entire domain (Marmorstein et al., 1992).

GCN4-p18-30, along with the longer peptides, folds as a stable, parallel, two-stranded, coiled coil. This indicates that the GCN4 leucine zipper contains stable subdomains (i.e., a cooperatively folded unit of structure larger than an isolated helix, but smaller than an entire domain; Rose, 1979; Richardson, 1981; Oas & Kim, 1988; Staley & Kim, 1990; Wu et al., 1993). It is striking that GCN4-p111-33 is predominantly unfolded, even though it is the same length as GCN4-p18-30 and differs only at three terminal residues. In addition, deletion of the three C-terminal residues from GCN4-p18-30 results in an unfolded peptide. Isolated helices are often only marginally stable (Scholtz & Baldwin, 1992), including those of a peptide corresponding to the bZIP region of GCN4 (Thompson et al., 1993). Our results indicate, therefore, that the formation of specific tertiary interactions give rise to the cooperative folding of GCN4-p18-30. Similar conclusions have been reached from studies of peptide fragments of BPTI, cytochrome *c* and myoglobin (Oas & Kim, 1988; Staley & Kim, 1990; Wu et al., 1993; Shin et al., 1993).

The framework model for protein folding emphasizes the formation of secondary structure (Kim & Baldwin, 1982). Although the intrinsic propensities of the amino acids to form  $\alpha$ -helices (for reviews, see Scholtz & Baldwin, 1992; Fersht & Serrano, 1993) or  $\beta$ -sheets (Kim & Berg, 1993; Minor & Kim, 1994; Smith et al., 1994) appear to be important determinants of protein folding, our results suggest that tertiary

packing interactions (even between two short  $\alpha$ -helices) can be energetically more important in establishing a cooperatively folded structure. This conclusion strongly reinforces the notion that protein folding can be understood, in large part, in terms of the formation of cooperatively folded subdomains, in which elements of secondary structure are stabilized by native-like tertiary interactions.

## ACKNOWLEDGEMENTS

We are especially indebted to L. P. McIntosh for many valuable discussions in all aspects of this work. We thank M. W. Burgess, S. Britt, M. A. Milhollen and R. Rutkowski for peptide synthesis and mass spectrometry and C. J. McKnight, B. A. Schulman and J. P. Staley for many helpful discussions. P.S.K. is a Pew Scholar in the Biomedical Sciences.

## METHODS

The recombinant peptides GCN4-p3 and GCN4-p3C were expressed in *Escherichia coli* strain BL21 (DE3) pLysS using the T7 system (Studier et al., 1990). A synthetic gene encoding GCN4-p3 was prepared with optimal codon usage for *Escherichia coli* (Grosjean & Fiers, 1982) and ligated into the *Xba*I/*Eco*R1 site of pET3a (Studier et al., 1990) as a precursor to an expression plasmid (pJH370) for a  $\lambda$  repressor-GCN4 leucine zipper fusion protein (Hu et al., 1990). The gene was then subcloned into the *Xba*I/*Eco*R1 site of the high copy expression vector pAED4 (Doering, 1992) using standard procedures (Sambrook et al., 1989) and called p4LZ. The plasmid for GCN4-p3C (called p4LZGGC) was prepared by oligonucleotide-directed mutagenesis of p4LZ (Zoller & Smith, 1982). Sequences were confirmed by DNA sequencing (Sanger et al., 1977).

Cells were grown directly from a single colony in LB media containing 100 mg/l ampicillin and 25 mg/l chloramphenicol to an optical density of 0.6 at 600 nm and induced with isopropyl- $\beta$ -D-thiogalactopyranoside (final concentration 0.5 mM). After 3 hours, cells were harvested by centrifugation and lysed by freezing followed by sonication in 50 mM Tris-base, 1 mM EDTA, pH 8.0. The pH of the lysate was reduced to 2 with HCl to precipitate impurities and the resulting mixture centrifuged. The soluble fraction was dialyzed against 10 mM sodium acetate, pH 4.0 and then loaded onto CM-Sephacrose CL-6B. Contaminants were eluted in 10 mM sodium acetate, pH 4.0 and the GCN4 peptide was then eluted in the same buffer containing 1 M NaCl. Final purification was by reverse-phase HPLC on a Vydac C18 column using a linear water/acetonitrile gradient containing 0.1 % trifluoroacetic acid. The yield of GCN4-p3 and GCN4-p3C from LB media was 60-90 and 40-60 mg/l respectively. Uniformly  $^{15}\text{N}$  labeled GCN4-p3C was prepared in the same way except cells were grown in M9T minimal media (McIntosh et al., 1987) containing 1 g/l  $(^{15}\text{NH}_4)_2\text{SO}_4$  to yield 25 mg/l HPLC purified  $^{15}\text{N}$  labeled GCN4-p3C. All other peptides were synthesized using Fmoc or t-Boc solid phase synthesis and purified as described previously (O'Shea et al., 1989, 1993). The identity of each peptide was confirmed by laser desorption mass spectrometry on a Finnigan LASERMAT and in all cases the expected and

observed molecular weights agreed to within 2 Da. The numbering of residues of the peptides follows the GCN4-p1 sequence (O'Shea et al., 1989).

CD spectra were acquired on Aviv 60DS or 62DS spectrometers. Samples were prepared in 50 mM sodium phosphate, 150 mM NaCl. Peptide concentrations were determined by tyrosine absorbance in 6 M guanidine hydrochloride assuming an extinction coefficient at 276 nm of  $1500 \text{ M}^{-1} \text{ cm}^{-1}$  (Edelhoch, 1967). Helix content was estimated from  $[\theta]_{222}$  by assuming that a value of  $-33 \times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$  corresponds to a helix content of 100 % for a 33-residue coiled coil (O'Shea et al., 1989) and by correcting for the length dependence of  $[\theta]_{222}$  (Chen et al., 1974). Thermal stability was determined at peptide concentrations (monomer) of 35  $\mu\text{M}$  or 1 mM by monitoring the change in  $[\theta]_{222}$  as a function of temperature. The temperature was increased in steps of 1 °C with an equilibration time of 120 s and a data collection time of 30 s. The  $T_m$  was determined from the minima of the first derivative of  $[\theta]_{222}$  with  $T^{-1}$ , where T is in K (Cantor & Schimmel, 1980). All thermal melts were reversible (the folding and unfolding curves were superimposable, with  $\geq 95$  % of the starting signal regained on cooling).

Apparent molecular weights were determined by sedimentation equilibrium on a Beckman XL-A ultracentrifuge at 0 °C. Samples were dialyzed against 50 mM sodium phosphate, 150 mM NaCl, pH 7.0 for at least 12 hours. Typically, three samples of total peptide concentration of 166, 332 and 553  $\mu\text{M}$  were run at rotor speeds of 35 000 or 40 000 rpm and the absorbance at 276 nm was monitored. One peptide (GCN4-p111-33) was not significantly folded at these concentrations; in this case, a concentration of 1 mM was used and the absorbance at 244 nm was monitored. Data were fit to an ideal model plot of  $\ln(\text{absorbance})$  versus radial distance squared. Significant deviation of residuals, indicative of deviation from this ideal model, were not observed except for the partially folded peptide GCN4-p111-33. Partial molar volumes and solvent densities were calculated as described by Laue et al. (1992).

NMR spectroscopy was performed on a Bruker AMX spectrometer operating at 500.1 MHz for  $^1\text{H}$ . Samples were typically 2 to 5 mM peptide in 50 mM sodium phosphate, 150 mM NaCl, pH 6.5 containing 10 %  $\text{D}_2\text{O}$  and internally referenced to zero ppm with trimethylsilylpropionic acid. NOESY (Macura et al., 1981) and 2D  $^1\text{H}$ - $^{15}\text{N}$  HSMQC-NOESY (Gronenborn et al., 1989; Zuiderweg, 1990) data sets consisted of 512  $t_1$  increments defined by 1024 complex points and 96 transients collected using time proportional phase incrementation (Marion & Wüthrich, 1983). Data were

acquired at 5 °C using a spectral width of 6024.1 Hz in the  $^1\text{H}$  dimension and 1250 Hz in the  $^{15}\text{N}$  dimension and a recycle delay of 1.5 s. The NOESY mixing time was 150 ms.  $^1\text{H}$  were decoupled in heteronuclear experiments using WALTZ-16 (Shaka et al., 1983). The water resonance was suppressed using a selective 1-1 echo sequence to avoid saturation transfer to the exchangeable amide NH protons (Sklenar & Bax, 1987). Data were resolution enhanced using a Gaussian function in  $t_2$  and a 50° shifted squared sine bell in  $t_1$  and zero filled prior to Fourier transformation to give a final digital resolution of 2.9 and 1.2 Hz/point in the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions respectively. Spectra were assigned using a main chain directed approach (Englander & Wand, 1987) using  $d_{\text{NN}}(i, i+1)$  NOEs and the Tyr 17 spin system as a unique reference point.

## REFERENCES

- Alber, T. (1992) *Curr. Opin. Genet. Dev.* 2, 205-210.
- Banner, D. W., Kokkinidis, M., & Tsernoglou, D. (1987) *J. Mol. Biol.* 196, 657-675.
- Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, New York.
- Chen, Y., Yang, J. T., & Chau, K. H. (1974) *Biochemistry* 13, 3350-3359.
- Creighton, T. E. (1993) *Proteins*, Freeman, New York.
- Cohen, C., & Parry, D. A. D. (1990) *Proteins* 7, 1-15.
- Crick, F. H. C. (1953) *Acta. Cryst.* 6, 689-697.
- Doering, D. S. (1992) *Ph. D. Thesis*, Massachusetts Institute of Technology.
- Dyson, H. J., & Wright, P. E. (1993) *Curr. Opin. Struct. Biol.* 3, 60-65.
- Edelhoch, H. (1967) *Biochemistry* 6, 1948-1954.
- Ellenberger, T. (1994) *Curr. Opin. Struct. Biol.* 4, 12-21.
- Englander, S. W., & Wand, A. J. (1987) *Biochemistry* 26, 5953-5958.
- Fersht, A. R., & Serrano, L. (1993) *Curr. Opin. Struct. Biol.* 3, 75-83.
- Goodman, E. M., & Kim, P. S. (1991) *Biochemistry* 30, 11615-11620.
- Greenfield, N. J., & Hitchcock-DeGregori, S. E. (1993) *Protein Sci.* 2, 1263-1273.
- Greenfield, N. J., Stafford, W. F., & Hitchcock-DeGregori, S. E. (1994) *Protein Sci.*, *in press*.
- Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett.* 243, 93-98.
- Grosjean, H., & Fiers, W. (1982) *Gene* 18, 199-209.
- Harbury, P. B., Zhang, T., Kim, P. S., & Alber, T. (1993) *Science* 262, 1401-1407.
- Hodges, R. S., Sodek, J., Smillie, L. B., & Jurasek, L. (1972) *Cold Spring Harbor Symp. Quant. Biol.* 37, 299-310.
- Holtzer, M. E., & Holtzer, A. (1992) *Biopolymers* 32, 1589-1591.
- Hu, J. C., O'Shea, E. K., Kim, P. S., & Sauer, R. T. (1990) *Science* 250, 1400-1403.
- Hu, J. C., & Sauer, R. T. (1992) *Nucleic Acids Mol. Biol.* 6, 82-101.
- Ishii, Y., Hitchcock-DeGregori, S., Mabuchi, K., & Lehrer, S. S. (1992) *Protein Sci.* 1, 1319-1325.
- Kim, C. A., & Berg, J. M. (1993) *Nature* 362, 267-70.
- Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem.* 51, 459-489.
- Landschulz, W. H., Johnson, P. F., & McKnight, S. L. (1988) *Science* 240, 1759-1764.
- Lau, S. Y. M., Taneja, A. K., & Hodges, R. S. (1984) *J. Biol. Chem.* 259, 13253-13261.

- Laue, T. M., Shah, B. D., Ridgeway, T. M., & Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S.E., Rowe, A.J., & Horton, J.C., Eds.) pp 90-125, The Royal Society of Chemistry, Cambridge.
- Lehrer, S. S. (1978) *J. Mol. Biol.* 118, 209-226.
- Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., & Eisenberg, D. (1993) *Science* 259, 1288-1293.
- Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson.* 43, 259-281.
- Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967-974.
- Marmorstein, R., Carey, M., Ptashne, M., & Harrison, S. C. (1992) *Nature* 356, 408-414.
- McIntosh, L. P., Griffey, R. H., Muchmore, D. C., Nielson, C. P., Redfield, A. G., & Dahlquist, F. W. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 1244-1248.
- McLachlan, A. D., & Stewart, M. (1975) *J. Mol. Biol.* 98, 293-304.
- Minor, D. L., & Kim, P. S. (1994) *Nature* 367, 660-663.
- O'Shea, E. K., Rutkowski, R., & Kim, P. S. (1989) *Science* 243, 538-542.
- O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. A. (1991) *Science* 254, 539-544.
- O'Shea, E. K., Lumb, K. J., & Kim, P. S. (1993) *Curr. Biol.* 3, 658-667.
- Oas, T. G., & Kim, P. S. (1988) *Nature* 336, 42-48.
- Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W., & Kim, P. S. (1990) *Biochemistry* 29, 2891-2894.
- Pauling, L. P., & Corey, R. B. (1953) *Nature* 171, 59-61.
- Richardson, J. S. (1981) *Adv. Protein Chem.* 34, 167-339.
- Rose, G. D. (1979) *J. Mol. Biol.* 134, 447-470.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct.* 21, 95-118.
- Shaka, A. J., Keeler, J., & Freeman, R. (1983) *J. Magn. Reson.* 53, 313-340.
- Shin, H., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry* 32, 6356-6364.
- Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., & Baldwin, R. L. (1987) *Nature* 326, 563-567.
- Sklenar, V., & Bax, A. (1987) *J. Magn. Reson.* 74, 469-479.
- Smith, C. K., Withka, J. M., & Regan, L. (1994) *Biochemistry*, in press.
- Staley, J. P., & Kim, P. S. (1990) *Nature* 344, 685-688.

- Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol.* 185, 60-89.
- Thompson, K. S., Vinson, C. R., & Freire, E. (1993) *Biochemistry* 32, 5491-5496.
- Wu, L. C., Laub, P. B., Elove, G. A., Carey, J., & Roder, H. (1993) *Biochemistry* 32, 10271-10276.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
- Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S. C., & Branton, D. (1993) *Science* 262, 2027-2030.
- Zoller, M. J., & Smith, M. (1982) *Nucleic Acids Res.* 10, 6487-6500.
- Zuiderweg, E. R. P. (1990) *J. Magn. Reson.* 86, 346-357.



## Figure Legends

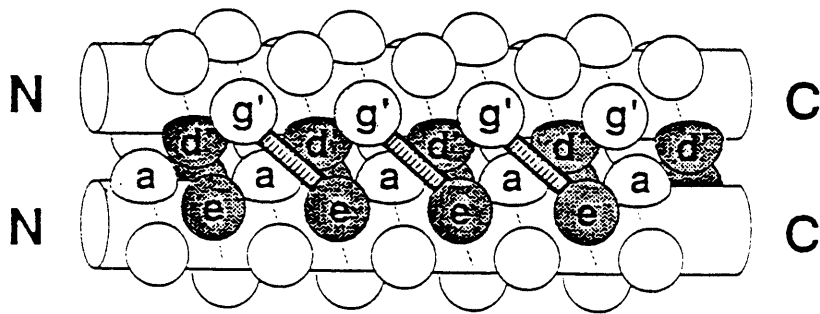
### Figure 1

(A) Schematic representation of a parallel, two-stranded coiled coil (O'Shea et al., 1991). A side view is shown. For simplicity, the supercoiling of the helices is not depicted. The sequences of coiled coils are characterized by a heptad repeat of seven amino acid residues, denoted **a** to **g**, with a 4, 3 repeat of predominantly hydrophobic residues at positions **a** and **d** (Hodges et al., 1972; McLachlan & Stewart, 1975). The hydrophobic interface between the two  $\alpha$ -helices is formed by residues at positions **a**, **d**, **e** and **g**. Prime (') refers to positions from the other helix. One set of packing interactions consists of side chains from positions **a**, **a'**, **g** and **g'**, whereas the second set consists of side chains from positions **d**, **d'**, **e**, and **e'** (O'Shea et al., 1991). Residues at positions **e** and **g** pack against positions **a** and **d**, as well as participating in interhelical electrostatic interactions, which are indicated with bridges (position **e** to **g'** of the preceding heptad).  $pK_a$  measurements in a designed heterodimeric coiled coil suggest, however, that interhelical salt bridges do not contribute significantly to the stability of a two-stranded coiled coil (O'Shea et al., 1993).

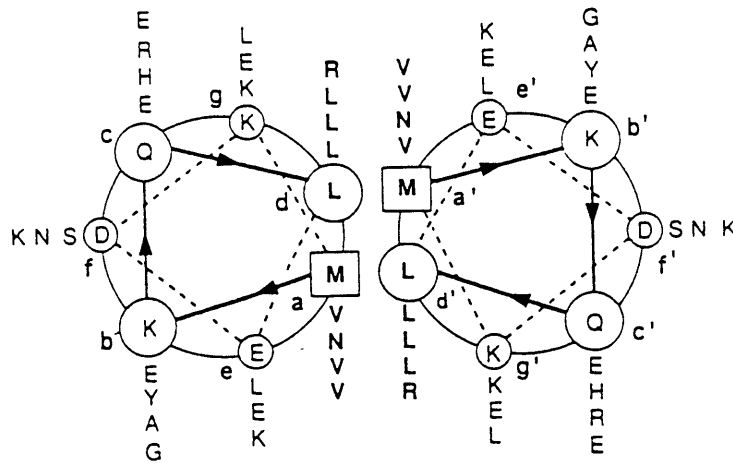
(B) Helical wheel representation of GCN4-p3, which corresponds to residues Met 2 to Arg 33 of GCN4-p1 (Table 1) and has an unblocked N-terminus. GCN4-p3 differs slightly from GCN4-p1 (O'Shea et al., 1989) which has a neutral (acetylated) N-terminus and begins at Arg 1, rather than Met 2. The view is from the N-terminus.

Figure 1

A



B



**Figure 2**

(A) CD spectra of GCN4-p1 and GCN4-p3 at 0 °C and pH 7. The minima at 208 and 222 nm indicate that both peptides are helical. However, GCN4-p3 (filled circles) is only approximately two-thirds helical by CD, in contrast to GCN4-p1 (open squares) which is essentially fully helical.

(B) Temperature dependence of  $[\theta]_{222}$  for GCN4-p3 (filled circles) and GCN4-p1 (open squares) at pH 7. Despite being only about two-thirds helical by CD, the temperature dependence of the CD signal of GCN4-p3 exhibits a folded baseline (where  $[\theta]_{222}$  shows little change with temperature) prior to the major unfolding transition region.

(C) Continuous sets of  $d_{NN}(i, i+1)$ ,  $d_{NN}(i, i+2)$  and  $d_{\alpha N}(i, i+3)$  NOEs indicate (Wüthrich, 1986) that GCN4-p3 is helical from Leu 5 to at least Leu 29. A solid bar indicates that the NOE was observed unambiguously, whereas a hashed bar indicates that a NOESY cross peak was observed but could not be assigned unambiguously.

(D) The amide NH region of the NOESY spectrum of GCN4-p3.  $d_{NN}(i, i+1)$  NOEs are observed from Leu 5 to Leu 29. For clarity, the contour plot has been made at a high level, so that the  $d_{NN}(i, i+2)$  cross peaks are not visible. The amide proton chemical shift degeneracy prevented unambiguous observation of the  $d_{NN}(28, 29)$  and the  $d_{NN}(i, i+1)$  NOEs for residues 30 to 33 (see also Figure 3).

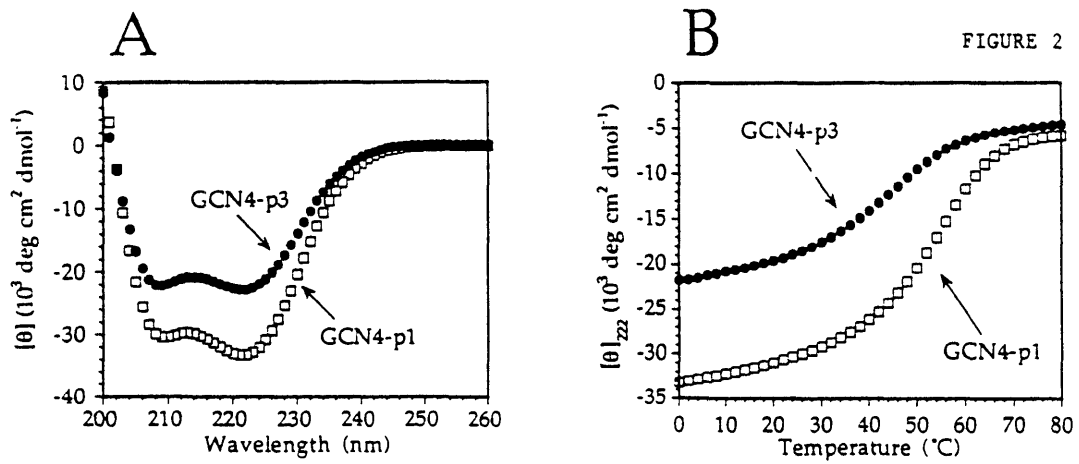
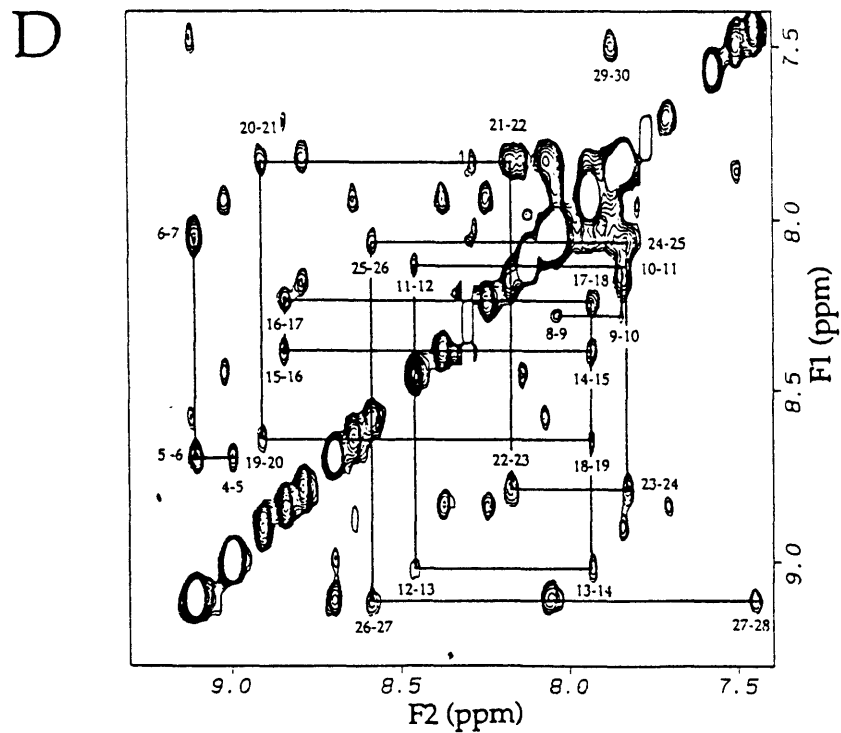
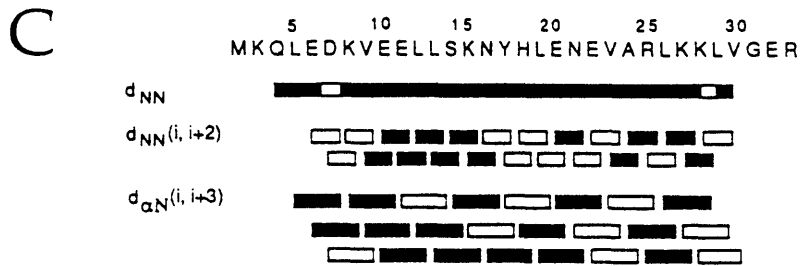


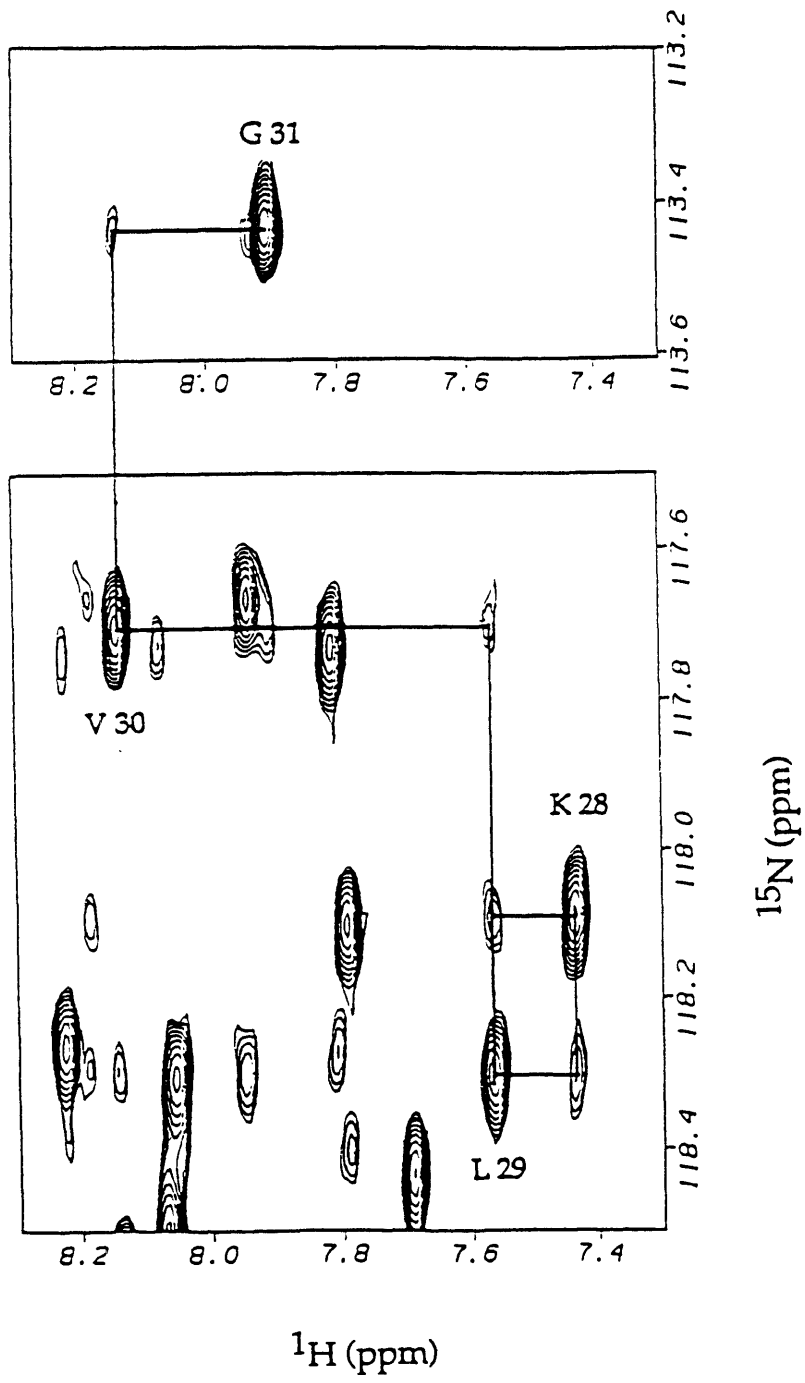
FIGURE 2



**Figure 3**

Region of the 2D  $^{15}\text{N}$ - $^1\text{H}$  HSMQC-NOESY spectrum of GCN4-p3C. In addition to the  $d_{\text{NN}}(i, i+1)$  NOEs observed in the NOESY spectrum of GCN4-p3 (Figure 2),  $d_{\text{NN}}(i, i+1)$  NOEs are observed between residues 29 to 31. NOEs are not observed between residues 31 to 33.

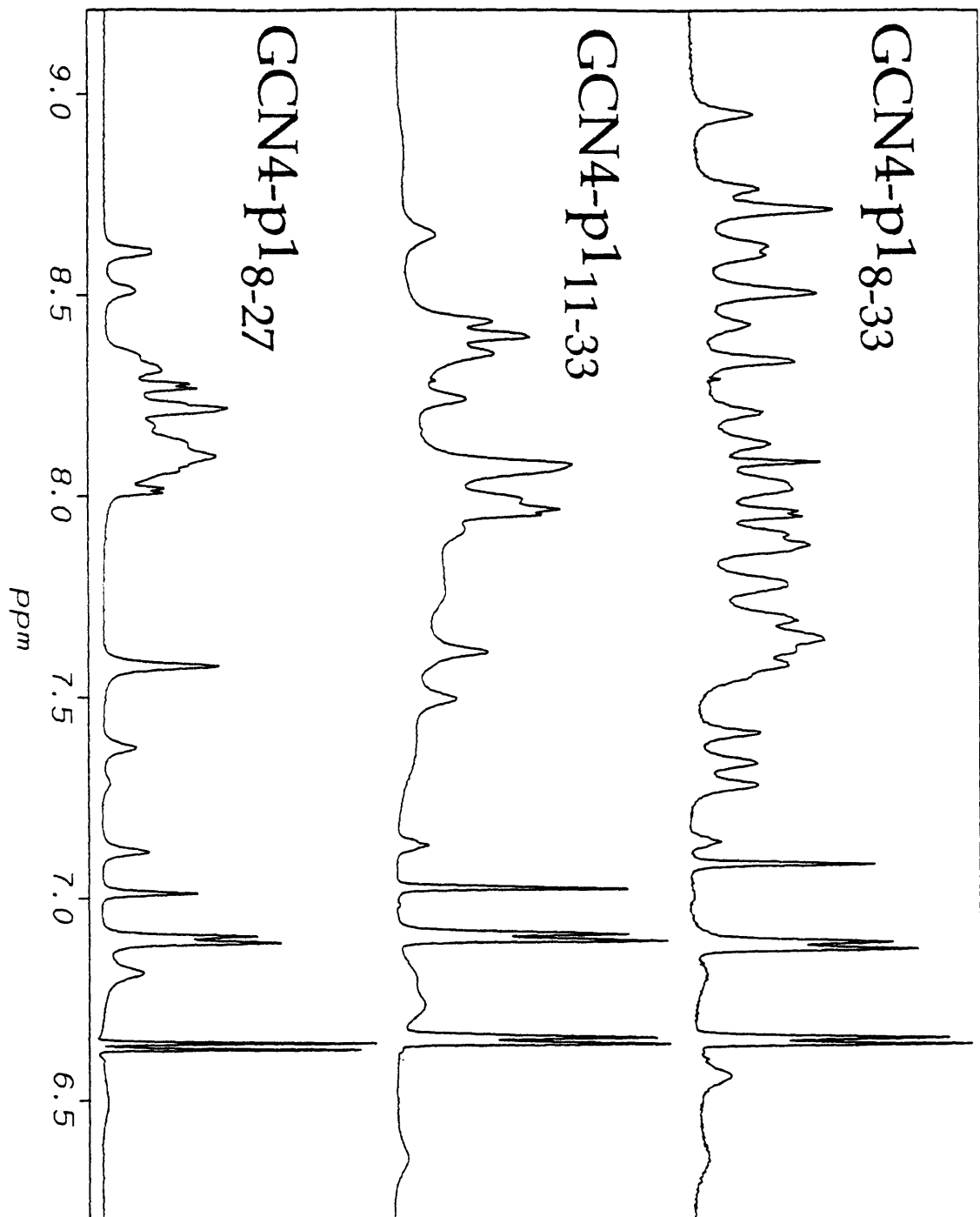
FIGURE 3



**Figure 4**

The amide NH and aromatic region of the 1D NMR spectrum of GCN4-p18-33 is typical of a folded, two-stranded coiled coil (Oas et al., 1990). In contrast, the amide NH resonances in the 1D NMR spectrum of GCN4-p111-33 are broad, perhaps as a result of chemical exchange on the intermediate timescale between multiple conformations or between the unfolded and folded states. The spectrum of GCN4-p18-27 indicates that the peptide is predominantly unfolded, with the amide proton chemical shifts close to those expected for an unstructured peptide (Wüthrich, 1986).

FIGURE 4





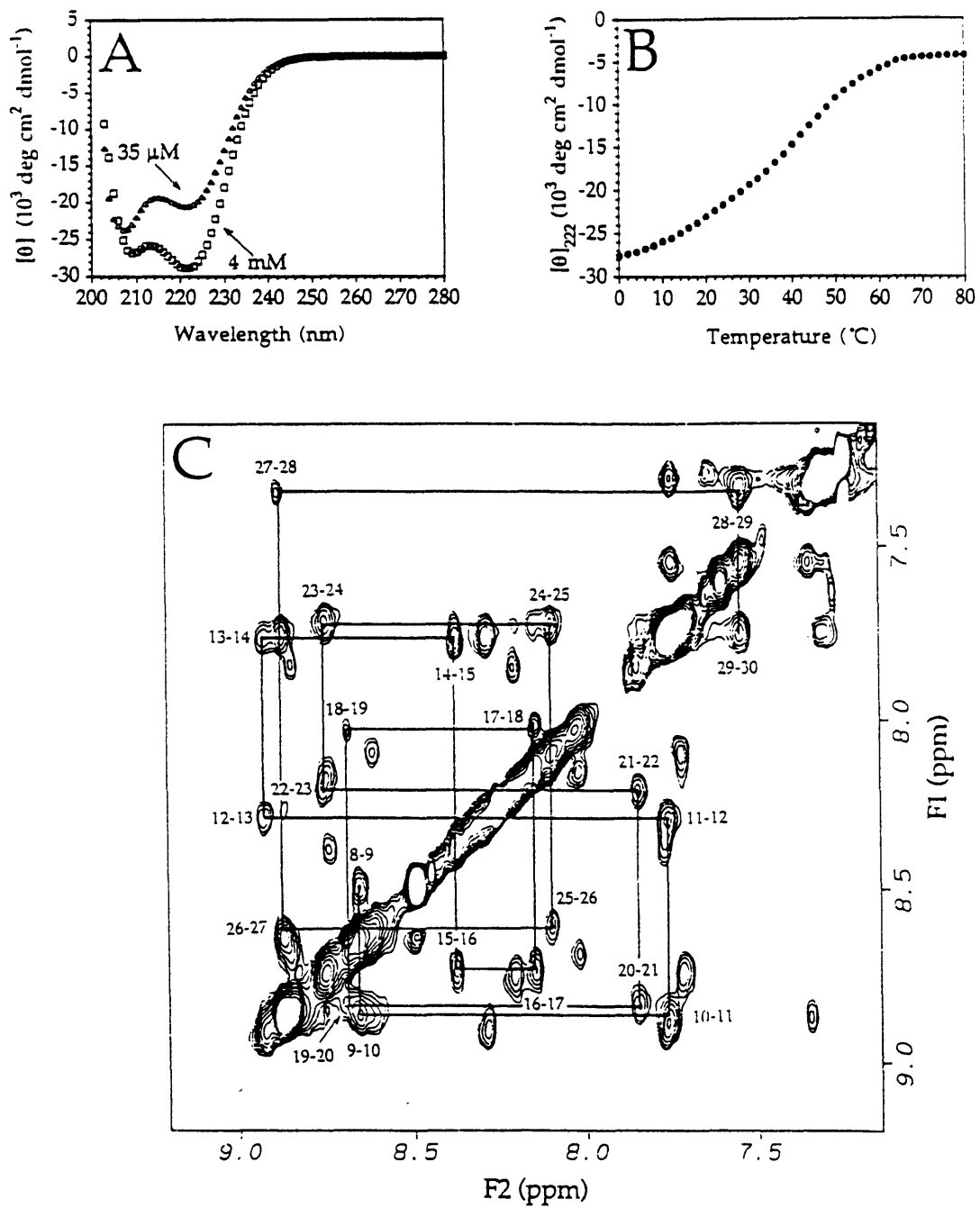
**Figure 5**

**(A)** CD spectra of GCN4-p18-30 at 0 °C and pH 7 indicate that the peptide is helical, and that the helix content is concentration dependent, consistent with self-association. At mM concentrations the peptide is greater than 90 % helical at low temperature. At 0 °C, GCN4-p18-30 is a dimer (Table 1).

**(B)** The thermal melt of 1 mM GCN4-p18-30. The peptide exhibits a cooperative unfolding transition.

**(C)** The amide NH region of the NOESY spectrum of GCN4-p18-30. dNN (i, i+1) NOEs that are indicative of helical structure are observed from Lys 8 to Val 30. For clarity, the contour plot has been made at a high level, so that the dNN (i, i+2) cross peaks are not visible.

FIGURE 5



### Table 1: Circular Dichroism and Sedimentation Equilibrium Data

<sup>a</sup>  $[\theta]_{222}$  was measured at 0 °C at peptide concentrations (monomer) of 35  $\mu\text{M}$  and 1 mM. GCN4-p1 $\Delta$ Ac, GCN4-p14-33, GCN4-p18-33 and GCN4-p18-30 are not fully helical at 35  $\mu\text{M}$ , but are >90 % helical at mM concentrations as judged by  $[\theta]_{222}$  and continuous sets of  $d_{\text{NN}}(i, i+1)$  NOEs (data shown for GCN4-p18-30 in Figure 5C).

<sup>b</sup> The  $T_{\text{m}}$  was determined at peptide concentrations (monomer) of 35  $\mu\text{M}$  and 1 mM and is accurate to  $\pm 1$  °C.

<sup>c</sup> Apparent molecular weights were determined at 0 °C. The expected value for a dimer is enclosed in parentheses.

<sup>d</sup> Data from O'Shea et al. (1989).

<sup>e</sup> The concentration dependence of  $[\theta]_{222}$  and the apparent molecular weight for a 1 mM sample of GCN4-p111-33 suggest that GCN4-p111-33 self-associates. The expected molecular weight indicated in parentheses is for a 1 mM solution of peptide, assuming an unfolded monomer-folded dimer equilibrium and a helix content of 45 % estimated from the value of  $[\theta]_{222}$ .

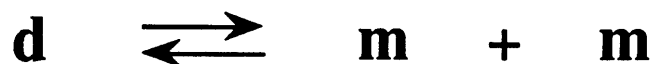
Table 1: Circular Dichroism and Sedimentation Equilibrium Data

peptide	sequence	$[\theta]_{222}^{\circ}$ ( $\times 10^3$ deg cm <sup>2</sup> dmol <sup>-1</sup> )		$T_m^b$ (°C)		molecular mass <sup>c</sup> (Da)
		35 $\mu$ M	1 mM	35 $\mu$ M	1 mM	
GCN4-p1	AcRMKQLEDKVEEILSKNYHLENEVARLKKLVGER	-33.2	-33.3	56	68	7900 (8080) <sup>d</sup>
GCN4-p3	MKQLEDKVEEILSKNYHLENEVARLKKLVGER	-22.0	-24.5	47	65	7790 (7681)
N-terminal modifications						
GCN4-p1ΔAc	RMKQLEDKVEEILSKNYHLENEVARLKKLVGER	-27.7	-30.0	49	67	7620 (7995)
GCN4-p3ΔAc	AcMKQLEDKVEEILSKNYHLENEVARLKKLVGER	-32.3	-33.3	51	69	7600 (7769)
N-terminal truncations						
GCN4-p1 <sub>4-33</sub>	AcQLEDKVEEILSKNYHLENEVARLKKLVGER	-31.1	-33.2	43	63	7120 (7246)
GCN4-p1 <sub>8-33</sub>	AcKVEEILSKNYHLENEVARLKKLVGER	-24.0	-33.2	35	50	5710 (6275)
GCN4-p1 <sub>11-33</sub>	AcEEILSKNYHLENEVARLKKLVGER	-8.8	-14.7 <sup>e</sup>	<0	<0	3600 (3900) <sup>e</sup>
C-terminal truncations						
GCN4-p1 <sub>8-30</sub>	AcKVEEILSKNYHLENEVARLKKLV	-20.7	-27.8	22	43	5160 (5588)
GCN4-p1 <sub>8-27</sub>	AcKVEEILSKNYHLENEVARLK	-6.5	-7.0	<0	<0	

## Appendix III

The Free Energy of unfolding 1 mole of dimer to 2 moles of monomer is expressed in units of Molar, and the Free Energy is calculated using the following equation:

## Bimolecular Equilibrium Equation



$f_f + f_u = 1$                        $f_f$  is fraction folded dimer and  $f_u$  is fraction unfolded monomer.

$C_f = 1/2 f_f C_T$                        $C_f$  is concentration folded dimer and  $C_T$  is total concentration.

$C_u = f_u C_T$                        $C_u$  is concentration unfolded monomer.

$$K_d = C_u^2 / C_f = (f_u C_T)^2 / (1/2 f_f C_T) = 2 f_u^2 C_T / (1 - f_u) = e^{-\Delta G/RT}$$

$K_d$     dissociation constant in units of Molar.

$\Delta G$     equilibrium free energy.

$R$       gas constant, 1.9872 cal / mol deg-K.

$T$       temperature, in Kelvin.

I.      To describe the equilibrium in terms of the observable variable ( $\theta_{obs}$ , CD signal), solve in terms of fraction monomer ( $f_u$ ).

$$K_d = 2 f_u^2 C_T / (1 - f_u) = e^{-\Delta G/RT}$$

Rearrange:

$$e^{-\Delta G/RT} / 2 C_T = f_u^2 / (1 - f_u)$$

Substitute  $e^{-\Delta G/RT}$  with  $K_d$  and cross multiply:

$$K_d (1 - f_u) = 2 C_T f_u^2$$

Or

$$2 C_T f_u^2 - K_d (1 - f_u) = 0$$

Rearrange to:

$$f_u^2 + K_d f_u / 2 C_T - K_d / 2 C_T = 0$$

Using the quadratic equation, solve for  $f_u$  :

$$f_u = 1/2 \{ -K_d / 2 C_T \pm [ (K_d / 2 C_T)^2 - 4 K_d / 2 C_T ]^{1/2} \}$$

Simplify to:

$$f_u = -K_d / 4 C_T + ( K_d^2 / 16 C_T^2 + K_d / 2 C_T )^{1/2}$$

To put it in terms of the observed CD signal (which varies), use a 2-state equation:

$$\theta_{obs} = f_f \theta_f + f_u \theta_u$$

$\theta_{obs}$ : the observed signal (CD signal, here).

$\theta_f$  and  $\theta_u$ : the folded dimer signal and unfolded monomer signal, respectively.

And

$$\theta_{obs} = ( 1 - f_u ) \theta_f + f_u \theta_u$$

$$\theta_{obs} = \theta_f - \theta_f f_u + f_u \theta_u$$

Solve for  $f_u$  :

$$f_u = ( \theta_{obs} - \theta_f ) / ( \theta_u - \theta_f )$$

From above:

$$f_u = -K_d / 4 C_T + ( K_d^2 / 16 C_T^2 + K_d / 2 C_T )^{1/2}$$

Therefore:

$$(\theta_{\text{obs}} - \theta_f) / (\theta_u - \theta_f) = -K_d / 4 C_T + (K_d^2 / 16 C_T^2 + K_d / 2 C_T)^{1/2}$$

Solve for  $\theta_{\text{obs}}$  :

$$\theta_{\text{obs}} = \theta_f + (\theta_u - \theta_f) [ -K_d / 4 C_T + (K_d^2 / 16 C_T^2 + K_d / 2 C_T)^{1/2} ]$$

II. For calculations of  $\Delta G$  from a Guanidine Unfolding transition:

Substitute:

$$K_d = e^{-\Delta G/RT}$$

Where,

$$\Delta G = \Delta G_{\text{H}_2\text{O}} - m [\text{GuHCl}]$$

$m$  is defined in (Pace, CN, [1986] Meth. Enz. 131, 266, 279) and  $[\text{GuHCl}]$  is the concentration of guanidine hydrochloride.

Thus:

$$K_d = e^{-\Delta G_{\text{H}_2\text{O}} - m [\text{GuHCl}] / RT}$$

In the Kalidograph application, make a general equation for the fit:

$$m3 + (m4 - m3) * ((-\exp(-m1 + m2 * m0)) / (4 * c) + ((\exp(-m1 + m2 * m0) * \exp(-m1 + m2 * m0)) / (16 * c * c) + (\exp(-m1 + m2 * m0)) / (2 * c))^{1/2})$$

$c$  = total protein concentration in Molar, here,  $c$  is  $34\mu\text{M}$  or  $0.000034 \text{ M}$

$m0$  = the  $x$  variable, here  $[\text{GuHCl}]$

$m1 = \Delta G_{\text{H}_2\text{O}} / RT$

$m2 = m / RT$

$m3$  = the folded baseline,  $\theta_f$ , which can be a constant or a line.

$m4$  = the unfolded baseline,  $\theta_u$ , which can be a constant or a line.



## Biographical Profile

**Chavela M. Carr**

Whitehead Institute for Biomedical Research  
Nine Cambridge Center, Cambridge, MA 02142  
Phone: (617) 258-5206  
FAX: (617) 258-7226

Residence:  
143 Albany St. Apt. 9  
Cambridge, MA 02139  
Phone: (617) 494-0588

### **RESEARCH EXPERIENCE:**

1989-present      Dr. Peter S. Kim's laboratory, MIT., Whitehead Institute, Cambridge, MA. Studies of the coiled-coil motif: structure and function in the proteins GCN4 and influenza hemagglutinin.

1987-1988      Dr. Douglas R. Cavener's laboratory, Vanderbilt University, Nashville, TN. Lineage of Gld expression during development of *Drosophila melanogaster*, *Drosophila pseudoobscura* and *Drosophila virilis*. Organ dissection, Northern analysis, and *in situ* hybridization techniques.

### **EDUCATION:**

1988-present      MIT, Biology, Ph. D. candidate  
1984-1988      Vanderbilt University, Molecular Biology, B. S., 1988

### **TEACHING EXPERIENCE:**

1/9/95      Lecture: IAP Biology Department Lecture Series, MIT  
2/10/94      Poster: Structural Biology Seminar series, Harvard, MIT  
11/12/93      Lecture: Topics in Computational Molecular Biology, MIT  
6/23/93      Poster: Gordon Conference, NH: Animal Cells and Viruses  
2/22/93      Lecture: Structure/Function Seminar series at MIT  
1992      Whitehead Education Program "Molecular Approaches to Human Diseases" instructor for high-school teachers.  
1991      Biochemistry Lab (7.011), TA, laboratory and recitation instructor  
1989      General Biochemistry (7.05), TA, recitation instructor.

### **ACHIEVEMENTS:**

1984      Phi Eta Sigma and Alpha Lambda Delta Freshman Honor Societies  
1988      Magna cum Laude, Vanderbilt University, Cumulative GPA 3.729  
1988      Phi Beta Kappa  
1988      Honors in Molecular Biology, Vanderbilt University,  
Thesis: The Development of Gld Regulation in *Drosophila*  
1988      Outstanding Research in Molecular Biology Award, Vanderbilt  
1988      Genetics Society of America Undergraduate Research Award  
1994      Sigma Xi, Massachusetts Institute of Technology Chapter

### **PERSONAL PROFILE:**

Born November 20, 1965, Indianapolis, IN, USA  
Female, marital status: single



Room 14-0551  
77 Massachusetts Avenue  
Cambridge, MA 02139  
Ph: 617.253.5668 Fax: 617.253.1690  
Email: docs@mit.edu  
<http://libraries.mit.edu/docs>

## **DISCLAIMER OF QUALITY**

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

**Some pages in the original document contain color pictures or graphics that will not scan or reproduce well.**