In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written, in his/her absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

_____

      Wookhyun Kim

Engineering of Elastin Using Noncanonical Amino Acids

By

Wookhyun Kim
Doctor of Philosophy

Department of Chemistry

_____
Vincent P. Conticello
Adviser

_____
Justin Gallivan
Committee Member

_____
Stefan Lutz
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

Engineering of Elastin Using Noncanonical Amino Acids

By

Wookhyun Kim
B.S., Dongguk University, 1997
M.S., Seoul National University, 1999

Adviser: Vincent P. Conticello, Ph.D.

An Abstract of
A dissertation submitted to the Faculty of the Graduate
School of Emory University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Chemistry

2007

ABSTRACT

Protein engineering has been widely used as a tool for globular protein design and modification, and also for engineering of protein-based materials. Conventional protein engineering approaches to protein-based materials have several limitations in that amino acid substitution in some cases alters the properties of the materials, and only the naturally occurring amino acids must be utilized. However, the use of noncanonical amino acids in protein engineering allows for maintaining the structure of the corresponding canonical amino acids, especially the imino acid proline, and introducing various functionalities into the side chain. Proline residues play important roles in protein structure and function. In order to more closely approximate the structural aspects of proline residue in a structurally important position, the substitution of proline with a proline analogue would be desirable for investigating contribution of proline to the conformational stability of the proteins.

The second chapter has focused on developing a series of proline auxotrophic *E. coli* expression strains that are competent for multi-site incorporation of a structurally diverse series of proline analogues into elastin-mimetic polypeptides and culture conditions that are compatible with high levels of analogue substitution within elastin-

mimetic sequences. In order to achieve these, hyperosmotic culture media has been used in addition to overexpression of wild-type *E. coli* prolyl-tRNA synthetase and variants derived from site-directed mutagenesis of active site residues. The efficacy of co-translational incorporation judged by protein yield depended on the structural similarity between proline and proline analogues. The reliability of the methods was examined by quantitative and qualitative analyses of analogue substitution.

The rest of this work describes the effects of fluoroproline substitution on the biophysical properties of elastin-mimetic polypeptides, which may be attributed to the influence of pyrrolidine ring conformation on the stability of turn structure and the resulting self-assembly. Structural analyses of epimeric pairs of fluoroproline derivatives revealed that stereoelectronic and steric effects altered the main-chain dihedral angles and the ring pucker preference. Biophysical and computational studies of fluoro-elastins provided evidence that these two factors contribute to observed biophysical differences that arise from the presence of fluoroprolines at the structurally critical positions in the polypeptide sequence.

ACKNOWLEDGEMENT

First, I would like to thank my adviser, Dr. Vincent P. Conticello, for his support and guidance. Without his advice and encouragement, this thesis work could not have seen the light. Vince allowed me to do interdisciplinary works in the lab. I have enjoyed things including synthetic works and biophysical studies. I would like to thank my committee members, Dr. Lutz and Dr. Gallivan, for taking the time to read this thesis and their comments. I also gratefully thank Dr. Conticello, Lutz, and Wu for their help with my postdoctoral position search.

My thanks are extended to all my lab members for their friendships. I appreciate the former graduate, Dr. R. Andrew McMillan at NASA, for his support in obtaining initial DSC data. My thanks are going to two undergraduates from Oglethorpe University, Melissa Evans and Anna George, for experiments in my first publication. Melissa Evans helped me when I first joined this lab. I want to give my heartfelt thanks to Dr. Shaoxiong Wu in the Emory University NMR Research Center. In fact, Dr. Wu has really been generous and supportive in obtaining excellent NMR data, and even planning my future. I would like to thank Dr. Jim Snyder for his support in computational studies of model peptides. I also would like to thank Dr. Kenneth Hardcastle and Dr. Robert P.

Engineering of Elastin Using Noncanonical Amino Acids

By

Wookhyun Kim
B.S., Dongguk University, 1997
M.S., Seoul National University, 1999

Adviser: Vincent P. Conticello, Ph.D.

A dissertation submitted to the Faculty of the Graduate
School of Emory University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Chemistry

2007

Table of Contents

List of Figures

List of Tables

List of Schemes

CHAPTER 1

Introduction

1. Proline-containing motifs

Proline-containing structural motifs are frequently found in peptides and proteins. The presence of a proline residue within a polypeptide chain can provide significant insight into the protein structure and properties as proline is restricted to certain secondary structure elements due to its constraint of main-chain torsion angles and its inability to form hydrogen bonds that stabilize regular α/β secondary structures.[1] In addition, proline plays an important role in protein folding and cellular signaling associated with the cis/trans isomerization of prolyl peptide bond that leads to a local conformational change.[2-4] Sequence-based structural analyses of proteins have suggested that proline residues are observed in polyproline (PPII) helices, turns, and loops with high frequency, but rarely in periodic secondary structures with the notable exception of prolines in transmembrane helical domains or as helical kinks.[5] For example, green fluorescent protein (GFP) contains 10 proline residues in amino acid sequence. The structure analysis of GFP revealed that the proline residues are exclusively positioned in non-periodic secondary structures (Figure 1).[6] The chromophore of GFP is protected from outside bulk water by a β-barrel structure consisting of 11 β-strands through the tight hydrogen-bonding interactions, in which proline residues are notably

A.
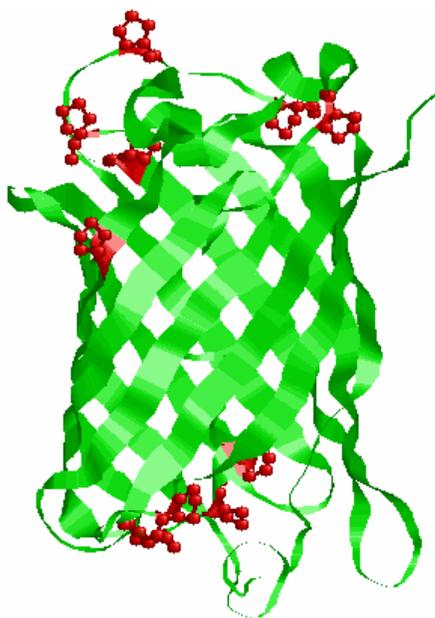
B.
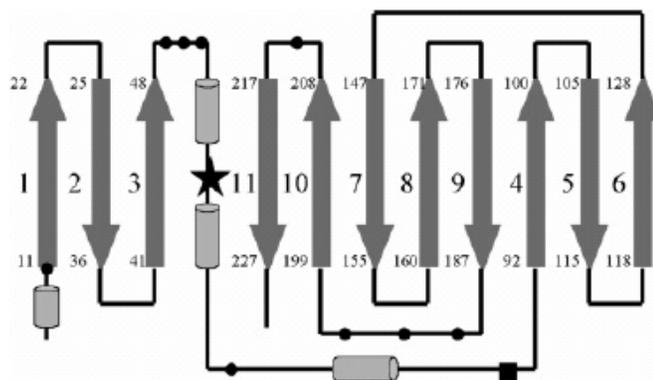
Figure 1. (**A**) Structure of green fluorescent protein (GFP) with proline residues (red), (**B**) secondary structure elements of GFP indicating the chromophore (star) proline residues in the trans conformation (filled circle), and in the cis conformation (square)[6]

absent.

Biologically important proline-containing structural domains include short ligand binding domains associated with signal transduction and modular structural domains of extracellular matrix (ECM) proteins.[5,7] The former largely consist of non-repeating proline-rich motifs and the latter of tandem repeats of oligopeptide motifs containing regularly positioned proline residues. The non-repeating proline-rich sequences adopt a PPII helical conformation that can interact with signaling domains. These PPII helices are exposed in the outer surfaces of proteins, linearly extended, and thus facilitate binding to the Src homology 3 (SH3) and WW domains. Structural studies of mammalian prion also suggested that a short proline motif forms a PPII helical structure to undergo further posttranslational modification, while other proline-containing repetitive oligopeptides at N-terminus region form turn structures in the presence of metals.[8]

The modular domains of fibrous proteins are composed of the tandem repeats of oligopeptides of 3-8 residues that may be repeated up to several hundred times. Many repeat domains have proline-containing motifs at structurally critical positions and adopt turn or PPII conformations which may be responsible for the mechanical and thermodynamic behavior of these proteins in their native state. For example, repeat sequences containing the Pro-Gly sequence motif in native protein polymers have

elastomeric behavior through periodic β-turn structures (Table 1).[9] There have been

many efforts made in order to biosynthesize artificial protein-based materials derived

from sequence of natural protein polymers.[10] These genetically engineered biomaterials

can be designed to emulate the native protein structure and can display distinctive

properties as a consequence of the defined sequence and chemical composition. Of those

protein-based materials, elastin is an attractive biomolecule for the construction of

synthetic biomaterials in that the elastomeric domains can be designed with diverse

repetitive sequences containing Pro-Gly motif using genetic engineering methods. In

addition, biologically functional peptide motifs can be incorporated co-translationally

between the elastin-mimetic sequences for the purpose of protein and cell binding. The

versatility of elastin-mimetic polypeptides has been realized a few decades ago and

potential applications are now being explored in drug delivery, tissue engineering and

other applications such as protein purification.[11] Elastin-based biomaterials have a

promising future in the fields of materials science and biomedical engineering.

Protein	Repeat Units
Elastin	Val-**Pro-Gly**-Gly Val-**Pro-Gly**-Val-Gly Ala-**Pro-Gly**-Val-Gly-Val Val-**Pro-Gly**-Phe-Gly-Val-Gly-Ala-Gly
Flagelliform Silk	Gly-**Pro-Gly**-Gly-Xaa (Xaa = Ser, Tyr, Ala, Val)
Dragline Silk	Gly-**Pro-Gly**-Gln-Gln Gly-**Pro-Gly**-Gly-Tyr Gly-Gly-Tyr-Gly-**Pro-Gly**-Ser
Glutenin HMW subunits	Gln-**Pro-Gly**-Gln-Gly-Gln
Resilin	Gly-Gly-Arg-Pro-Ser-Asp-Ser-Tyr-Gly-Ala-**Pro-Gly**-Gly-Gly-Asn Gly-Tyr-Ser-Gly-Gly-Arg-**Pro-Gly**-Gly-Gln-Asp-Leu-Gly
Mussel Byssus	Gly-**Pro-Gly**-Gly-Gly

Table 1. Repeat sequences of native protein-based elastomeric materials containing the "Pro-Gly" sequence motif.[9]

2. Elastin-mimetic polypeptides

Elastin, along with collagen, is a major protein component of extracellular matrix (ECM), which is found in many tissues such as skin, lung, artery, ligament and cartilage. The primary sequence analysis of the soluble precursor protein, tropoelastin, has revealed that the native elastin is composed of elastomeric domains and cross-linking domains. The elastomeric domains are rich in hydrophobic amino acids, and have highly repetitive tetra-, penta-, and hexapeptides sequences (VPGG, VPGVG, and APGVGV, respectively). The cross-linking domains are rich in lysines, which are involved chemically in the cross-linking reactions. Mature elastin is extremely insoluble and forms filamentous fibers due to interaction between hydrophobic structural segments that are held in proximity in the molecular network of the hydrophobic domain and the cross-linking domain.

The thermodynamic and viscoelastic properties of the elastomeric domains can be emulated by elastin-mimetic polypeptides generated through a concatenation of the native oligopeptide sequence motifs; the most common of which is the pentapeptide (Val-Pro-Gly-Val-Gly).[11] Polypeptides based on these pentameric repeat sequences undergo reversible, temperature-dependent, hydrophobic assembly from aqueous solution in

analogy to the phase behavior of native tropoelastin. This process results in spontaneous phase separation of the polypeptide above a lower critical solution temperature (LCST), which coincides with a conformational rearrangement of the local secondary structure within the pentapeptide motifs.

Precise determination of the detailed molecular structure of the native elastin is difficult using high-resolution structural analysis, such as NMR or X-ray crystallographic methods, because the molecule is insoluble above the transition temperature and conformationally mobile in solution below the temperature. Biophysical studies of elastin-mimetic polypeptides have suggested the structural development within the pentapeptide repeats during the temperature change.[12,13] The pentameric sequence units undergo a conformational transition from a random coil to a type II β-turn as the temperature approaches the T_t (Figure 2).[14] The facility with which this process occurs critically influences the thermodynamics of the phase transition and is crucial for assembly of the physiologically relevant coacervate state of native elastin. However, molecular dynamics simulations of (Val-Pro-Gly-Val-Gly) concatemers and spectroscopic investigations of native elastin-mimetic model peptides also suggested the presence of alternative conformations that may contribute to structural ensemble above the T_t, and thus influence the elastomeric properties of the material.[15] The relative

importance of these alternative conformations to the macroscopic mechanical properties

of elastin and elastin-mimetic materials should be considered.
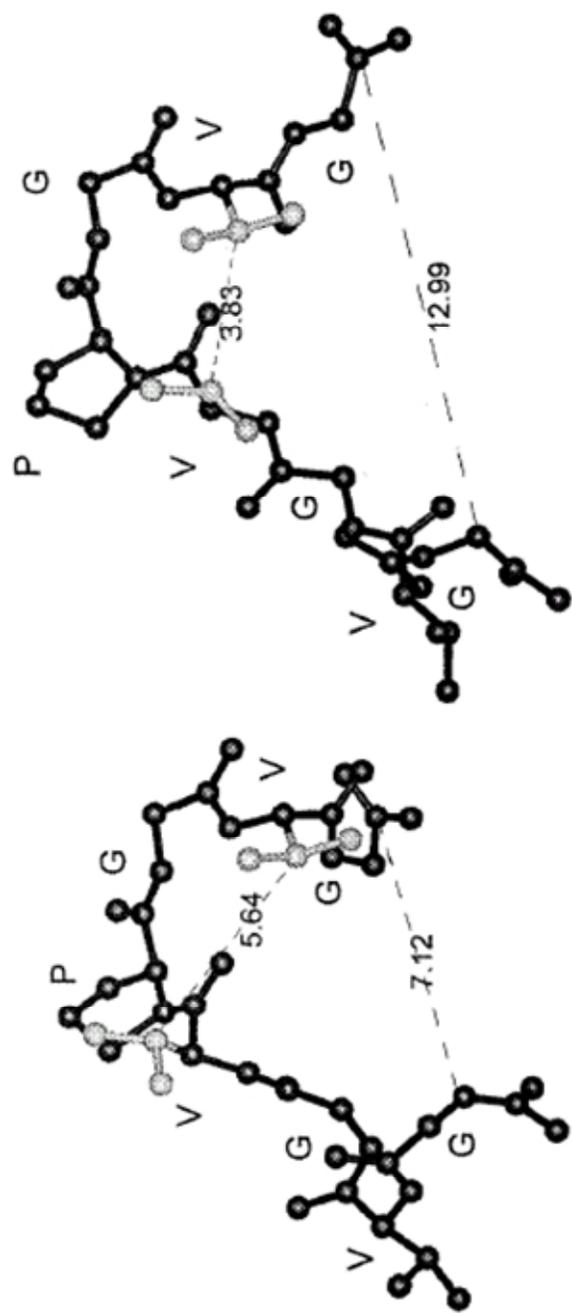
Figure 2. Molecular models of peptide GVGVPGVG in the low temperature conformation (left) and the high temperature conformation (right).[15]

3. Amino acid substitutions in engineering elastin

Amino acid substitutions at specific positions of the pentapeptide repeat have been examined for understanding the structure-property relationships of elastin-based biomaterials.[16-19] Limited success has been observed for substitution of valine 1 by isoleucine and leucine, in which the additional methylene group lowered the transition temperature.[16] Proline at position 2 plays a critical structural role in polypentapeptide as a type II β-turn structure (Figure 3). It has been identified that substitution of glycine 3 by alanine results in the alteration of elastomeric property (VPGVG) to plastic property (VPAVG) which is irreversible phase transition.[17] Valine 4 has been also systematically substituted for other naturally occurring amino acids except for proline.[18] The temperature of coacervation (T_t) of elastin-mimetic sequences could be altered through doping of guest residues at position 4 of the pentapeptide repeat in a manner that depended on the chemical identity of amino acids. Hydrophilic residues shifted the T_t to higher temperature and hydrophobic residues shifted T_t to lower temperatures (Figure 4). While it was possible to make substitution by other amino acids at positions 1, 3, and 4 without significant alteration in structure, substitution of the proline residue substitution was not considered due to its importance in the turn formation and subsequent assembly.

The use of unnatural imino acids represents an alternative method as this substitution is

not likely to significantly perturb the native protein structure.
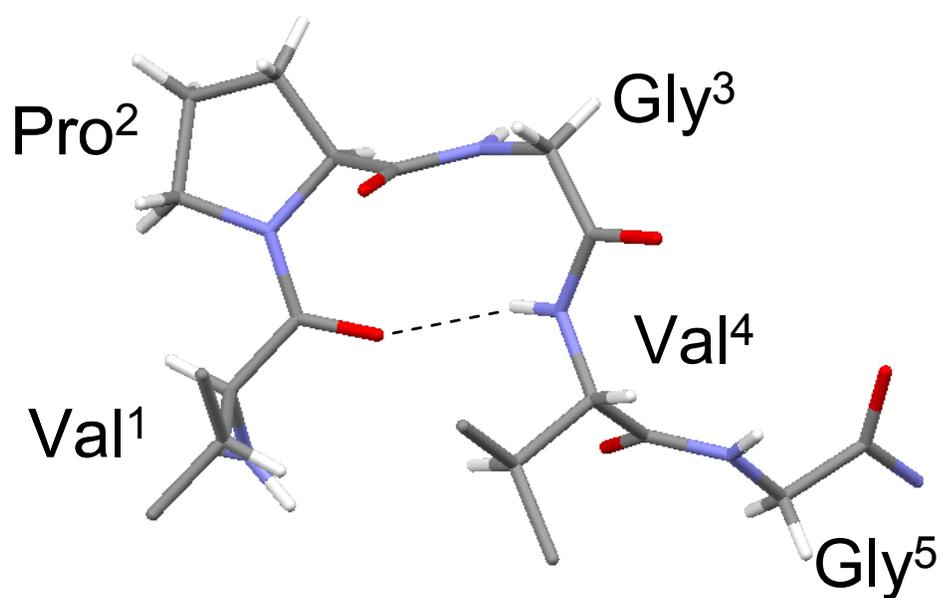
Figure 3. Type II β-turn structure of Pro^2-Gly^3 sequence motif with hydrogen bond between Val^4 N-H and Val^1 C=O obtained from the crystal structure of cyclo-$(VPGVG)_3$
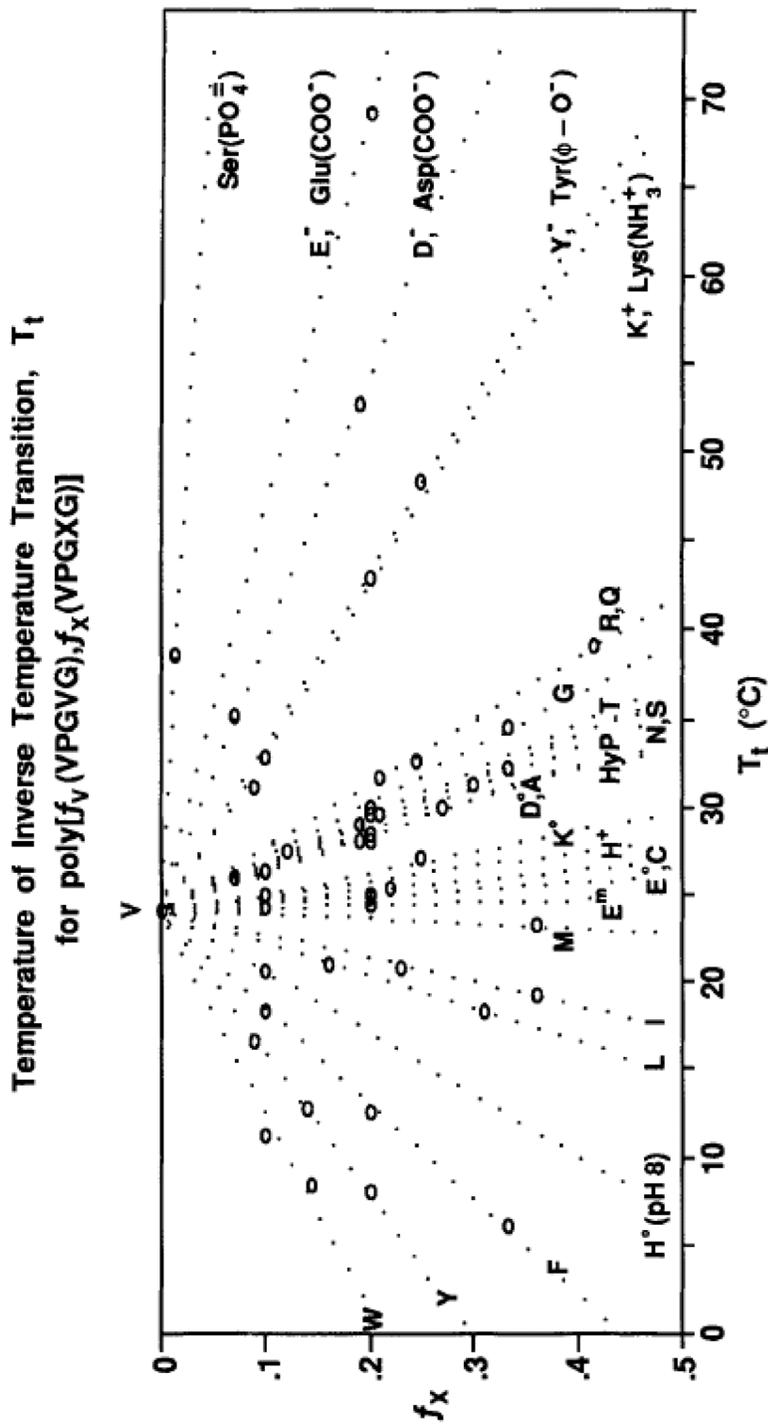
Figure 4. Substitution of valine 4 with hydrophobic or hydrophilic amino acid residues (Xaa) altering transition temperature (T_t) to lower or higher, respectively. A host-guest study showed dependence of T_t on the fraction of guest residues for elastin-mimetic polypeptides.

4. Introduction of noncanonical amino acids as protein engineering method

Historically, mutagenesis of a wild-type amino acid with other naturally occurring amino acids has been used as a technique for protein structural analysis and engineering. This technique is confined, however, by the genetically encoded 20 functional groups. Then, researchers considered the non-genetically encoded amino acids as the new building blocks for *de novo* protein engineering.[20-22] To date a variety of noncanonical amino acids has been available and incorporated into different protein species. New chemical functionalities and remarkable structural effects can be observed upon incorporation of carefully selected noncanonical amino acids in place of canonical amino acids. In addition, applications to protein engineering are beginning to be observed in diverse areas such as the enhancement of protein stability, alteration of fluorescent properties, multi-wavelength anomalous diffraction (MAD) experiments in structural biology, selective labeling proteins or cells and so on.[20-22]

To perform this study, there are two strategies developed for introduction of noncanonical amino acids into proteins. One strategy is a multi-site incorporation. The multiple specific residues encoding degenerate codons are reassigned by noncanonical amino acids through the use of a bacterial auxotrophic host strain with the wild-type or

mutant amino-acyl tRNA synthetase (selective pressure incorporation or codon reassignment).[20,21] This method has been used for the global incorporation of analogues into artificial proteins or protein-based biomaterials (Figure 3).[23] An alternative strategy employs single-site incorporation, which allows a single residue to be replaced with a rational design. This strategy is efficiently achieved by nonsense suppression (stop codons) or frameshift suppression (four-base or five-base codons), relying on orthogonality to introduce a synthetase/tRNA pair from other species (Figure 5).[22]

Although these methods seem to be powerful for incorporating noncanonical amino acids into proteins, in fact, problems still remain such as a low protein yield, possibility of incomplete substitution, truncation, and weak recognition and charging of the amino acid analogue to the tRNA by the synthetase. None of the currently developed methods can be a complete and general method for both single-site and multi-site substitution. Those strategies are complementary to each other and appropriate methods should be used selectively according to the purpose of studies. This thesis will focus on a multi-site incorporation of proline analogues into protein-based materials and proteins. Therefore, in the next chapter, we will begin to discuss establishing systematic methods for global substitution of proline residues in elastin-mimetic sequences with a structurally diverse set of proline analogues.
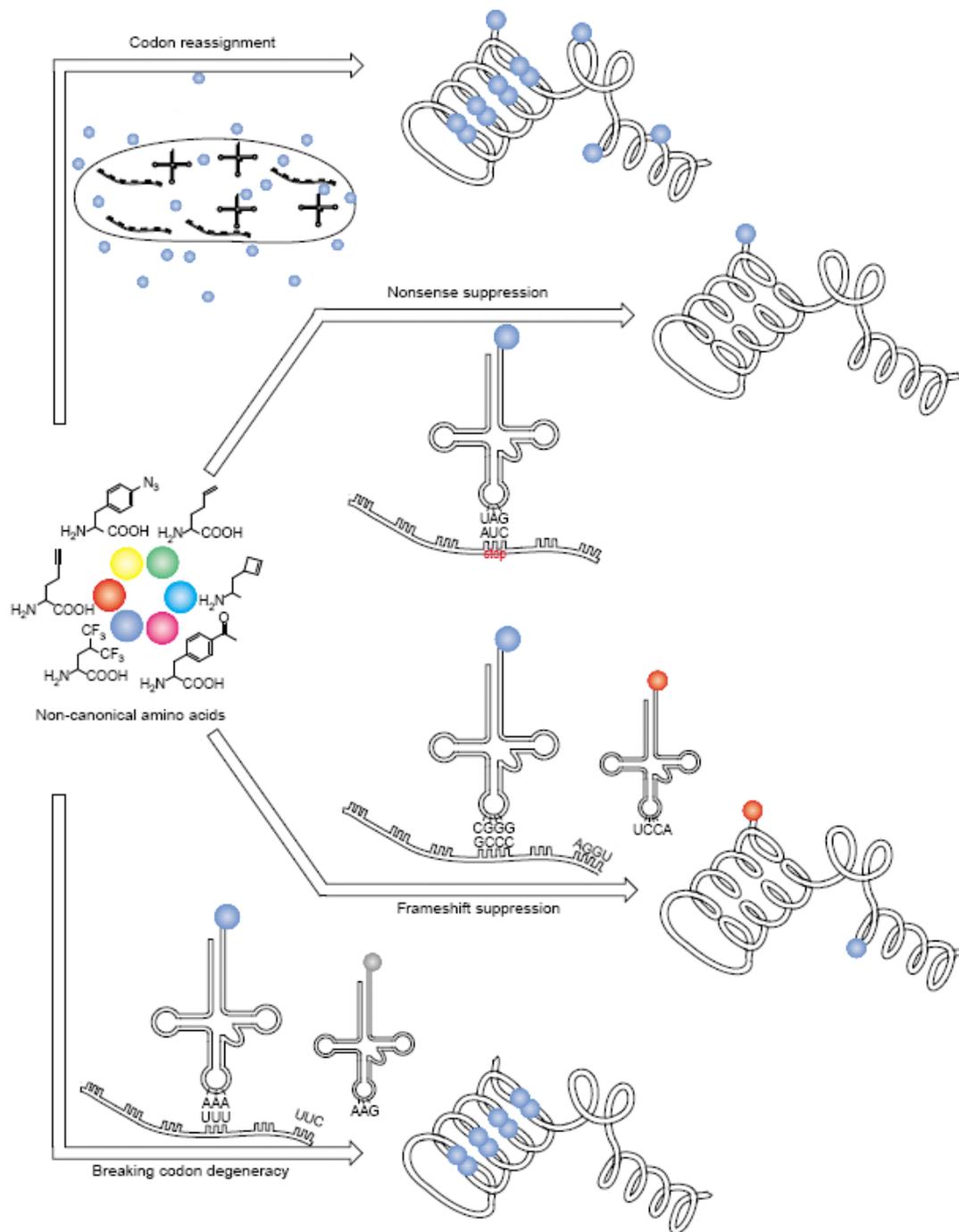
Figure 5. Methods developed for the introduction of noncanonical amino acids into proteins.[21] Global incorporation is enabled by selective pressure incorporation, codon reassignment and breaking codon degeneracy. Single-site incorporation is worked by nonsense suppression and frameshift (4-base and 5-base codons) suppression.

5. References

[1]   MacArthur, M. W.; Thornton, J. M., *J. Mol. Biol.* **1991**, 218, 2, 397-412.

[2]   Wedemeyer, W. J.; Welker, E.; Scheraga, H. A., *Biochemistry* **2002**, 41, 50, 14637-14644.

[3]   Andreotti, A. H., *Biochemistry* **2003**, 42, 32, 9515-9524.

[4]   Lummis, S. C. R.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A., *Nature* **2005**, 438, 7065, 248-252.

[5]   Reiersen, H.; Rees, A. R., *Trends Biochem. Sci.* **2001**, 26, 11, 679-684.

[6]   Enoki, S.; Saeki, K.; Maki, K.; Kuwajima, K., *Biochemistry* **2004**, 43, 44, 14238-14248.

[7]   Kay, B. K.; Williamson, M. P.; Sudol, M., *FASEB J.* **2000**, 14, 2, 231-241.

[8]   Gill, A. C.; Ritchie, M. A.; Hunt, L. G.; Steane, S. E.; Davies, K. G.; Bocking, S. P.; Rhie, A. G. O.; Bennett, A. D.; Hope, J., *EMBO J.* **2000**, 19, 20, 5324-5331.

[9]   Tatham, A. S.; Shewry, P. R., *Trends Biochem. Sci.* **2000**, 25, 11, 567-571.

[10]   van Hest, J. C. M.; Tirrell, D. A., *Chem. Commun.* **2001**, 19, 1897-1904.

[11]   McGrath, K.; Kaplan, D.; Editors, *Protein-Based Materials*. 1997; p 429 pp.

[12]   Urry, D. W.; Shaw, R. G.; Prasad, K. U., *Biochem. Biophys. Res. Commun.*

1985, 130, 1, 50-7.

[13]   Thomas, G. J., Jr.; Prescott, B.; Urry, D. W., *Biopolymers* **1987**, 26, 6, 921-34.

[14]   Reiersen, H.; Clarke, A. R.; Rees, A. R., *J. Mol. Biol.* **1998**, 283, 1, 255-264.

[15]   Li, B.; Alonso, D. O. V.; Daggett, V., *J. Mol. Biol.* **2001**, 305, 3, 581-592.

[16]   Urry, D. W.; Long, M. M.; Harris, R. D.; Prasad, K. U., *Biopolymers* **1986**, 25, 10, 1939-53.

[17]   Urry, D. W.; Walker, J. T.; Rapaka, R. S.; Prasad, K. U., *Polymer Prepr.* **1983**, 24, 1, 3-4.

[18]   Urry, D. W.; Luan, C. H.; Parker, T. M.; Gowda, D. C.; Prasad, K. U.; Reid, M. C.; Safavy, A., *J. Am. Chem. Soc.* **1991**, 113, 11, 4346-8.

[19]   Urry, D. W.; Trapane, T. L.; Long, M. M.; Prasad, K. U., *J. Chem. Soc., Faraday Trans. 1* **1983**, 79, 4, 853-68.

[20]   Budisa, N., *Angew Chem. Int. Ed. Engl.* **2004**, 43, 47, 6426-6463.

[21]   Link, A. J.; Mock, M. L.; Tirrell, D. A., *Curr. Opin. Biotechnol.* **2003**, 14, 6, 603-609.

[22]   Wang, L.; Xie, J.; Schultz, P. G., *Annu. Rev. Biophys. Biomol. Struct.* **2006**, 35, 225-249.

[23]   Deming, T. J.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A., *Polym. Mater. Sci.*

Eng. **1994**, 71, 673-4.

CHAPTER 2

Co-translational Incorporation of a Structurally Diverse Series

of Proline Analogues in an *E. coli* Expression System.

Introduction

The imino acid proline occupies a unique niche in protein structural biology as a consequence of its distinctive structural and conformational properties vis-à-vis the other canonical amino acids.[1] In contrast to the latter, proline has an endocyclic amino group that narrowly limits its accessible conformational space, as well as that of surrounding residues when placed in the context of a polypeptide chain, and prevents participation of the prolyl amide group in the hydrogen bonding interactions that normally act to stabilize protein structure.[2-3] Therefore, proline residues are not easily accommodated within periodic, hydrogen-bonded secondary structure and are often associated with aperiodic features such as turns and loops or non-α/β periodic structures. Nonetheless, proline serves multiple roles of critical importance for protein structure and function,[4] which include delimitation of periodic secondary structure elements,[5] alteration of the polypeptide backbone trajectory through the formation of reverse turns or helical kinks,[6-7] restriction of conformational entropy in the folded and unfolded states,[8] preferential destabilization of mis-folded protein structures,[9] and conformational modulation via the thermodynamically accessible *cis/trans* isomerization of the Xaa-Pro peptide bond.[10] In addition, proline residues are often observed as central features within repetitive

oligopeptide motifs[11] in which the unique conformational features of the imino acid largely determine the structural and functional properties of the repetitive domain. These proline-rich domains usually occur as either intrinsically unstructured elements[12] associated with protein-protein and protein-ligand recognition,[13] or as the structural components of extracellular matrix proteins.[14-15]

Despite the structural importance of proline residues, few experimental methods are available to directly interrogate local conformational effects that arise due to the presence of proline residues in polypeptide sequences on the structure and function of the corresponding native, folded proteins. Directed mutagenesis techniques have been applied routinely to replace proline with other canonical amino acids, however these studies are limited in that the intrinsic structural differences between proline and the other canonical amino acids must be carefully considered during interpretation of the experimental results. As a point of contrast, scanning proline mutagenesis has been employed to identify functionally critical regions of protein secondary structure through the disruptive effect of the proline residues on hydrogen-bonding interactions and local peptide conformation.[16] Recently, chemosynthetic[17] and biosynthetic[18] methods have been described that permit substitution of non-canonical amino acids into full-length, native protein sequences in place of canonical amino acids. In principle, these methods

enable the substitution of proline residues in polypeptide sequences with non-canonical

imino acids that more closely mimic the structural and conformational properties of the

canonical amino acid. These investigations can provide important information on the role

of proline residues within specific structural contexts, particularly under conditions in

which the differences in protein structure and function can be interpreted on the basis of

the often subtle stereoelectronic distinctions that are observed between the canonical

amino acid and its analogues.[19-24] Of the available methods, the biosynthetic approach

holds the most promise for large-scale synthesis of native proteins in which specific

canonical amino acids, or sets of amino acids, have been replaced with structurally

similar amino acid analogues.[18,25] Several bacterial expression systems have been

described that enable global substitution of proline residues in protein sequences with

imino acid analogues,[26-30] however these expression systems have neither been

optimized with respect to efficacy of analogue incorporation nor defined in terms of the

accessible scope of structural analogues that might be accommodated in a conventional

biosynthetic system. We describe herein methods for the multi-site substitution of

canonical proline residues with a structurally diverse set of non-canonical imino acids

that differ substantively in stereoelectronic properties (Scheme 1). We envision that

incorporation of these analogues may be employed to address specific structural

questions regarding the multiple roles of proline residues in native protein sequences

through comparison of the differential effect of the various proline analogues on the

macromolecular structure and function of biosynthetic polypeptides.

Experimental Section

Materials and Methods.

All chemical reagents were purchased from either Fisher Scientific, Inc. (Pittsburgh, PA) or Sigma Chemical Co. (St. Louis, MO) unless otherwise noted. Proline analogues **2**, **3**, **6**, **10**, **11** and **14** were purchased from Bachem Bioscience, Inc. (King of Prussia, PA), **7**, **8** and **12** were obtained from Sigma Chemical Co., and **13** was obtained from Fluka. (2*R,3R*)-3-fluoroproline, **4**, and (2*R,3S)*-3-fluoroproline, **5**, were synthesized from (2*S,3S)*-3-hydroxyproline (Acros Organics, Inc.) using a modification of the method of Demange et al.[46] (2*S*)-4,4-Difluoroproline, **9**, was synthesized from *N*-tert-butoxycarbonyl-(2*S,4S)*-4-hydroxyproline methyl ester (Bachem Bioscience, Inc.) using a modification of the method of Demange, et al.[46] 4-(dimethyl)silaproline, **15**, was synthesized using a modification of the method of Martinez, et al.[31] 2,4-methanoproline, **16**, was synthesized from serine using the method of Clardy, et al.[32] 2-azabicyclo-[2.1.1]hexane, **17**, was a gift from Professor Grant Krow of the Temple University. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Research Products International Corp. (Prospect, IL). Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs, Inc. (Beverly, MA), shrimp alkaline phosphatase

was obtained from Roche Applied Science (Indianapolis, IN), and Platinum *Pfx* DNA polymerase was obtained from Invitrogen Corp (Carlsbad, CA). Plasmid pPROTetE.133 was obtained from BD Biosciences, Inc. (Palo Alto, CA) and plasmids pQE-60 and pQE-80L were obtained from Qiagen, Inc (Valencia, CA). Plasmid pCS-364, which encodes the native *E. coli* prolyl-tRNA synthetase as an *N*-terminal hexahistidine fusion in plasmid pQE-30,[45] was a gift from Professor Karin-Musier Forsyth of the University of Minnesota and plasmid pPROLarA.231 was obtained from Professor Rik Myers of the University of Miami. *E. coli* strain DG99 was purchased from American Type Culture Collection (ATCC# 47041) and strains CAG18515[47] and UQ27[48] were obtained from the *E. coli* Genetic Stock Center at Yale University. *E. coli* strain UMM5[41] was provided by Professor Charles Deutch of Arizona State University. Synthetic oligonucleotides were purchased from either Sigma-Genosys, Inc (The Woodlands, TX) or Integrated DNA Technologies (Coralville, IA) and were used as received. TALON metal affinity resin was purchased from BD Biosciences, Inc. NMM medium was prepared according to the protocol of Budisa, et al,[49] with the exception that proline was not added to the medium prior to cell culture. Procedures for the manipulation of DNA, the transformation of competent cells, and the growth and induction of bacterial cultures were adapted from the published literature[50] or instructions supplied by manufacturers.

Reagents for the manipulation of bacteria and DNA were sterilized by either autoclave or passage through a 0.22 μm filter. Enzymatic reactions were performed in the reagent buffers supplied by the manufacturer. Site-directed mutagenesis was performed using the Quick-Change mutagenesis technique from gene-specific oligonucleotide primers.

Physical and Analytical Measurements.

DNA sequence analyses were performed at the Emory University Core DNA Sequencing Facility on a Perkin-Elmer ABI Prism model 377 DNA sequencer. Amino acid compositional analyses were performed at the W.M. Keck Foundation Biotechnology Resource Laboratory of Yale University. Protein electrophoresis was performed on 10-15% gradient discontinuous SDS polyacrylamide gels on a PhastSystem from Amersham Pharmacia Biotech and visualized via a silver staining procedure.

The molar masses of elastin analogues were determined by MALDI-TOF MS on an Applied Biosystems Voyager System 428 mass spectrometer in the positive linear mode. The matrices, 2-(4-hydroxyphenylazo)benzoic acid (HABA) or 4-hydroxy-3-methoxycinnamic acid, were used at a concentration of 10 mg/ml in a 50:50 mixture of water and 2-propanol. The protein solution (1 mg/ml in distilled water) was mixed with the matrix solution in a ratio of 1:10 and dried under vacuum or air. Bovine serum

albumin was used as a standard for external calibration.

Solution NMR spectra were acquired with either a Varian INOVA 600 instrument (599.742 MHz, ^{1}H) or a Varian UNITY 600 instrument (564.044 MHz, ^{19}F) equipped with a 5-mm ^{1}H/^{19}F probe. Spectra were collected at 4 °C on specimens consisting of 10 mg of protein dissolved in 70:30 sterile H_2O/D_2O in which the pH was adjusted to 2.7 to retard amide proton exchange rates. Chemical shifts for ^{1}H NMR spectra were referenced and reported relative to internal sodium 2,2-dimethyl-2-silapenta-5-sulfonate (0.0 ppm). Standard solvent suppression techniques were employed to reduce signal due to the residual protons of H_2O in the ^{1}H NMR of aqueous solutions. Chemical shifts for the ^{19}F NMR spectra are referenced and reported relative to external sample of aqueous (10 % v/v) trifluoroacetic acid (0.0 ppm).

Plasmid Construction.

The plasmid pRAM2 was employed as a source of the gene encoding the **elastin-1** sequence.[31] Double digestion of pRAM2 with *Nco* I and *Bam*H I afforded a duplex DNA cassette of approximately 1300 bp, which was inserted into the compatible *Nco* I/*Bgl* II sites of plasmid pQE-60 to generate plasmid pAG1. An *Eco*R I/*Hin*D III cassette derived from pAG1 was excised, isolated, and cloned into the compatible sites of plasmid

pQE-80L to afford pAG2, which incorporated a copy of the overproducing repressor

allele *lacI^q* to ensure tight control of the basal level of transcription prior to induction

with IPTG. Plasmid constructs encoding variants of the *E. coli* prolyl-tRNA synthetase

were derived from plasmid pCS-364. The internal *Kpn* I restriction site within the *E. coli*

ProRS gene was removed via site-directed mutagenesis, which introduced a silent

mutation into the coding sequence of the prolyl-tRNA synthetase. Mismatch primers

were employed to introduce the C443G mutation within the activation site of the wild-

type *E. coli* ProRS gene in *Kpn* I-negative mutant of pCS-364. The identity of the

plasmid constructs was confirmed via double-stranded DNA sequence analysis. The

wild-type and mutant ProRS genes were amplified from the plasmid constructs using

polymerase chain reaction in which unique *Kpn* I and *Xba* I restriction sites were

incorporated at the 5'- and 3'-termini, respectively, via gene-specific oligonucleotide

primers. The upstream PCR primer incorporated a new ATG initiation codon and

annealed downstream of the sequence encoding the hexa-histidine tag encoded within the

DNA target to prevent inclusion of this sequence within the PCR products. An acceptor

plasmid for the ProRS genes was constructed from ligation of the *Avr* II/*Spe* I fragment

of pPROTetE.133, containing the transcriptional/translational control elements, multiple

cloning site, and chloramphenicol resistance gene, to the corresponding fragment of

pPROLarA.231, containing the p15A origin of replication. The ProRS genes were cloned

as *Kpn* I/*Xba* I DNA cassettes into the corresponding sites within the polylinker of the

acceptor plasmid pME1 to generate plasmids pWK1 and pWK2, which encoded wild-

type ProRS, and the C443G mutant, respectively, under control of the $P_{L}tet$ promoter.

These synthetic plasmid constructs were screened for prolyl-tRNA synthetase activity via

the ability to rescue a temperature-sensitive ProRS phenotype in *E. coli* strain UQ27. The

plasmids were transformed into competent cells of *E. coli* UQ27 and cultured at 30 °C.

LB solid media (34 μg/mL chloramphenicol) was then streaked with single colonies of

the transformants and identical plates were incubated at permissive (30 °C) and non-

permissive (42 °C) temperatures. The acceptor plasmid pME1 lacking a ProRS gene was

employed as a negative control. Plasmids that displayed the appropriate growth

phenotype were characterized via DNA sequence analysis to ascertain the identity of the

ProRS sequence.

Bacterial Growth and Expression.

The expression vector encoding elastin-mimetic protein sequence, pAG2, was

transformed into *E. coli* strain DG99 or co-transformed with either pWK1 or pWK2 into

E. coli strains CAG18515 or UMM5 to generate the expression strains employed for

these studies. Single colonies of the expression strains were inoculated into sterile LB broth (50 mL) supplemented with the appropriate antibiotics (100 µg/mL ampicillin and 34 µg/mL chloramphenicol) as required for plasmid maintenance. The overnight culture was centrifuged at 4000g for 10 min to isolate the cells, which were re-suspended in sterile NMM medium (1 L) supplemented with the appropriate antibiotics. The proline concentration was adjusted to 0.3 mM from a sterile 100× stock solution. The culture was incubated with agitation (225 rpm) at 37 °C until the OD_{600} reached between 0.8 and 1.0 absorbance units, the cells were collected by centrifugation at 4000g for 10 min. The cell pellet was washed with cold (4 °C), sterile 0.9% aqueous NaCl twice (2 × 100 mL) and resuspended in sterile NMM containing antibiotics but without proline supplementation. After incubation at 37 °C for 30 min, the proline analogues were added to a final concentration of 0.5 mM from sterile 100× stock solutions. For expressions involving proline analogues **7-12**, the osmolarity of the culture was adjusted prior to induction via addition of the appropriate osmolytes to a final concentration of either 600 mM for NaCl or 800 mM for sucrose. An aliquot of aqueous 1.0 M IPTG was added to the cultures to a final concentration of 1 mM to induce synthesis of the elastin-mimetic protein. After a 3 h induction period, the cells were harvested by centrifugation at 4000g and 4 °C for 20 min. The cell pellet was resuspended in lysis buffer (50 mL, 50 mM sodium phosphate,

300 mM NaCl, pH 7.0) and stored at -80 °C.

Protein Purification.

The frozen cells were lysed by three freeze/thaw cycles. Lysozyme (1 mg/mL), EDTA-free protease inhibitor cocktail, benzonase (25 units/mL), and $MgCl_2$ (1 mM) were added to the lysate and the mixture was incubated with shaking at 4 °C overnight. The cell lysate was centrifuged at 40,000g for 30 min at 4 °C. Supernatant and pellet were separated and analyzed by SDS-PAGE to determine the location of the target protein. For soluble elastin-mimetic proteins derived from **1, 2, 3, 4, 7, 8** and **10**, the supernatant was loaded onto cobalt charged TALON resin (5 mL) and washed of lysis buffer (50 mL) containing 20 mM imidazole. The target protein was eluted with elution buffer (20 mL, 50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole, pH 7.0) and dialyzed (MWCO = 10 kDa) against distilled deionized water (5 × 4 L). The dialysate was lyophilized to produce a white spongy solid. For elastin-mimetic polypeptides derived from **5**, **6**, **9**, **11**, and **12**, the insoluble target protein was resuspended in denaturing lysis buffer (100 mL, 50 mM sodium phosphate, 300 mM NaCl, 6 M urea, pH 7.0) and shaken at 4 °C overnight. The resulting mixture was centrifuged at 40,000g and 4 °C for 30 min. SDS-PAGE analysis indicated that the

majority of the target protein dissolved under these conditions. The soluble fraction was

loaded onto cobalt-charged TALON (10 mL) resin that had been previously equilibrated

with denaturing lysis buffer. The target protein was washed with denaturing lysis buffer

(100 mL) containing 20 mM imidazole and eluted with denaturing lysis buffer (40 mL)

containing 250 mM imidazole. The eluted target protein was dialyzed (MWCO = 10 kDa)

against a diminishing urea step gradient from 6 M to 1 M, and, subsequently, against

distilled water (5 × 4 L). Lyophilization of the dialysate produced the elastin-mimetic

polypeptides as white spongy solids.

Results and Discussion

The elastin-mimetic polypeptide, **elastin-1** (Scheme 1), was employed as a test

substrate for assessing the efficacy of multi-site incorporation of a wide range of proline

analogues using a high-level, IPTG-inducible expression system.[33] The repetitive

domain of **elastin-1** comprises a concatenated series of pentapeptide repeats based on the

consensus sequence (Val-Pro-Gly-Val/Ile-Gly) in which proline residues constitute 20 %

of the amino acids within the repetitive domain (80 prolines/polypeptide chain). The high

density of proline residues and the uniform structural environment of these residues

within the polypeptide sequence should provide a well-defined context for evaluating the

efficacy of analogue incorporation in terms of its effect on protein yield and analogue

substitution level. Moreover, biophysical studies of elastin and elastin-mimetic

polypeptides have suggested a crucial structural role for the proline residues within the

pentapeptide repeats that may be essential for the normal physiological function of the

native polypeptide material.[34-35] Thus, substitution of proline analogues into the

repetitive domain of **elastin-1** may provide critical information on the local structural

parameters that influence the thermodynamics and kinetics of elastin assembly under

physiologically relevant conditions. The construction of the synthetic gene encoding the

A. $MGH_{10}S_2GHID_4KHM\ [(VPGVG)_4VPGIG]_{16}V$

B.
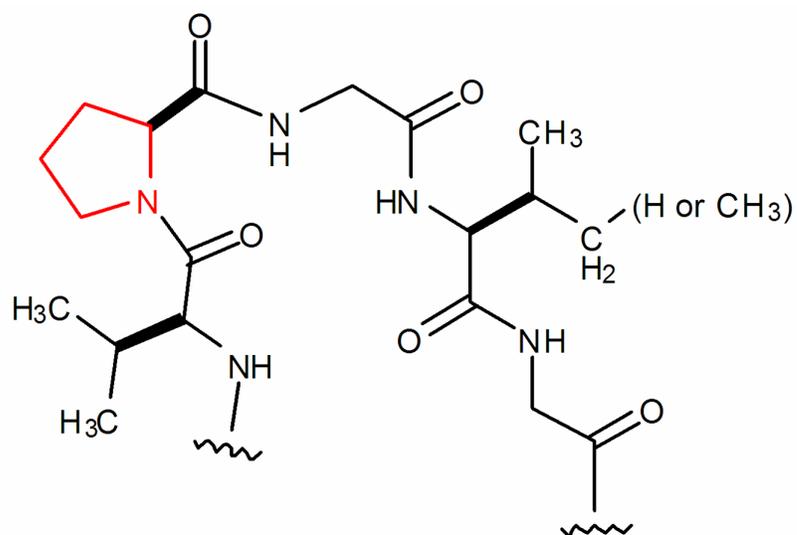
Scheme 1. (A) Complete amino acid sequence of **elastin-1** and **(B)** chemical structure of the pentapeptide repeat unit depicting proline residue as a red color.

Strain	Genotype	Comments
DG99	F⁻, *thi*-1, *endA*1, *supE44*, *hsdR*17, *proC*::Tn10, *lacI^q*, ˌ *lacZ)M15*	
CAG18515	F⁻, *proA*3096::Tn10Kan, *rph*-1	**Reference 49**
UMM5	*putA*1::Tn5, *proC*24, *metB*1, *relA*1, *spoT*1, *bglF*18::IS150	**Reference 43**
UQ27	*proS*127(ts), *lacZ*4,lam⁻, *argG*75	**Reference 50**
Plasmids		
pRAM2	pET-19b/**elastin-1**	**Reference 33**
pAG1	pQE60/**elastin-1**	
pAG2	pQE80L/**elastin-1**	*lacI^q*
pME1	pPROTetE.133/pPROLarA.231	**Cm^R, p15A ori**
pCS-364	pQE30/*proS*	**Reference 47**
pWK1	pME1/*proS*	
pWK2	pME1/*proS*(C443G)	

Table 1. *E. coli* strains and plasmids.

elastin-1 sequence has been previously reported in a pET-19b-derived expression system.[33] In order to maintain tighter control over the basal level of gene expression, this DNA cassette was sub-cloned into a pQE-derived expression vector, pAG2, under the control of a phage T5 promoter that was inducible with IPTG (Table 1). A copy of the over-producing lactose repressor allele *lacI^q* was incorporated into the plasmid to maintain transcriptional silencing of the target sequence prior to induction. The decahistidine tag was retained from the original construct to facilitate purification of the target protein from the endogenous host proteins via immobilized metal affinity chromatography.

The facility with which amino acid analogues are incorporated into native proteins in place of a canonical amino acid often depends on the degree of structural similarity between the two molecular species.[18] Native aminoacyl-tRNA synthetases must discriminate between closely related canonical amino acids and have evolved selective editing mechanisms that operate at both the pre-transfer and post-transfer stages to ensure high fidelity aminoacylation of tRNA substrates with the cognate amino acid.[36-37] However, structurally similar amino acids can often be incorporated into proteins in place of a canonical amino acid under selective pressure in which the host bacterium is depleted of the canonical amino acid and supplemented with the analogue.
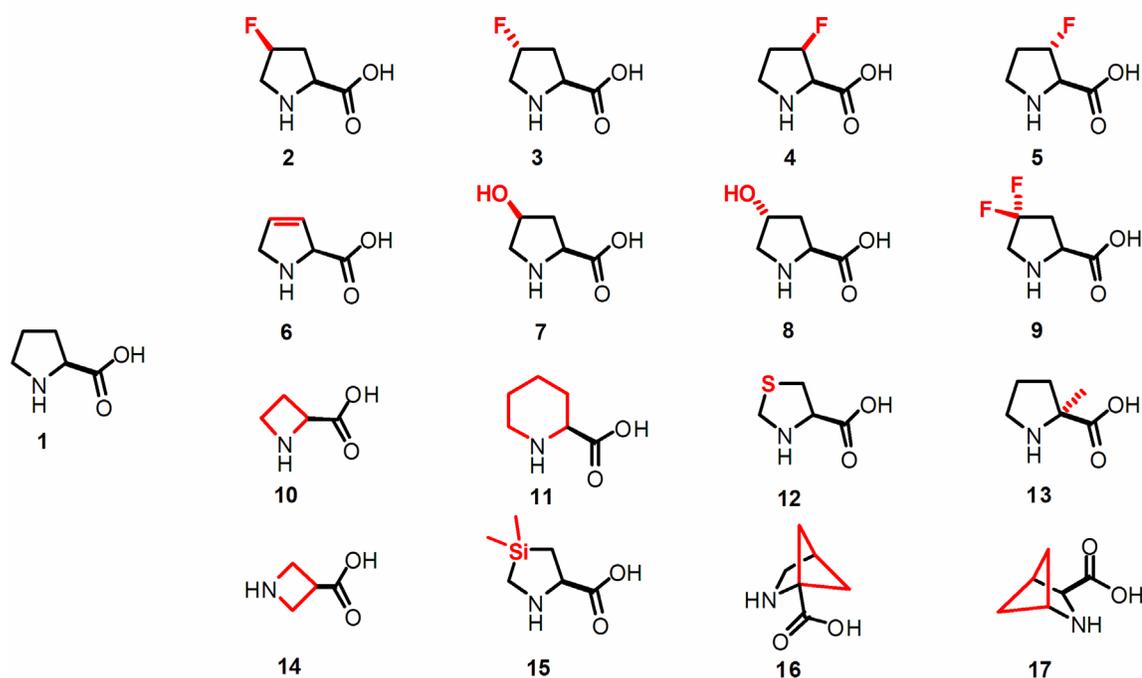
Scheme 2. Chemical structures of proline (**1**) and the imino acid analogues (**2-17**) that were tested for use as substrates for biosynthetic incorporation into the elastin-mimetic protein sequence, **elastin-1**: **1**, (*2S*)-proline; **2**, (*2S,4S*)-4-fluoroproline; **3**, (*2S,4R*)-4-fluoroproline; **4**, (2R,3R)-3-fluoroproline; **5**, (2R,3S)-3-fluoroproline; **6**, (*2S*)-3,4-dehydroproline; **7**, (*2S,4S*)-4-hydroxyproline; 8, (*2S,4R*)-4-hydroxyproline; **9**, (*2S*)-4,4-difluoroproline; **10**, (*2S*)-azetidine-2-carboxylic acid; **11**, (*2S*)-piperidine-2-carboxylic acid; **12**, (*4R*)-1,3-thiazolidine-4-carboxylic acid; **13**, 2-methylproline; **14**, (3S)-azetidine-3-carboxylic acid; **15**, 4-(dimethyl)silaproline; **16**, 2,4-methanoproline; and **17**, 2-azabicyclo-[2.1.1]hexane.

As the structural divergence increases between a cognate amino acid and an analogue, the efficiency of incorporation of the analogue decreases as judged by reduced protein yield and/or appearance of products associated with amino acid starvation. Given the structural diversity of the proline analogues in scheme 2, a single technique is unlikely to suffice for incorporation of the entire set of analogues. Three different methods were explored to assess the efficacy of incorporation of various proline analogues into the target polypeptide, in which the host cell physiology was altered to an increasing degree to accommodate more structurally divergent proline analogues. The result of this process was the definition of conditions for the biosynthesis of a series of elastin derivatives, **elastin-2** through **elastin-12**, based on substitution of the canonical proline residues encoded in the **elastin-1** sequence with the respective imino acid analogues **2-12** (Scheme 2). The simplest experimental method employed the auxotrophic *E. coli* strain DG99 (*proC*::Tn10, Tc^R) as a host organism under conditions of proline depletion and analogue supplementation.[26-30] This approach worked well for incorporation of the mono-4-fluoroprolines, **2**, **3**, and **4** and 3,4-dehydroproline, **6**, as the native genomic background activity of prolyl-tRNA synthetase within the host bacterium was probably sufficient to recognize and charge these structurally similar proline analogues to the tRNA^{Pro} at a level of efficiency approaching that of canonical amino acid (Figure 1a). In contrast,

structurally less similar analogues, i.e., most of the imino acids in Scheme 2, were not readily incorporated into the elastin sequence via this approach.

An alternative method that has been successfully employed for the multi-site substitution of weakly recognized amino acid analogues in response to sense codons involves the co-expression of the wild-type aminoacyl-tRNA synthetase,[38] or in particularly difficult cases, a mutant synthetase that displays more permissive substrate selectivity.[39] Both of these strategies were investigated to facilitate incorporation of the more recalcitrant proline analogues into the target protein substrate. An expression vector was constructed in which variants of *E. coli* prolyl-tRNA synthetase could be expressed under the control of an orthogonal promoter system. This vector system was derived from the modular series of bacterial plasmids reported initially by Lutz and Bujard.[40] Plasmid pME1 contained the transcriptional/translational control elements, multiple cloning site, and chloramphenicol resistance marker of plasmid pPROTetE.133 and the compatible p15A replicon of plasmid pPROLarA.231 (Table 1). Duplex DNA cassettes encoding variants of *E. coli* prolyl-tRNA synthetase were amplified via PCR and cloned as *Kpn* I/*Xba* I fragment into the polylinker of pME1. The ProRS expression cassettes were placed under the control of the P_Ltet promoter and were expressed constitutively from the respective plasmids in *E. coli* strains that lacked the *tet* repressor. The incorporation of
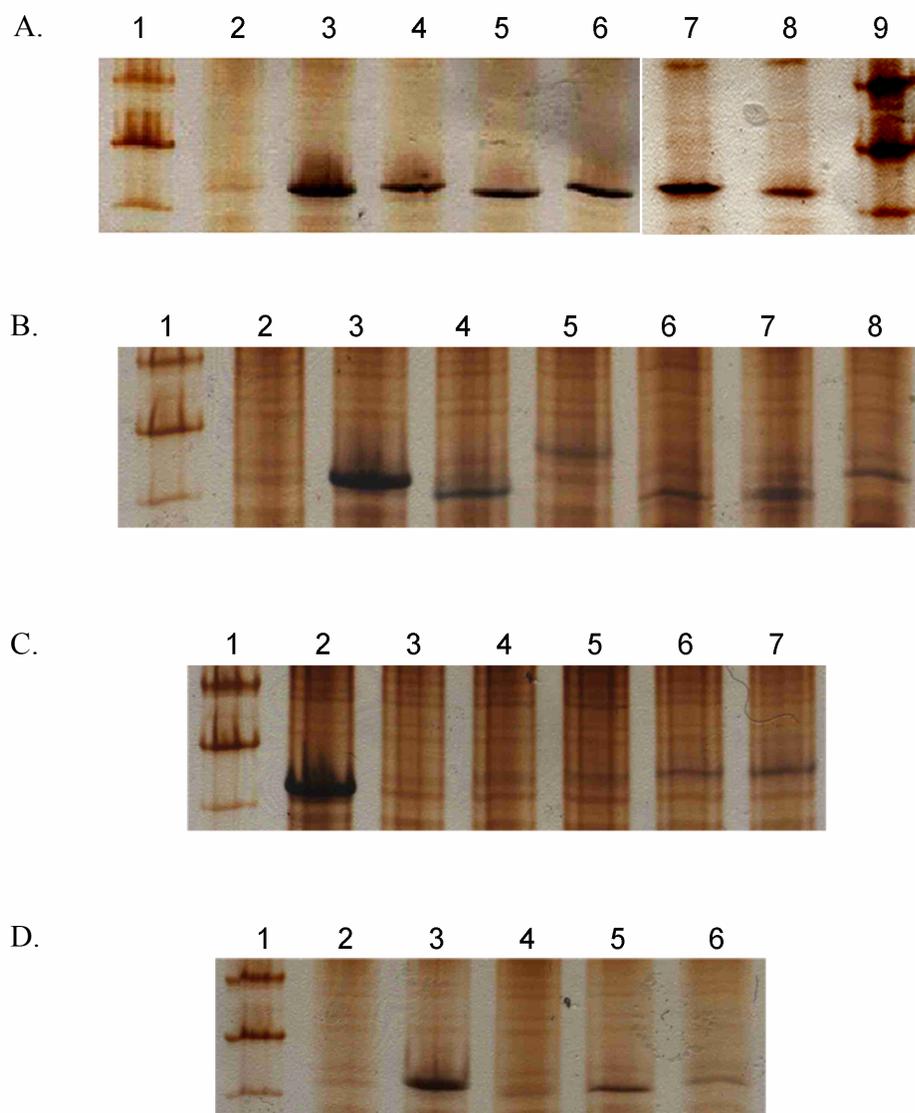
Figure 1. SDS-PAGE analysis of whole cell lysates derived from expression cultures for production of elastin analogues in the presence of proline derivatives **1-12**. (**A**) Incorporation of proline analogues using the genomic background activity of prolyl-tRNA synthetase. Expression cultures were derived from *E. coli* strains DG99[pAG2] and CAG18515[pAG2][pWK1] after 3 h induction with 1 mM IPTG: lane 1 and 9, molecular weight standards (25 kDa, 35 kDa, and 50 kDa); lane 2, proline deficient (negative) control; lane 3, proline supplemented (positive) control; lane 4, (2*S*,4*S*)-4-fluoroproline (**2**); lane 5, (2*S*,4*R*)-4-fluoroproline (**3**); lane 6, (2*S*)-3,4-dehydroproline (**6**);

lane 7, (2*R*,3*R*)-3-fluoroproline (**4**); and lane 8, (2*R*,3*S*)-3-fluoroproline (**5**). (**B**) Incorporation of proline analogues under conditions of co-expression of wild-type prolyl-tRNA synthetase. Expression cultures were derived from *E. coli* CAG18515[pAG2][pWK1] after 3 h induction with 1 mM IPTG in hyperosmotic (600 mM NaCl) media. Lane 1, molecular weight standards; lane 2, negative control; lane 3, positive control; lane 4, (2*S*,4*S*)-4-hydroxyproline (**7**); lane 5, (2*S*,4*R*)-4-hydroxyproline (**8**); lane 6 (2*S*)-4,4-difluoroproline (**9**); lane 7, (4*R*)-1,3-thiazolidine-4-carboxylic acid (**12**); and lane 8, (2*S*)-azetidine-2-carboxylic acid (**10**). Note that the expression culture for analogue **12** (Lane 7) was performed using *E. coli* strain UMM5[pAG2][pWK1] under 3 h induction in media supplemented 800 mM sucrose rather than 600 mM NaCl. In addition, the incorporation of proline analogues **8** and **10** into the target polypeptide resulted in noticeable shifts in the gel mobility as detected via SDS-PAGE analysis (*cf.* lanes 5 and 8). A similar correspondence between hydroxyproline substitution level and decreased electrophoretic mobility has been observed previously for recombinant glutathione-*S*-transferase expressed under conditions of hydroxyproline supplementation in the auxotrophic *E. coli* strain JM108 (Δ(*lac-pro*AB)).[26] (**C**) Dependence of target protein accumulation on osmolyte (NaCl) concentration for expression cultures derived from CAG18515[pAG2][pWK1] after 3 h induction with 1 mM IPTG. Lane 1, molecular weight standards; lane 2, positive control; and lanes 3-7; expression in the presence of 0.5 mM (2*S*,4*R*)-4-hydroxyproline (**8**) under conditions of increasing hyperosmolarity in culture medium supplemented to final concentrations of 8.5 mM, 100 mM, 200 mM, 400 mM, and 600 mM NaCl, respectively. (**D**) Effect of wild-type and mutant prolyl-tRNA synthetase expression on proline analogue incorporation. Expression cultures were derived from *E. coli* strain CAG18515[pAG2] expression cultures after 3 h induction with 1 mM IPTG. Lane 1, molecular weight standards; lane 2, negative control, lane 3, positive control; lane 4, supplementation with 0.5 mM (2*S*)-piperidine-2-carboxylic acid (**11**) in 600 mM NaCl medium with co-expression of wild-type ProRS from pWK1; lane 5, supplementation with 0.5 mM (2*S*)-piperidine-2-carboxylic acid (**11**) in 600 mM NaCl medium with co-expression of the C443G mutant ProRS from pWK2.

proline analogues into the **elastin-1** sequence was assayed in the auxotrophic *E. coli* strain CAG18515 *(proA*::Tn10Kan, Kn[R]*)* under conditions in which recombinant wild-type ProRS was constitutively expressed from plasmid pWK1 (Table 1). Electrophoretic analysis of whole cell lysates derived from these expression cultures indicated that most of the proline analogues, with the exception of **11** and **12**, displayed a detectable level of target protein expression with respect to the proline-deficient negative control (Figure 1b). Moreover, hyperosmotic concentrations of osmolytes such as sucrose or sodium chloride in the culture media enhanced the level of accumulation of the target polypeptides in concentration-dependent manner (Figure 1c). Hyperosmotic expression cultures have been employed previously to facilitate co-translational incorporation of (*2S,4R*)-4-hydroxyproline, **8**, into test proteins in response to proline codons in an *E. coli* host system.[26] This approach relies on the hyperosmotically induced up-regulation of the biosynthesis of endogenous low affinity proline transporters (the *putP*, *proP*, and *proU* gene products) in *E. coli*.[41] These proteins can recognize and transport proline analogues into the bacterium, enhancing their cellular concentration and, thus, their functional efficacy as substrates for endogenous prolyl-tRNA synthetase.[26]

However, neither **11** nor **12** could be incorporated into elastin-mimetic polypeptide substrate using expression strain CAG18515[pWK1] even under conditions

of hyperosmolarity. The inability of the latter substrate to be incorporated into **elastin-1**

is particularly puzzling as co-translational incorporation of **12** into proteins has been

reported for *in vivo*[28] and *in vitro*[42] bacterial expression systems derived from *E. coli*.

Recent experimental evidence has suggested that several proline analogues, particularly **6**

and **12**, are subject to intracellular oxidative degradation through the action of

endogenous *E. coli* enzymes such as *L*-proline dehydrogenase (*putA*) and Δ^1-pyrroline-5-

carboxylate reductase (*proC*).[43] The co-translational incorporation of **12** into the elastin

sequence was explored in *E. coli* strain UMM5, which contains genetic mutations in the

chromosomal loci associated with these two enzymes that impair their activity towards

oxidative degradation of **6** and **12** (Table 1). Despite the use of this mutant strain,

expression of **elastin-12** could only be detected under conditions in which the wild-type

ProRS was co-expressed from plasmid pWK1 in hyperosmotic media, with the best

yields of isolated protein resulting from culture in media supplemented with 800 mM

sucrose (Figure 1b). However, the isolated yield of **elastin-12** under these conditions was

comparable to that of a thiaproline-substituted analogue of recombinant human annexin

V,[28] which contained a much lower content of proline residues than did the elastin test

substrate. In addition, *E. coli* strain UMM5 improved the isolated yield of the **elastin-6**

derivative by approximately an order of magnitude in comparison to the *E. coli* strain

DG99 even though the level of target protein accumulation appeared to be comparable between the two strains as judged via SDS-PAGE analysis of whole cell lysates from expression cultures conducted under similar conditions. These observations emphasized the important effect of host cell physiology on target protein accumulation, particularly under conditions in which chemically reactive amino acid analogues are employed as substrates for *in vivo* ribosomal protein synthesis.

In contrast to the other analogues, (*2S*)-piperidine-2-carboxylic acid, **11**, could not be incorporated into **elastin-1** using any of the auxotrophic strains even with co-expression of the wild-type ProRS under hyperosmotic conditions. Previous reports indicated that **11** accumulated intracellularly in *E. coli* under hyperosmotic conditions as a consequence of the action of the osmotically regulated proline transporters derived from expression of *proP* and *proU*.[44] However, analogue **11** did not act as an effective substrate to activate ATP-PP$_i$ exchange in the presence of *E. coli* ProRS in an *in vitro* assay,[42c] nor did the analogue inhibit ATP-PP$_i$ exchange of *E. coli* ProRS in the presence of the cognate amino acid, proline. These data suggested that this analogue could not serve as an effective substrate for the aminoacylation reaction, which correlated with the inability of **11** to support the biosynthesis of an **elastin-11** derivative from *in vivo* expression systems that utilized wild-type *E. coli* ProRS activity. We hypothesized that

the larger, six-membered piperidinyl ring of **11** could not be accommodated within the activation site of native *E. coli* ProRS. If this were the situation, then analogue **11** would be effectively precluded from engaging in the aminoacylation reaction with tRNA^{Pro} isoacceptors, and, therefore, could not be co-translationally incorporated into the target polypeptide in response to proline codons.

A representative steric model for the environment surrounding proline in the activation site of *E. coli* prolyl-tRNA synthetase can be estimated from the crystal structures of the corresponding ProRS enzymes from *T. thermophilus* (PDB: 1H4T) and *M. thermautotrophicus* (PDB: 1NJ5).[45-46] Both structures have been determined with proline or proline-derived substrate analogues bound in the activation site, and, in both cases, relatively close contacts are observed between the *exo* face of the proline ring and the terminal heavy atom of a side chain corresponding to amino acids that resides at a homologous position within the activation site of the prolyl-tRNA synthetases, i.e., Ser288 of *T. thermophilus* ProRS and Cys265 of *M. thermautotrophicus* ProRS. The positioning of this residue may inhibit the ability of the prolyl-tRNA synthetase to accommodate more sterically demanding proline analogues such as **11** within the activation site. The homologous amino acid within the sequence of *E. coli* ProRS (Cys443) has been implicated as a structurally critical residue within the activation site as

chemical derivatization of the sulfhydryl side chain abrogates aminoacylation activity.[47]

However, site-directed mutagenesis experiments indicated that this cysteine residue may

be substituted by glycine, alanine, or serine with relatively minor attenuation of the

aminoacylation activity. In particular, the C443G mutant of *E. coli* prolyl-tRNA

synthetase maintained near-native levels of aminoacylation activity ($k_{cat}/K_M \approx 26$ % of

wild-type ProRS),[47] although the available volume within the activation site should be

significantly increased in the mutant as a consequence of the large difference in steric

demand between the Cα-substituents of glycine and cysteine.

The C443G mutant of *E. col*i prolyl-tRNA synthetase was prepared via site-

directed mutagenesis and cloned into the expression plasmid pME1 (Table 1). The ability

of analogue **11** to support protein biosynthesis of **elastin-11** was examined in *E. coli*

strain CAG18515 under conditions in which wild-type and mutant ProRS variants were

expressed from the plasmids pWK1 and pWK2, respectively (Figure 1d). Electrophoretic

analysis of whole cell lysates from these expression cultures indicated the accumulation

of a new protein that migrated at the expected molar mass versus the molecular weight

standards, but only under conditions in which the mutant ProRS was co-expressed within

the bacterial culture. To our knowledge, this experiment represents the first example of a

bacterial expression system that is competent for incorporation of the ring-expanded

proline analogue **11**. The expression level of the **elastin-11** derivative in the presence of

analogue **11** approaches that observed for other proline analogues within the series and

suggests that the substitution process is relatively efficient despite the high density of

encoded proline sites in the coding sequence of **elastin-1**. Hyperosmotic conditions (600

mM NaCl) within the expression media provided the highest accumulation levels for the

target protein, **elastin-11**. This result coincided with the previously reported observation

that the intracellular transport and accumulation of **11** was facilitated in the presence of

high concentrations of extracellular osmolytes.

Elastin derivatives **elastin-1** through **elastin-12** could be purified to homogeneity

from the endogenous proteins of the bacterial host using immobilized metal affinity

chromatography. The isolated yields of the polypeptides (Table 2) ranged from

approximately 15 mg/L to 50 mg/L for fully induced expression cultures in modified

minimal medium (NMM) under conventional batch fermentation conditions in shake

flask culture. The protein yields are quite respectable especially in consideration of the

high density of proline residues in the target polypeptide sequence. These synthetic

elastin derivatives represent among the highest effective concentration of amino acid

analogues that have been incorporated into a test substrate using a biosynthetic approach.

Since the average frequency of occurrence of proline residues in a typical protein

sequence is significantly lower than that encoded within the amino acid sequence of

elastin-1, the methods described herein should be sufficient to effect global substitution

of proline residues with analogues **2-12** in high expressed protein yield for more

conventional polypeptides. Amino acid compositional analysis (Table 2) and MALDI-

TOF mass spectrometry (Table 3 and Figure 2) of the elastin derivatives suggested

virtually complete substitution of proline with the respective imino acid analogue. Little

residual proline content was detected in the recombinant target proteins **elastin-2** through

elastin-12, and the molecular ions within the mass spectra corresponded well with the

calculated masses for elastin derivatives in which the encoded proline residues were

completely substituted with the respective proline analogues (Table 3). Spectroscopic

analyses of selected elastin derivatives demonstrated structural features that were

commensurate with high levels of incorporation of the respective imino acid analogues

(Figure 3) in comparison to the canonical amino acid within the same structural context.

Elastin Derivative	Mass Yield (mg/L)	% Proline Analogue
Theoretical	–	18.9
Elastin-1[1]	46.0	19.9
Elastin-2[1]	50.8	21.6
Elastin-3[1]	45.0	21.4
Elastin-4[2]	53.5	19.6
Elastin-5[2]	48.7	n/a
Elastin-6[3]	34.4	18.6
Elastin-7[4]	41.0	21.3
Elastin-8[4]	26.0	21.3
Elastin-9[4]	15.2	23.1
Elastin-10[4]	16.5	19.4[7]
Elastin-11[5]	16.0	13.3[8]
Elastin-12[6]	27.3	20.1

Table 2. Isolated yields and proline analogue incorporation levels for purified elastin derivatives under optimized expression conditions for cotranslational incorporation of imino acid analogues **1-12**.

[1]Expression from *E. coli* strain DG99[pAG2]. [2]Expression from *E. coli* strain CAG18515[pAG2][pWK1]. [3]Expression from *E. coli* strain UMM5[pAG2]. [4]Expression from *E. coli* strain CAG18515[pAG2][pWK1] at 600 mM NaCl. [5]Expression from *E. coli* strain UMM5[pAG2][pWK1] at 800 mM sucrose. [6]Expression from *E. coli* strain CAG18515[pAG2][pWK2] at 600 mM NaCl. [7]The computed value for the proline analogue (2S)-azetidine-2-carboxylic acid reflects the sum of peaks in the chromatogram that resulted from hydrolytic cleavage of the analogue under the experimental conditions employed in the amino acid analysis. The major product eluted at the position expected for homo-serine, which could have arisen from hydrolytic cleavage of the azetidine ring under the acidic conditions employed in the amino acid analysis. [8]The peak in the chromatogram associated with (2S)-piperidine-2-carboxylic acid could not be effectively integrated due to elution during the buffer change between methionine and valine. The residual proline content (0.12 %) could be computed from the chromatogram as an effective upper limit on the canonical amino acid substitution, which suggested that the majority of the encoded proline sites were occupied by the amino acid analogue.

Elastin Derivative	Calculated m/z[1]	Observed m/z	│Δm/z│ (% error)
Elastin-1	35866.36	35851.85	14.51 (0.04)
Elastin-2	37305.60	37224.26	81.34 (0.22)
Elastin-3	37305.60	37225.46	80.14 (0.21)
Elastin-4	37305.60	37286.30	19.30 (0.05)
Elastin-5	37305.60	37231.90	73.70 (0.20)
Elastin-6	35705.10	35726.39	21.29 (0.06)
Elastin-7	37146.31	37121.18	25.13 (0.07)
Elastin-8	37146.31	37115.23	31.08 (0.08)
Elastin-9	38744.84	38690.87	53.97 (0.14)
Elastin-10	34744.22	34716.04	28.18 (0.08)
Elastin-11	36988.50	36950.08	38.42 (0.10)
Elastin-12	37309.02	37351.87	42.85 (0.11)

Table 3. MALDI-TOF mass spectrometric data for purified elastin derivatives incorporating proline analogues **1-12**.

[1]Molar masses for the elastin derivatives were calculated based on complete substitution of proline with the corresponding amino acid analogue. The calculated molar mass assumes proteolytic cleavage of the *N*-terminal methionine residue as a consequence of the endogenous activity of *E. coli* methionyl-aminopeptidase.[53]
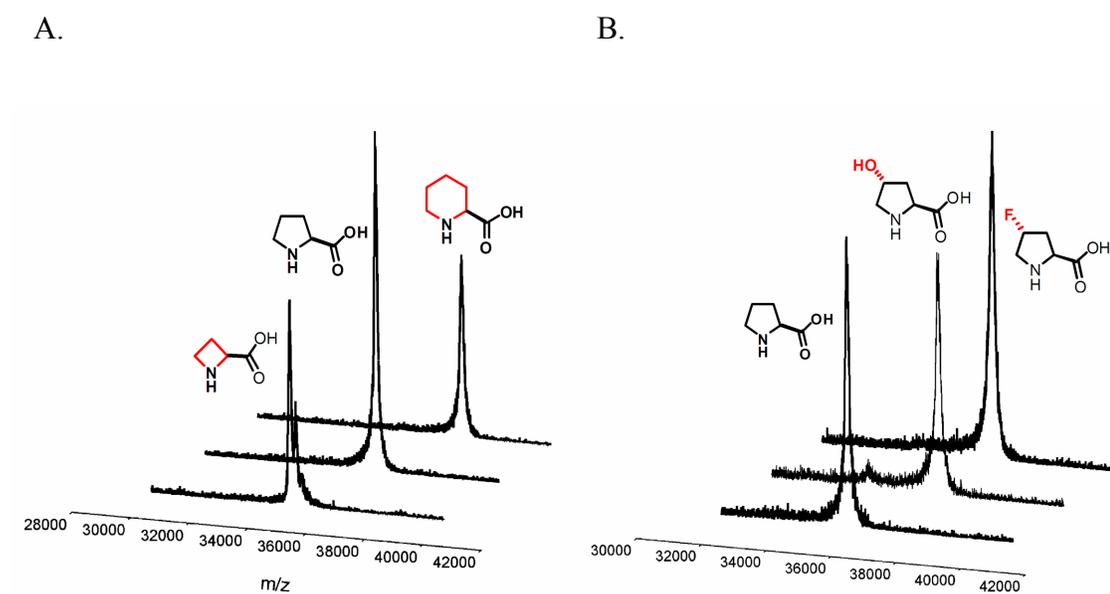
Figure 2. Comparative MALDI-TOF mass spectrometric analysis of selected elastin derivatives. (**A**) MALDI-TOF mass spectra of elastin polypeptides incorporating contracted and expanded ring analogues in place of the canonical proline residues in the repetitive pentapeptide domains. (**B**) MALDI-TOF mass spectra of elastin polypeptides incorporating functionalized analogues of proline in place of the canonical proline residues in the repetitive polypeptide domains. Note that the observed differences in molar mass between the elastin derivatives are commensurate with a high level of substitution of the canonical amino acid with the corresponding analogue.
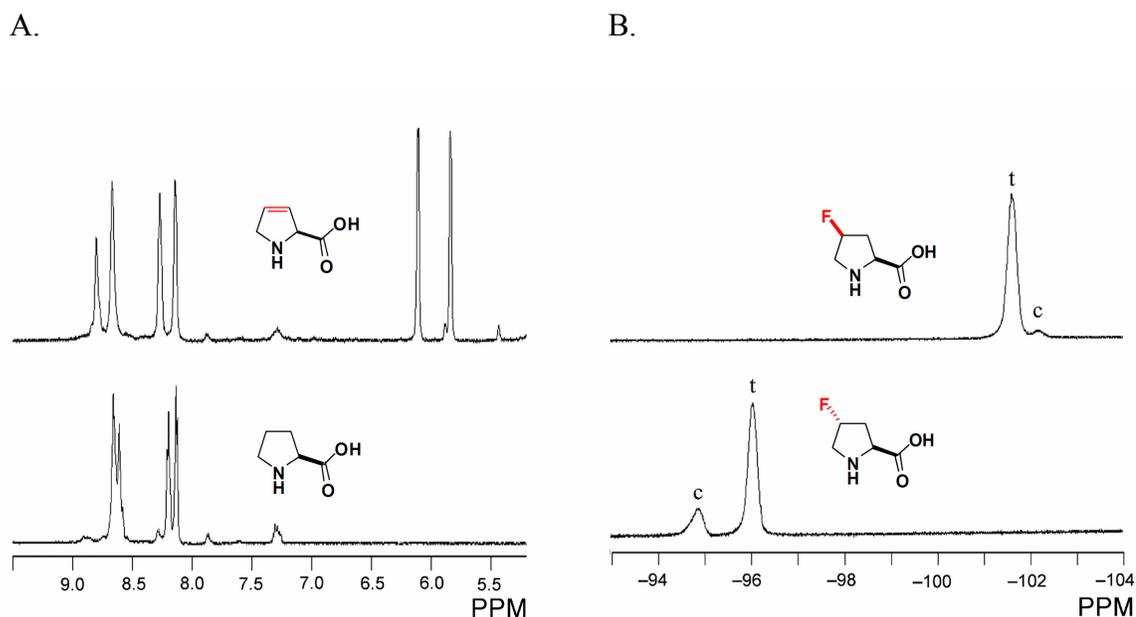
Figure 3. NMR spectroscopic analyses of selected elastin derivatives. (**A**) ^{1}H NMR spectra of **elastin-1** (lower spectrum) and **elastin-6** (upper spectrum) in the downfield region encompassing the amide proton and olefinic proton resonances. Note the presence of the olefinic protons in the upper spectrum that are consistent with incorporation of (2*S*)-3,4-dehydroproline at a level commensurate with that of the amide protons of the repeat unit. The minor resonances within the spectrum can be attributed to the presence of the *cis* configuration of the Val-Pro peptide bond. (**B**) ^{19}F NMR spectra of **elastin-2** (lower spectrum) and **elastin-3** (upper spectrum) indicating the incorporation of (2*S*,4*S*)-4-fluoroproline and (2*S*,4*R*)-4-fluoroproline, respectively, into the recombinant polypeptide. The symbols **t** and **c** refer to the *trans* and *cis* isomers, respectively, of the Val-Pro peptide bonds within the repeat units. These spectroscopic assignments were based on comparison with ^{19}F NMR chemical shift values for 4-fluoroproline model compounds as described in the literature.[19,48] The greater relative abundance of the cis peptidyl bond isomer in **elastin-2** versus **elastin-3** reflects the increased thermodynamic preference for this configuration in (2*S*,4*S*)-4-fluoroproline derivatives.[19,27]

Conclusion

In summary, we have defined a series of *E. coli* expression strains that are competent for the incorporation of a structurally diverse set of proline analogues and developed culture conditions that are compatible with high levels of analogue substitution within a test protein substrate. These bacterial strains may be employed to assay the efficacy of incorporation of novel proline analogues into recombinant proteins or to create variant polypeptides in which native protein sequences have been globally substituted with imino acid analogues in response to proline codons. Moreover, these methods may be used to interrogate the effect of imino acid substitution on protein structure and function and may be particularly informative in the context of structural comparison of a series of imino acid-substituted proteins with respect to the stereoelectronic differences between the incorporated proline analogues. Preliminary studies of the elastin derivatives have indicated that the incorporation of imino acid analogues into the polypeptide can have quite dramatic effects on the thermodynamic and conformational properties associated with the self-assembly of elastin-mimetic polypeptides.

References

[1]   H. Reiersen, A.R. Rees, *Trends Biochem Sci*. **2001**, *26*, 679-684.

[2]   M.W. MacArthur, J.M. Thornton, *J. Mol. Biol.* **1991**, *218*, 397-412.

[3]   U. Reimer, G. Scherer, M. Drewello, S. Kruber, M. Schutkowski, G. Fischer, *J. Mol. Biol.* **1998**, *279*, 449-460.

[4]   S.J. Eyles, L.M. Gierasch, *J. Mol. Biol.* **2000**, *301*, 737-747.

[5]   a) A.R.Viguera, L. Serrano, *Protein Sci.* **1999**, *8*, 1733-1742. b) J. Prieto, L. Serrano, *J. Mol. Biol.* **1997**, *274*, 276-288. c) K. Gunasekaran, H.A. Nagarajaram, C. Ramakrishnan, P. Balaram, *J. Mol. Biol*. **1998**, *275*, 917-932.

[6]   a) S.S. Zimmerman, H.A. Scheraga, *Biopolymers* **1977**, *16*, 811-843. b) C.M. Wilmot, J.M. Thornton, *J. Mol. Biol*. **1988**, *203*, 221-232. c) E.G. Hutchinson, J.M. Thornton, *Protein Sci*. **1994**, *3*, 207-221.

[7]   F.S. Cordes, J.N. Bright, M.S. Sansom, *J. Mol. Biol*. **2002**, *323*, 951-960.

[8]   Y. Xu, T. Hyde, X. Wang, M. Bhate, B. Brodsky, J. Baum, *Biochemistry* **2003**, *42*, 8696-8703.

[9]   W.C. Wigley, M.J. Corboy, T.D. Cutler, P.H. Thibodeau, J. Oldan, M.G. Lee, J. Rizo, J.F. Hunt, P.J. Thomas, *Nat. Struct. Biol*. **2002**, *9*, 381-388.

[10]   a) W.J. Wedemeyer, E. Welker, H.A. Scheraga, *Biochemistry*. **2002**, *41*, 14637-14644. b) A.H. Andreotti, *Biochemistry*. **2003**, *42*, 9515-9524.

[11]   M.P. Williamson, *Biochem. J*. **1994**, *297* (Pt 2), 249-260.

[12]   B.K. Kay, M.P. Williamson, M. Sudol, *FASEB J*. **2000**, *14*, 231-241.

[13]   a) P. Tompa, *Bioessays* **2003**, *25*, 847-55. b) P. Tompa, *Trends Biochem. Sci*. **2002**, *27*, 527-533.

[14]   B. Brodsky, J.A. Ramshaw, *Matrix Biol*. **1997**, *15*, 545-554.

[15]   J. Rosenbloom, W.R. Abrams, R. Mecham, *FASEB J*. **1993**, *7*, 1208-1218.

[16]   a) A.D. Williams, E. Portelius, I. Kheterpal, J.T. Guo, K.D. Cook, Y. Xu, R. Wetzel, *J. Mol. Biol*. **2004**, *335*, 833-842. b) K. Sugase, Y. Oyama, K. Kitano, H. Akutsu, M. Ishiguro, *Bioorg. Med. Chem. Lett*. **2002**, *12*, 1245-1247.

[17]   T.W. Muir, *Annu. Rev. Biochem*. **2003**, *72*, 249-289.

[18]   a) A.J. Link, M.L. Mock, D.A. Tirrell, *Curr. Opin. Biotechnol*. **2003**, *14*, 603-609. b) L. Wang, P.G. Schultz, *Chem. Commun*. **2002**, 1-11. c) N. Budisa, C. Minks, S. Alefelder, W. Wenger, F. Dong, L. Moroder, R. Huber, *FASEB J*. **1999**, *13*, 41-51.

[19]   a) J.A. Hodges, R.T. Raines, *J. Am. Chem. Soc*. **2003**, *125*, 9262-9263. b) M.L. DeRider, S.J. Wilkens, M.J. Waddell, L.E. Bretscher, F. Weinhold, R.T. Raines,

Markley, J.L. *J. Am. Chem. Soc*. **2002**, *124*, 2497-2505. c) L.E. Bretscher, C.L. Jenkins, K.M. Taylor, M.L. DeRider, R.T. Raines, *J. Am. Chem. Soc*. **2001**, *123*, 777-778. d) S.K. Holmgren, L.E. Bretscher, K.M. Taylor, R.T. Raines, *Chem. Biol*. **1999**, *6*, 63-70. e) S.K. Holmgren, K.M. Taylor, L.E. Bretscher, R.T. Raines, *Nature*. **1998**, *392*, 666-667. f) E.S. Eberhardt, N. Panasik, Jr., R.T. Raines, *J. Am. Chem. Soc*. **1996**, *118*, 12261-12266. g) N. Panasik, Jr., E.S. Eberhardt, A.S. Edison, D.R. Powell, R.T. Raines, *Int. J. Pept. Protein Res*. **1994**, *44*, 262-269.

[20]   D. Barth, A.G. Milbradt, C. Renner, L. Moroder, *Chembiochem*. **2004**, *5*, 79-86.

[21]   A.V. Persikov, J.A. Ramshaw, A. Kirkpatrick, B. Brodsky, *J. Am. Chem. Soc*. **2003**, *125*, 11500-11501.

[22]   a) M. Umashankara, I.R. Babu, K.N. Ganesh, *Chem. Commun*. **2003**, 2606-2607. b) I.R. Babu, K.N. Ganesh, *J. Am. Chem. Soc*. **2001**, *123*, 2079-2080.

[23]   a) R. Improta, C. Benzi, V. Barone, *J. Am. Chem. Soc*. **2001**, *123*, 12568-12577. b) R. Improta, F. Mele, O. Crescenzi, C. Benzi, V. Barone, *J. Am. Chem. Soc*. **2002**, *124*, 7857-7865. c) C. Benzi, R. Improta, G. Scalmani, V. Barone, *J. Comput. Chem*. **2002**, *23*, 341-350.

[24]   S.D. Mooney, P.A. Kollman, T.E. Klein, *Biopolymers*. **2002**, *64*, 63-71.

[25]   For an alternative chemosynthetic approach to imino acid incorporation into proteins, please see: B.L. Nilsson, R.J. Hondal, M.B. Soellner, R.T. Raines, *J. Am. Chem. Soc*. **2003**, *125*, 5268-5269.

[26]   D.D. Buechter, D.N. Paolella, B.S. Leslie, M.S. Brown, K.A. Mehos, E.A. Gruskin, *J. Biol. Chem.* **2003**, *278*, 645-650.

[27]   C. Renner, S. Alefelder, J.H. Bae, N. Budisa, R. Huber, L. Moroder, *Angew. Chem. Int. Ed. Engl*. 2001, *40*, 923-925.

[28]   N. Budisa, C. Minks, F.J. Medrano, J. Lutz, R. Huber, L. Moroder, *Proc. Natl. Acad. Sci. U S A*. **1998**, *95*, 455-459.

[29]   a) T.J. Deming, M.J. Fournier, T.L. Mason, D.A. Tirrell, *Macromolecules* **1996**, *29*, 1442-1444. b) T.J. Deming, M.J. Fournier, T.L. Mason, D.A. Tirrell, *J. Macromol. Sci., Pure Appl. Chem.*, **1997**, *A34*, 2143-2150.

[30]   a) L. Fowden, M.H. Richmond, *Biochim Biophys. Acta* **1963**, *71*, 459-461. b) L. Fowden, S. Neale, H. Tristam, *Nature* **1963**, *199*, 35-38.

[31]   B. Vivet, F. Cavelier, J. Martinez, *Eur. J. Org. Chem.* **2000**, 807-811.

[32]   P. Hughes, J. Clardy, *J. Org. Chem.* **1988**, 53, 4793-4796

[33]   a) R.A. McMillan, T.A.T. Lee, V.P. Conticello, *Macromolecules* **1999**, *32*, 3643-3648. b) R.A. McMillan, V.P. Conticello, *Macromolecules* **2000**, *33*, 4809-4821.

[34]   D.W. Urry in *Protein-Based Materials*. (Eds. K.P. McGrath, D. Kaplan), Birkhauser, Boston, **1997**, pp 133-177.

[35]   B. Li, D.O. Alonso, V. Daggett, *J. Mol. Biol*. **2001**, *305*, 581-592.

[36]   M. Ibba, D. Soll, *Science* **1999**, *286*, 1893-1897.

[37]   a) S. Fukai, O. Nureki, S. Sekine, A. Shimada, J. Tao, D.G. Vassylyev, S. Yokoyama, *Cell* **2000**, *103*, 793-803. b) T.L. Lincecum, Jr., M. Tukalo, A. Yaremchuk, R.S. Mursinna, A.M. Williams, B.S. Sproat, W. Van Den Eynde, A. Link, S. Van Calenbergh, M. Grotli, S.A. Martinis, S. Cusack, *Mol. Cell*. **2003**, *11*, 951-963.

[38]   a) J.C. van Hest, D.A. Tirrell, *FEBS Lett*. **1998**, 428, 68-70. b) K.L. Kiick, J.C. van Hest, D.A. Tirrell, *Angew. Chem. Int. Ed. Engl*. **2000**, *39*, 2148-2152. c) Y. Tang, G. Ghirlanda, W.A. Petka, T. Nakajima, W.F. DeGrado, D.A. Tirrell, *Angew. Chem. Int. Ed. Engl*. **2001**, *40*, 1494-1496. d) K.L. Kiick, R. Weberskirch, D.A. Tirrell, *FEBS Lett*. **2001**, *502*, 25-30. e) Y. Tang, D.A. Tirrell, *J. Am. Chem. Soc*. **2001**, *123*, 11089-11090. f) K.L. Kiick, E. Saxon, D.A. Tirrell, C.R. Bertozzi, *Proc. Natl. Acad. Sci. U S A*. **2002**, *99*, 19-24.

[39]   a) N. Sharma, R. Furter, P. Kast, D.A. Tirrell, *FEBS Lett*. **2000**, *467*, 37-40. b) D. Datta, P. Wang, I.S. Carrico, S.L. Mayo, D.A. Tirrell, *J. Am. Chem. Soc*. **2002**,

124, 5652-5653. c) K. Kirshenbaum, I.S. Carrico, D.A. Tirrell, *Chembiochem*.

2002, *3*, 235-237.

[40]   R. Lutz, H. Bujard, *Nucleic Acids Res*. **1997**, *25*, 1203-1210.

[41]   S. Grothe, R.L. Krogsrud, D.J. McClellan, J.L. Milner, J.M. Wood, *J. Bacteriol*.

1986, *166*, 253-259.

[42]   a) S. Kothakota, E. Yoshikawa, O. Murphy, T.L. Mason, D.A. Tirrell, M.J.

Fournier, *J. Polym. Sci. A* **1995**, *33*, 1267-1274. b) S.N. Kara-Murza, E.D.

Khodorovskaya, V.I. Mazurov, *Biokhimiya* **1974**, *39*, 599-607. c) T.S. Papas,

A.H. Mehler, *J. Biol. Chem*. **1970**, *245*, 1588-1595.

[43]   C.E. Deutch, J.L. Klarstrom, C.L. Link, D.L. Ricciardi, *Curr. Microbiol*. **2001**,

42, 442-446.

[44]   G. Gouesbet, M. Jebbar, R. Talibart, T. Bernard, C. Blanco, *Microbiology* **1994**,

140, 2415-2422.

[45]   a) A. Yaremchuk, M. Tukalo, M. Grotli, S.J. Cusack, *Mol. Biol*. **2001**, *309*, 989-

1002. b) A. Yaremchuk, S. Cusack, M. Tukalo, *EMBO J*. **2000**, *19*, 4745-4758.

[46]   S. Kamtekar, W.D. Kennedy, J. Wang, C. Stathopoulos, D. Soll, T.A. Steitz,

Proc. Natl. Acad. Sci. U S A. **2003**, *100*, 1673-1678.

[47]   C. Stehlin, D.H. Heacock, 2nd; H. Liu, K. Musier-Forsyth, *Biochemistry* **1997**, *36*, 2932-2938.

[48]   a) L. Demange, J. Cluzeau, A. Menez, C. Dugave, *Tetrahedron Lett.* **2001**, *42*, 651–653. b) L. Demange, A. Menez, C. Dugave, *Tetrahedron Lett*. **1998**, *39*, 1169-1172.

[49]   M. Singer, T.A. Baker, G. Schnitzler, S.M. Deischel, M. Goel, W. Dove, K.J. Jaacks, A.D. Grossman, J.W. Erickson, C.A. Gross, *Microbiol. Rev*. **1989**, *53*, 1-24.

[50]   K. Bohman, L.A. Isaksson, *Mol. Gen. Genet*. **1980**, *177*, 603-605.

[51]   N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann, R. Huber, *Eur*. *J. Biochem*. **1995**, *230*, 788-796.

[52]   J. Sambrook, D.W. Russell, D.W. *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory, Cold Spring Harbor, **2001**.

[53]   P.H. Hirel, M.J. Schmitter, P. Dessen, G. Fayat, S. Blanquet, *Proc. Natl. Acad. Sci. U S A*, **1989**, *86*, 8247-8251.

CHAPTER 3

A Stereoelectronic Effect on Turn Formation Due to Proline

Substitution in Elastin-Mimetic Polypeptides.

Introduction

Stereoelectronic effects have been identified recently as contributing factors to the conformational stability of proteins, the influence of which can be particularly dramatic within specific structural contexts.[1] For example, structural investigations of collagen-mimetic peptides have demonstrated that the thermodynamics of self-assembly of the native triple helical structure can be modulated by stereoelectronic effects that arise from introduction of electro-negative substituents onto the proline ring.[2-6] In particular, the thermodynamic stability of the triple helix can be interpreted in terms of the effect of the electro-negative substituent on the conformation of the pyrrolidine ring as it influences the local secondary structure of the collagen peptide. The results of these studies have provided insight into the role of post-translational hydroxylation of proline residues in stabilizing the triple helical structure of native collagens,[7] as well as a rationale for the design of non-native collagen analogues with enhanced stabilities for biomedical applications. We present results herein that suggest that stereoelectronic effects are not limited to collagen-mimetic peptide sequences and may be a more general feature of protein sequences in which proline residues occupy a critical structural role.[8] We provide evidence that elastin-mimetic polypeptide sequences are susceptible to

stereoelectronic effects that alter the thermodynamics of self-assembly of the protein, which can be rationalized in terms of the effect of proline substitution on the energetics of the β-turn conformation that develops within the pentapeptide structural repeats above the phase transition.

The proteinaceous material elastin is the primary structural component underlying the elastomeric mechanical response of compliant tissues in vertebrates,[9] and is, therefore, of considerable interest as a biomaterial for tissue engineering applications.[10] Elastin consists of a cross-linked matrix of a precursor protein, tropoelastin, which is characterized by a modular sequence of alternating, structurally distinct elastomeric and cross-linkable domains. The elastomeric domains comprise structurally similar oligopeptide motifs that are tandemly repeated in the native protein sequence. The local secondary structure and macromolecular thermodynamic and viscoelastic properties of the elastomeric domains can be emulated in synthetic polypeptides that are composed of a concatenated sequence of the native oligopeptide motifs; the most common of which is the pentapeptide [Val-Pro-Gly-Val-Gly].[11] Polypeptides based on these pentameric repeat sequences undergo reversible, temperature-dependent, hydrophobic assembly from aqueous solution in analogy to the phase behavior of native tropoelastin. This process results in spontaneous phase

separation of the polypeptide above a lower critical solution temperature, T_t, which coincides with a conformational rearrangement of the local secondary structure of the pentapeptide motifs. Biophysical studies of elastin and elastin-mimetic polypeptides have indicated a crucial role for the proline residues with regard to structure development within the pentapeptide repeats.[12-17] In particular, the proline residues participate in the formation of type II β-turn structures within the [Val-Pro-Gly-Val-Gly] repeat units, which increase in population above the phase transition of the polypeptide. The facility with which this process occurs critically influences the thermodynamics of the phase transition and is essential for assembly of the physiologically relevant state of native elastin. However, due to the limited solubility and high mobility of native elastin sequences in the functionally relevant coacervate state, direct methods for protein structural determination can not be applied easily to investigate the structural role of the proline residues on elastin assembly. In analogy with structural investigations of proline substitution in collagen-mimetic peptides,[2-6] the incorporation of structurally modified proline analogues with altered stereoelectronic properties into the pentapeptide repeats of elastin-mimetic polypeptides may provide insight into the relationship between the local structural parameters that define the β-turn conformation and the macromolecular thermodynamics of elastin assembly.

We recently reported biosynthetic methods for the co-translational incorporation of a structurally diverse series of proline analogues into an elastin-mimetic polypeptide, **elastin-1**, which afforded the elastin analogues, **elastin-2** and **elastin-3**, in which the canonical proline residues were substituted with (2*S*,4*S*)-4-fluoroproline and (2*S*,4*R*)-4-fluoroproline, respectively (Scheme 1).[18] These polypeptides were envisioned as potentially useful substrates to assess the role of stereoelectronic effects due to proline substitution on elastin assembly. The stereoelectronic gauche interaction between the vicinal fluoro-amide substituents[19] of these substituted proline derivatives has been shown to strongly influence the conformational energetics of the pyrrolidine ring. The (4*S*)- and (4*R*)-fluoroproline epimers preferentially adopt alternative conformations of the pyrrolidine ring corresponding to the C^γ-*endo* pucker and the C^γ-*exo* pucker, respectively.[1,20-22] Structural investigations of fluoroproline substitution on collagen-mimetic peptides have established that the pyrrolidine ring conformation exerts an influence on the conformational thermodynamics of a polypeptide chain through its effect on the local dihedral angles (ϕ, ψ, ω) associated with the substituted proline residues.[2-5] We hypothesized that elastin-mimetic polypeptides might also be susceptible to stereoelectronic effects that manifest themselves through their influence on the local conformation within the pentapeptide repeats. The strongest evidence for this hypothesis

arises from consideration of the crystal structure of a cyclic trimer of the elastin repeat sequence, *cyclo*-(Val-Pro-Gly-Val-Gly)$_3$.[23] The individual pentapeptide units within this structure adopt a type II β-turn conformation in which the pyrrolidine rings of the proline residues at the (*i+1*) positions of the turn uniformly display a C$^\gamma$-*exo* ring pucker. On the basis of this observation, we anticipated that incorporation of the 4-fluoroproline epimers into the pentapeptide repeats of the elastin-mimetic polypeptide sequence, **elastin-1**, might introduce a stereochemical bias with respect to the development of the type II β-turn conformation among the pentapeptide units during the elastin assembly. This investigation would provide structural insight into the importance of the proline ring conformation on the thermodynamics of the phase transition, which can have important implications for the *de novo* design of elastin-based biomaterials for applications in tissue engineering[24-27] and biotechnology.[28-30]

Experimental Section

Materials and Methods.

Protein samples of **elastin-1**, **elastin-2**, and **elastin-3** were prepared via bacterial fermentation and purified via immobilized metal affinity chromatography as previously reported.[18] The noncanonical amino acids (2*S*,4*S*)-4-fluoroproline and (2*S*,4*R*)-4-fluoroproline were purchased from Bachem Bioscience, Inc. (King of Prussia, PA). The ^{15}N-labeled proteins were prepared from fermentations of *E. coli* expression strain DG99(*proC*::Tn10)/pAG2 in M9 minimal medium under supplementation with $^{15}NH_4Cl$ (7.5 mM) and the appropriate proline derivative. The ^{15}N labeled proteins were purified as described for unlabeled proteins. Protein solutions for analytical measurements were prepared from lyophilized specimens that were dissolved at the appropriate concentration in distilled, deionized water at 4 °C. Quantitative amino acid analysis was performed on aliquots of protein stock solutions (ca. 0.2-0.3 mg/mL) at the W. M. Keck Foundation Biotechnology Resource Laboratory of Yale University to provide accurate concentrations for CD spectroscopic experiments.

Physical and Analytical Measurements.

The inverse temperature transitions of the elastin polypeptides were monitored as a function of temperature using an ultrasensitive differential scanning calorimeter (VP-DSC MicroCal, LLC, Northampton, MA). Protein samples were dissolved in distilled, deionized water at 4 °C in concentrations ranging from 0.5 to 2 mg/mL, degassed under dynamic vacuum and scanned from 5 to 60 °C at a rate of 60°/h. DSC data were processed using the program Origin (MicroCal, LLC, Northampton, MA), and T_t, ΔH, and ΔC_p values were calculated by curve fitting to the simplest appropriate model using the Levenberg/Marquardt nonlinear least-squares method. Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a PFD-425S Peltier temperature control unit in 0.2 mm sealed quartz cells at concentrations of 5.8 μM (**elastin-1**), 8.8 μM (**elastin-2**), and 6.4 μM (**elastin-3**) in distilled, deionized water. Temperature/wavelength CD-scans were performed within the temperature range from 5 °C to 65 °C with equilibration for 5 min at each temperature. The reversibility of the CD spectra was confirmed by scanning in the opposite direction of decreasing temperature with a similar equilibration period. Minimal hysteresis was observed between the forward and reverse scans under the conditions of polypeptide concentration employed in this study. Spectra were obtained from 260 to 190 nm at a resolution of 0.2 nm and at a scanning speed of 50 nm/min. The CD curves represented the average of five

measurements and were smoothed using the means-movement method on the interval

analysis of the spectral manager program. CD data are reported as mean residue

ellipticity ($[\theta]$, deg cm^2 dmol^{-1}) in which the molar masses of the polypeptides **elastin-2**

and **elastin-3** were calculated on the basis of complete substitution of the canonical

proline residues with the respective amino acid analogue. NMR spectra were acquired on

a Varian INOVA 600 (^{1}H, 599.74 MHz; ^{15}N, 60.78 MHz) at 4 °C. The NMR samples

were prepared by dissolving the respective polypeptides in a H_2O/D_2O (70:30) mixture at

a concentration of 10 mg/mL. The pH of the specimens was adjusted to 2.7 to retard

amide proton exchange on the NMR time scale. Chemical shifts for ^{1}H NMR spectra

were referenced and reported relative to internal sodium 2,2-dimethyl-2-silapenta-5-

sulfonate (0.0 ppm). Chemical shifts for ^{15}N NMR spectra were referenced and reported

relative to 1 M urea (^{15}N, 98%+) in dimethyl sulfoxide as an external standard (77.0 ppm).

Standard solvent suppression techniques were employed to reduce signal due to the

residual protons of H_2O in the ^{1}H NMR of aqueous solutions of the polypeptides. Two-

dimensional ^{1}H-^{1}H NOESY NMR spectra were acquired in phase-sensitive mode using

the hypercomplex method with a mixing time of 200 ms at a spectral width of 6799.8 Hz.

Spectra were collected with 512 t_1 increments and 2048 complex data points with 32

scans. The ^{1}H-^{15}N HSQC NMR spectra were acquired in phase-sensitive mode in which

[15]N decoupling was applied for data acquisition. The spectral width was 6200.1 Hz for the ^{1}H channel and 2791.0 Hz for the ^{15}N channel. The data matrices contained 512 t_1 increments with 2048 complex points. The spectra were acquired with 32 scans per t_1 increment. The two-dimensional NMR data were further processed using the program NutsPro (Acorn NMR, Inc.).

Computational Methods.

The second and third amino acids in the [Val-Pro-Gly-Val-Gly] moiety that serve as turn elements for the pentapeptide repeats were capped as follows: CH_3-C(=O)-Xaa-Gly-NH-CH_3 (**PG**). The latter was constructed in Macromodel[58] and conformationally manipulated to produce three reverse turns (β_I, β_{II}, and inverse γ). Each turn type was modified to generate the corresponding C^γ-*endo* and C^γ-*exo* puckered proline conformers. The structures were further modified to add a fluorine atom at C-4 of the proline rings to produce the corresponding (4*S*)- and (4*R*)-fluoroproline substituted structures. The corresponding 18 (**PG**) structures were optimized using the AMBER* force field to provide input to density functional theory (DFT) calculations. Several sets of optimizations and single-point calculations were performed on each structure as listed in Tables 2-4 employing the Gaussian-03 suite of programs.[59] The first phase of

calculations optimized geometries using the Beck3LYP functional and the double-ξ 6-31G* basis set (Beck3LYP/6-31G*//Beck3LYP/6-31G*). The second phase utilized the latter structures as input to a DFT triple-ξ (plus a diffuse function) optimization for the same set of structures (Beck3LYP/6-311+G*//Beck3LYP/6-311+G*). The third phase employed (**PG**) geometries from both levels of calculation as input to a single-point energy evaluation using the more extensive triple-ξ Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-31G* and Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-311+G* basis sets. Finally, chemical shifts for ^{15}N in N-H of the central Gly engaged in the β-turns corresponding to Gly[3] of the VPGVG repeat units were estimated from the Gauge-Independent Atomic Orbital (GIAO) method as implemented in Gaussian-03 in combination with the MPW1PW91 and PBE protocols and calculated relative to an absolute shielding of 244.6 ppm for liquid ammonia at 25 °C.[60] The MPW1PW91 functional61 has been shown to provide reliable chemical shifts for ^{13}C, while the parameter-free PBE model[62] proved superior for ^{15}N shifts in a six-way comparison. For the comparisons in Table 1, chemical shifts for ^{15}N in the Gly residue are expressed in ppm relative to 244.6 ppm. For example, for the fluorinated *exo*-F$_{eq}$ type II β-turn, the Beck3LYP/6-311+G*//Beck3LYP/6-311+G* optimized structure delivers a raw isotropic chemical shift of 132.4 ppm. The NH$_3$-adjusted relative shift was obtained from the equation $\delta(exo\text{-}F_{eq}) = \delta(NH_3) - \delta(N\text{-}$

H) = 244.6 - 132.4 = 112.2 ppm. The corresponding isotropic PBE value of 121.4 ppm is

then $\delta(NH_3) - \delta(N-H) = 244.6 - 121.4 = 123.2$ ppm.

Results and Discussion

Biosynthesis of Elastin Analogues.

Elastin-mimetic polypeptides **elastin-1**, **elastin-2**, and **elastin-3** (Scheme 1) were prepared as previously described using the *E. coli* auxotrophic host strain DG99 (*proC*::Tn10) upon expression of the target gene from a modified pQE-80 plasmid in the presence of the respective proline derivative.[18] A similar protocol was employed to produce variants of **elastin-1**, **-2**, and **-3** in which the nitrogen atoms of the non-proline residues within the respective polypeptides were uniformly labeled with the ^{15}N isotope for ^{1}H-^{15}N HSQC NMR spectroscopic experiments. The target proteins were purified to homogeneity from the endogenous proteins of the bacterial host as fusions to an *N*-terminal decahistidine sequence using immobilized metal affinity chromatography. The decahistidine leader sequence was retained in the elastin derivatives employed in subsequent calorimetric and spectroscopic analyses. Due to its small size relative to the repetitive elastin sequence, it was anticipated that the leader sequence would not significantly influence the physical properties of the macromolecules and was shown to behave accordingly. The isolated yields of the polypeptides were approximately 45-50 mg/L for fully induced expression cultures in modified minimal medium (NMM) under

A. **Elastin-1**: $MGH_{10}S_2GHID_4KHM[(VPGVG)_4VPGIG]_{16}V$

B.
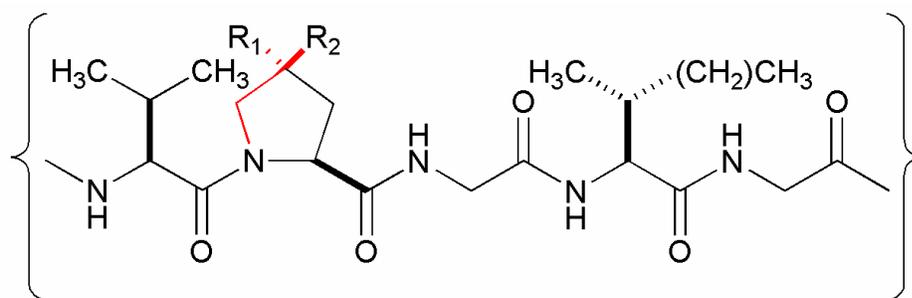
Elastin-1: R_1 = H, R_2 = H; Elastin-2: R_1 = H, R_2 = F; Elastin-3: R_1 = F; R_2 = H

Scheme 1. (**A**) Complete amino acid sequence of the elastin-mimetic polypeptide, **elastin-1**. (**B**) Structural formulae of the pentapeptide repeat units corresponding to the elastomeric domains of **elastin-1, elastin-2**, and **elastin-3**. The bond vectors that define the stereoelectronic fluoro-amide gauche interactions within the pyrrolidine ring are highlighted in color.

conventional batch fermentation conditions in shake flask culture. The protein yields of

elastin-2 and **elastin-3** were comparable to the parent sequence **elastin-1** and are quite

respectable, especially in consideration of the high analogue content within the respective

polypeptide sequences. Amino acid compositional analysis and MALDI-TOF mass

spectrometry of the elastin derivatives indicated virtually complete substitution of proline

with the respective imino acid analogue as previously described.[18] Little residual proline

content was detected in the recombinant target proteins, and the molecular ions within the

mass spectra corresponded well with the calculated masses for elastin derivatives in

which the encoded proline residues were completely substituted with the respective

proline analogues. The ^{19}F NMR analyses of **elastin-2** and **elastin-3** demonstrated

spectroscopic features that were commensurate with high levels of incorporation of the

respective fluoroproline derivative in comparison to the canonical amino acid within the

same structural context.[18]

Calorimetric Measurement of the Elastin Phase Transition.

Differential scanning calorimetry (DSC) provides a convenient method for

determination of the thermodynamic parameters associated with the elastin assembly

from aqueous solution.[33] DSC measurements on dilute aqueous solutions (ca. 1.0

mg/mL) of the elastin-mimetic polypeptides **elastin-1**, **-2**, and **-3** indicated that the value

of the transition temperature, T_t, associated with the endothermic transition depended

dramatically on the identity of the proline analogue that had been incorporated into the

polypeptide sequence (Figure 1). The (4S)- and (4R)- fluoroproline-substituted elastins

displayed transition temperatures that were shifted toward higher (41.1 °C) and lower

(21.6 °C) temperatures, respectively, in comparison with the T_t of the parent polypeptide,

elastin-1 (33.1 °C). In support of the observed trend in T_t, the van't Hoff enthalpies

calculated from the DSC data decrease in the order: **elastin-3** (+908 ± 16 kJ mol^{-1}) >

elastin-1 (+879 ± 9 kJ mol^{-1}) > **elastin-2** (+782 ± 13 kJ mol^{-1}).[34] These results are in

agreement with previously observed endothermic transition enthalpies (ΔH) for elastin

mimetic polymers[11] and block-copolymers[35] and indicate an entropy-driven (positive

ΔS) process consistent with hydrophobic or other solvation-related interactions. These

calorimetric data imply that the entropic driving force associated with the hydrophobic

assembly process is greatest for **elastin-3**. Significantly in this regard, **elastin-3** is the

only polypeptide that, at low concentration (≤1.0 mg/mL), displays a measurable

decrease in apparent heat capacity (ΔC_p = -24.5 ± 0.92 kJ mol^{-1} K^{-1} or -0.66 J K^{-1} g^{-1})

above the transition temperature (as indicated by the change in baseline for **elastin-3** in

Figure 1). This heat capacity decrement is usually indicative of the formation of
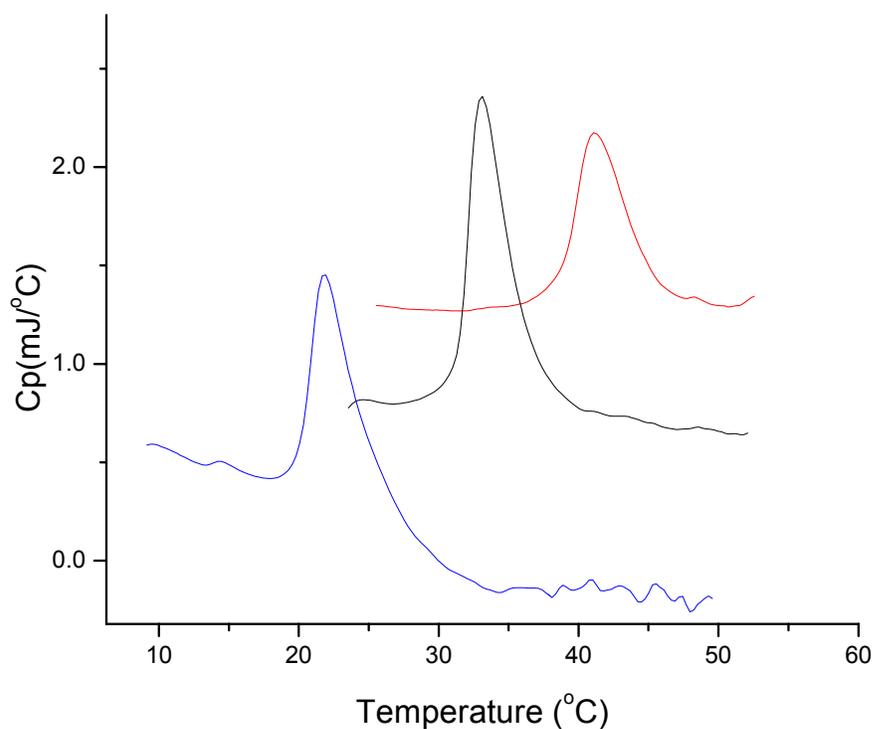
Figure 1. Raw differential scanning calorimetry data for dilute aqueous solutions (1.0 mg/mL) of **elastin-1** (black), **elastin-2** (red), and **elastin-3** (blue). These results are in agreement with previously observed endothermic transition enthalpies (ΔH) for elastin-mimetic polymers and block-copolymers,[22] and indicate an entropy-driven (positive ΔS) process consistent with hydrophobic or other solvation-related interactions. Note the pronounced decrease in baseline for **elastin-3** subsequent to the calorimetric transition, which corresponds to the observed heat capacity decrement (negative ΔC_p) associated with the endothermic process.

noncovalent interactions associated with a hydrophobic assembly process. The value of

ΔC_p for **elastin-3** compares well to the average values seen upon folding of globular

proteins from the denatured state to a more compact native fold (mean $\Delta C_p \approx$ -0.48 ± 0.14

J K^{-1} g^{-1} at 25 °C),[36,37] which suggests that the aggregation of polypeptide chains

observed for **elastin-3** resembles, at least in part, the burial of peptide groups in a

compact protein fold. We suggest that the differences in thermodynamic parameters

observed between **elastin-2** and **elastin-3**, particularly vis-à-vis the parent polypeptide

elastin-1, may be attributed to the influence of stereoelectronic effects due to the

presence of the substituted proline residues on the energetics of elastin assembly.

Previous calorimetric and turbidimetric studies of the self-assembly of elastin analogues

demonstrated that large differences could be observed in the transition temperature of

elastin-mimetic polypeptides due to amino acid substitutions in the pentapeptide repeat

sequence.[38,39] The structurally non-critical fourth position of the elastin repeat sequence

is tolerant of nonconservative substitutions such that variant polypeptides can be

synthesized in which guest residues replace the canonical Val^4 residue within the

pentapeptides at differing levels of fractional incorporation. These amino acid

substitutions alter the position of the phase transition of the elastin analogues in a manner

that depends strongly on the polarity of the amino acid side chain of the guest residues

and the mole fraction of the variant pentapeptides within the elastin sequence. The observed differences in thermodynamic parameters of the variant elastin polypeptide sequence with respect to the consensus sequence have been rationalized in terms of the effect of the amino acid substitutions on the hydration shell of the elastin polypeptides rather than intrinsic structural effects. However, we believe that this phenomenon cannot account for the significant difference in transition temperature observed between **elastin-2** and **elastin-3**, *which are identical in composition, sequence, and molar mass and are distinguished only with respect to the configuration of the asymmetric center at C-4 on the structurally critical proline derivative.*

Conformational Analysis of Elastin Polypeptides.

If indeed stereoelectronic effects give rise to the observed differences in the thermodynamics of assembly among the elastin derivatives, these intrinsic structural effects should become apparent through their influence on the polypeptide conformation. Circular dichroism (CD) spectroscopy was employed to examine the implications of this hypothesis, particularly in the context of secondary structure development within the pentapeptide repeats during the thermal transition in dilute aqueous solution. Prior CD studies of elastin-mimetic polypeptides indicated that a conformational rearrangement

occurred as the temperature increased through the transition point, which corresponded to a conversion of the local secondary structure of the pentapeptide repeats from a random coil conformation to a more ordered type II β-turn conformation.[12,15] This conformational transition can be detected in the CD spectra of **elastin-1** as a function of temperature in that the random coil signature (negative ellipticity near 195 nm) is gradually replaced with the type II β-turn signature (positive ellipticity near 207 nm) as the temperature is increased through the transition point (Figure 2).[12,15,40] Similar behavior is observed in the CD spectra of **elastin-3**, with the exception that the onset of the conformational transition occurs at a lower temperature and that a more fully developed β-turn conformation is observed at higher temperatures. In contrast, the CD spectra of **elastin-2** differ significantly from those of the other two elastin-mimetic polypeptides in that evidence of an alternative conformation is detected even at lower temperatures as a shoulder at approximately 220 nm on the major absorbance. In addition, the CD minimum associated with the random coil absorption is shifted to slightly higher wavelengths in comparison to those of **elastin-1** and **elastin-3**. As the random coil signature diminishes at higher temperatures, the longer wavelength feature dominates the CD spectra of **elastin-2**. A weak positive ellipticity is also observed near 208 nm in the CD spectra of **elastin-2** at higher temperatures that might indicate the presence of a type
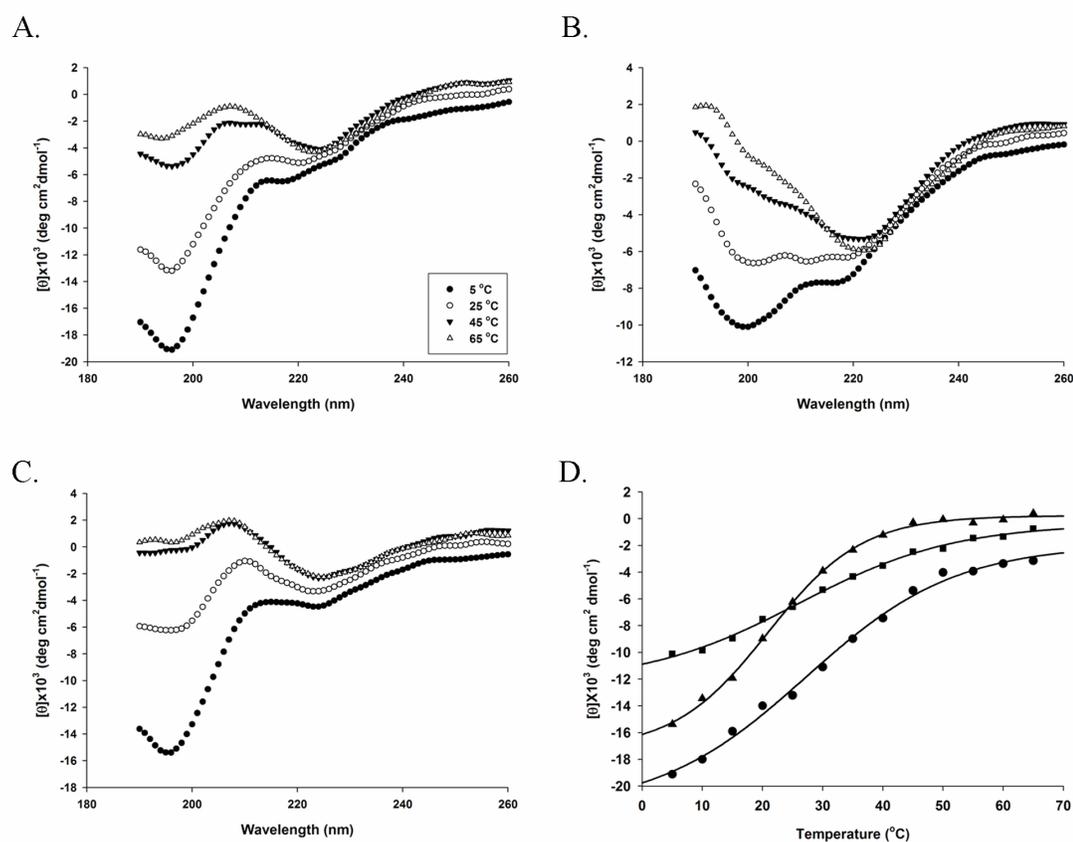
Figure 2. (**A**), (**B**), and (**C**) Circular dichroism spectral manifolds depicting the thermally induced conformational transitions for the polypeptides **elastin-1**, **elastin-2**, and **elastin-3**, respectively, at representative temperatures within proximity of the phase transition. (**D**) Thermal transition profiles monitoring the disappearance of the random coil conformation ($[\theta]_{195}$) in the CD spectra of **elastin-1** (circle), **elastin-2** (square), and **elastin-3** (triangle).

II β-turn conformation. However, the breadth of the CD traces observed at higher temperatures suggests that the (VPGVG) repeat unit significantly populates alternative conformations in addition to a type II β-turn (vide infra) as the temperature approaches the transition point. Indeed, the thermal transition curves for the disappearance of the random coil conformation, as judged by the change in mean residue ellipticity at the wavelength associated with the random coil minimum, $[\theta]_{195}$, suggest the presence of a two-state transition for **elastin-1** and **elastin-3**, but not for **elastin-2**, under the experimental conditions (Figure 2D). The mathematical fits of the thermal transitions for **elastin-1** and **elastin-3** provided estimates for the respective transition temperatures (T_t) of 29.1 °C and 22.8 °C that approximate the corresponding T_t values calculated from DSC measurements of the respective polypeptides.[40] Similar structural analyses of modified collagen-mimetic peptides suggested that the molecular origin of the observed differences in thermodynamic and conformational properties between the elastin analogues might be rationalized in terms of the stereoelectronic effect of the *fluoro* substituents on the conformation of the pyrrolidine ring, particularly as it affects the local secondary structure of the polypeptide.[2-5] These stereoelectronic effects might be manifested within the elastin structure either through the influence of the pyrrolidine ring pucker on the *cis/trans* conformational equilibrium of the Val-Pro peptide bond, through

its influence on the local (ϕ, ψ) torsional angles of the proline residue, or through a combination of the two effects.[1] Two-dimensional ^{1}H-^{1}H NOESY NMR spectroscopy of the elastin-mimetic polypeptides indicated that the major peptidyl bond conformation corresponded to the *trans* isomer as detected through the strong Val(Hα)-Pro(Hδ) correlation for all three elastin analogues (Figure 3, 4).[41] Minor peaks corresponding to the *cis* isomer of the Val-Pro peptide bond could be identified for **elastin-2**. However, even in the latter situation, in which the (4S)-fluoroproline should display a greater propensity for the *cis* prolyl peptide bond isomer,[8,21] the integrated intensity of the minor peaks corresponded to less than 20% of the total peptide pool. Similarly, the previously reported ^{19}F NMR spectra of **elastin-2** and **elastin-3** provided evidence for the presence of a minor *cis*-peptidyl isomer at a fractional content consistent with that observed from the respective ^{1}H NMR spectra.[18] The strong preference for the *trans* prolyl peptide bond isomer within these elastin derivatives coincides with the previously described propensity of Xaa-Pro peptidyl bonds to favor the *trans* orientation when the preceding residues have β-branched side chains such as the valine residues within the elastin repeats.[42] Although, especially in the case of **elastin-2**, the presence of a *cis* Val-Pro peptidyl bond isomer may play a minor role in the conformational equilibria that influence the elastin assembly, the NOESY-NMR spectroscopic data suggest that the
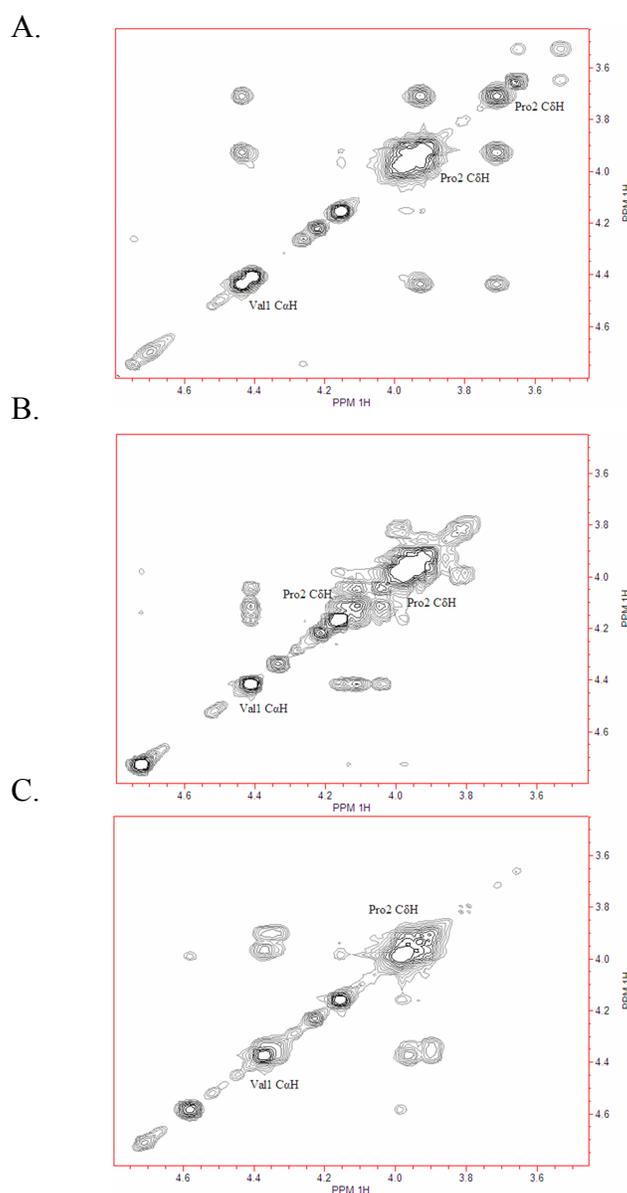
Figure 3. Expansion of the two-dimensional ^{1}H-^{1}H NOESY NMR spectra depicting the Val(αH)-Pro(δH) cross-peaks associated with the *trans* configuration of the Val-Pro peptidyl bond for **elastin-1** (**A**), **elastin-2** (**B**), and **elastin-3** (**C**). Note the absence of a strong Val(αH)-Pro(αH) cross-peak corresponding to a *cis* Val-Pro peptidyl bond configuration.

A.
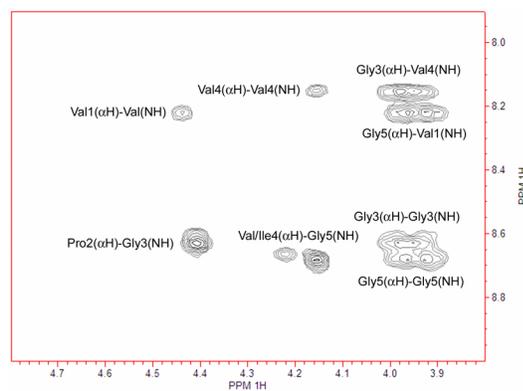
B.
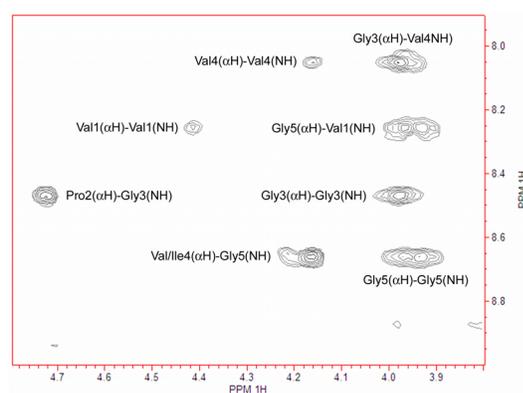
C.
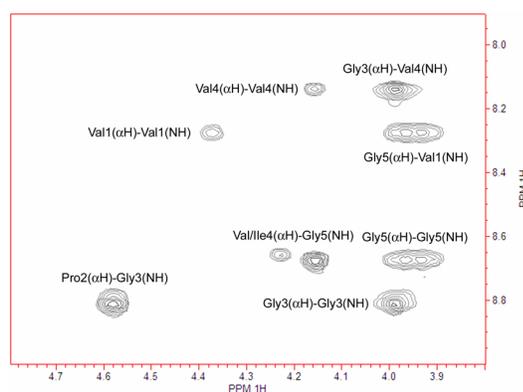
Figure 4. Expansion of the two-dimensional ^{1}H-^{1}H NOESY NMR spectra depicting the HN-Hα cross-peaks that were employed to confirm the spectroscopic assignments of the non-proline residues in the pentapeptide repeats of the elastomeric domains of **elastin-1** (**A**), **elastin-2** (**B**), and **elastin-3** (**C**).

observed differences in the macromolecular properties between these elastin derivatives

detected from CD and DSC studies do not arise primarily as a consequence of prolyl

peptidyl bond isomerism. Alternatively, stereoelectronic effects on protein structure can

stem from the influence of the pyrrolidine ring pucker on the local (ϕ,ψ) angles of the

proline residue. Despite the structural similarity of (4S)- and (4R)-fluoroproline, the

stereoelectronic fluorine-amide gauche effect alters the preferred ring pucker between the

two isomers, which is manifested in subtle but distinct differences in the thermodynamic

preferences of the main chain dihedral angles associated with the proline residue.

Computational investigations of the methyl esters of the N-acetyl-4-fluoroproline epimers

have shown significant differences in the dependence of the conformational energy on the

value of the dihedral angle ψ, particularly notable for the (4S)-fluoroproline derivative in

the C$^\gamma$-*endo* pucker of the pyrrolidine ring.[1] We suggest that the differences in

macromolecular behavior observed among the elastin analogues can be interpreted in

terms of the closeness of correspondence between the preferred values of the ψ angle for

the substituted proline derivatives and the preferred range of ψ angles associated with

proline residues in the (i + 1) position of a type II β-turn conformation (ideal β_{II} ϕ,ψ: -60°,

120°). Prior structural investigations of elastin-mimetic polypeptides have demonstrated

that the phase transition is closely associated with the formation of type II β-turn

structures within the pentapeptide repeats, in which the Pro²-Gly³ residues of the repeats occupy the (i + 1) and (i + 2) positions of the turn sequence, respectively.[12-17] A crystallographic structural determination on the model elastin peptide, *cyclo*-(VPGVG)₃,[24] indicated that the (ϕ,ψ) dihedral angles observed for the proline residues, (-53°, 140°), are within the expected range for the (i + 1) position of a type II β-turn conformation. We noted above that the conformation of the pyrrolidine rings of proline residues within the crystal structure of *cyclo*-(VPGVG)₃ corresponds to a C$^\gamma$-*exo* pucker. Significantly, the minimal energy dihedral angles calculated for the *exo* isomer of the model compound, Ac-(4*R*)-F-Pro-OMe (-59.22°, 140.79°),[1] closely correspond to those observed for the proline residues in the crystal structure of *cyclo*-(VPGVG)₃. In contrast, the unsubstituted proline residue displays a slight thermodynamic preference for the C$^\gamma$-*endo* ring pucker, in which the dihedral angles diverge from those preferred for a type II β-turn.[1,8] Thus, a stereoelectronic effect may preorganize the conformation of the (4*R*)-fluoroproline residues in **elastin-3** for the transition to a type II β-turn, thereby facilitating the self-assembly process vis-à-vis **elastin-1**. How then can one explain the anomalous conformational behavior of **elastin-2**? The (4*S*)-fluoroproline residue displays a strong thermodynamic preference for the C$^\gamma$-*endo* ring pucker as a consequence of the gauche effect between the vicinal fluorine and amide substituents.[1,8,21-23] However,

computational studies of the conformational energetics for the model compound Ac-(4*S*)-

F-Pro-OMe indicated that a 1,3-diaxial interaction between the fluoro and carboxyl

substituents on the pyrrolidine ring significantly distorts the minimal energy values of the

(ϕ,ψ) angles for the substituted proline residue in the C$^\gamma$-*endo* conformation (-76.44°,

171.95°)[1] away from the typical values of the ψ angle associated with the type II β-turn

conformation. In addition, calculations of total energy and steric exchange energy versus

ψ for fragments derived from this model compound exhibited energetic maxima at

values of ψ corresponding to those associated with the type II β-turn conformation.

Taken in combination, these considerations suggest that the C$^\gamma$-*endo* pucker of

pyrrolidine ring in (4*S*)-fluoroproline energetically destabilizes the type II γ-turn structure

relative to alternative conformations that are energetically accessible for the pentapeptide

repeats, which results in the elevated transition temperature and anomalous

conformational behavior observed for **elastin-2**. Complementary spectroscopic evidence

suggests that significant conformational differences occur between the **elastin-2** and

elastin-3, which presumably originate from an altered population of pyrrolidine ring

conformers among the substituted proline residues. Conformational information

regarding the local environment about the proline residue can be gleaned from the ^{1}H-^{15}N

HSQC NMR spectra of the elastin analogues (Figure 5).
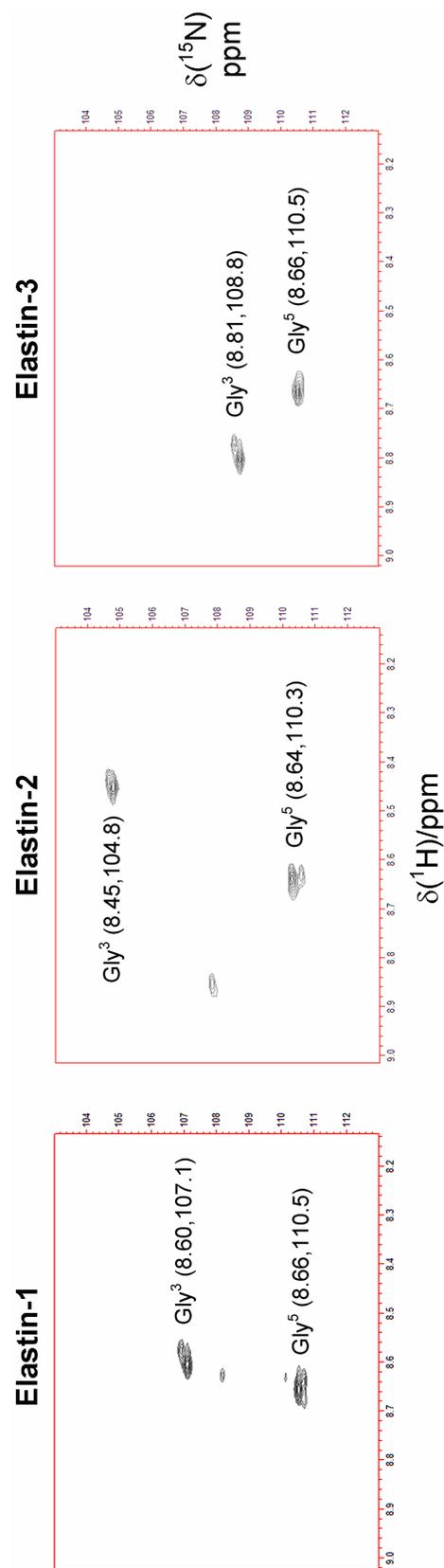
Figure 5. Comparison of the ^{1}H-^{15}N HSQC NMR spectra of **elastin-1**, **elastin-2**, and **elastin-3**, indicating the positions of the chemical shifts associated with the glycine residues that occur in the repetitive domain. Spectra were acquired at 4 °C as described in the Experimental Section on elastin specimens that had been labeled with ^{15}N at the amide positions corresponding to non-proline residues, i.e., valine, glycine, and isoleucine, within the repetitive domains of the polypeptides. The unlabeled cross-peaks in the spectra correspond to minor species that might be associated with the presence of *cis*-peptidyl isomers of the Val-Pro bond, and, consequently do not show correlations to the major *trans* peptidyl bond isomer under the experimental conditions.

The ^{1}H and ^{15}N chemical shifts of the Gly[3] amide group should depend strongly on the ψ angle of the preceding proline residue[43,44] and, therefore, should be sensitive to local structural perturbations that arise from stereoelectronic effects. The values of these chemical shifts depend dramatically on the identity of the proline derivative within the elastin-mimetic polypeptide. The ^{1}H and ^{15}N chemical shifts of Gly[3] are displaced to high field for **elastin-2** and to low field for **elastin-3** relative to the corresponding chemical shift values for **elastin-1**. These shifts reflect incipient structural development within the Pro[2]-Gly[3] structural unit, which even at 4 °C does not correspond to a completely denatured random coil conformation on the basis of the values of $[\theta]_{195}$ observed in the CD spectra (Figure 2).[15,45] We may infer from the observed distribution of Gly[3] chemical shifts among the elastin analogues that the Pro[2] residues experience different conformational environments even at temperatures below the respective transition points. In contrast, the values of the ^{1}H and ^{15}N chemical shifts for the Gly[5] amide group, which does not participate in the incipient β-turn conformation, do not significantly vary among the elastin derivatives. The downfield displacement of the Gly[3] [15]N and ^{1}H chemical shifts for **elastin-3** with respect to the corresponding resonances of **elastin-1** is consistent with increased occupancy of the type II β-turn conformation on the basis of empirical chemical shift correlations.[43,44] However, the observed upfield shift of the ^{1}H and [15]N

resonances derived from the Gly[3] residues in the HSQC NMR spectra of **elastin-2** cannot be interpreted in terms of chain conformations commonly associated with the elastin pentapeptide repeats on the basis of spectroscopic[12,15] and computational[16] analyses and may reflect with increased occupancy of alternative turn structures in which the local conformation of the proline residue is shifted toward less positive values of the ψ angle. Arad and Goodman have described an equilibrium between β-turn and γ-turn conformations within the pentapeptide repeats on the basis of spectroscopic analyses of elastin-mimetic peptides.[46] The relative population of the γ-turn conformation can be increased in depsipeptide analogues of the pentapeptides in which Val[4] has been replaced with the non-hydrogen bonding isostere, 2-hydroxyisovaleric acid. In addition, molecular dynamics simulations of (Val-Pro-Gly-Val-Gly)$_{18}$ at temperatures above the transition point provide evidence for a minor population of proline residues with negative ψ values that are consistent with type I β-turn structures.[16] The CD and NMR spectroscopic data for **elastin-2** suggest that the presence of the (4S)-fluoroproline residue may tip the conformational equilibria of the pentapeptide repeats toward alternative turn structures, vis-à-vis **elastin-1** and **elastin-3**, through a relative destabilization of the type II β-turn conformation. The upfield chemical shift of the ^{15}N signal for the Gly[3] amide of **elastin-2** is consistent with a greater population of type I β-turn structures based on the calculated

[15]N chemical shifts for different turn structures of the model turn segment (MeCO-Xaa-Gly-NHMe) (vide infra) and empirical ^{15}N chemical shift correlations. In addition, the CD spectra of **elastin-2** display an increase in the magnitude of the negative ellipticity absorption at 222 nm near the phase transition, which occurs at the expense of the positive ellipticity absorption at 207 nm associated with type II β-turn formation. Qualitative comparison of the CD spectra of the elastin peptides to conformational deconvolutions of the CD spectra of model type I and type II β-turn structures[47] suggests that the observed difference in CD spectroscopic behavior for **elastin-2** may be attributable to an increase in type I β-turn conformation relative to type II β-turn conformation with respect to **elastin-1** and **elastin-3**. Taken together, the spectroscopic data may indicate that the relative populations of the different turn structures associated with the conformational transition of the (VPGVG) units depends on the structural identity of the proline analogue and can be rationalized on the basis of stereoelectronic effects.

Computational Studies of (Pro-Gly) Turns.

To assist interpretation of the structural consequences of introduction of (4*S*)-fluoroproline versus (4*R*)-fluoroproline into these elastin-mimetic polypeptides, we have

used density functional theory (DFT) to model three possible turn types (β_I, β_{II}, and inverse γ) derived from peptide segments corresponding to the turnforming residues of the elastin repeat unit (VPGVG). The pentapeptide repeat unit was truncated to a capped moiety corresponding to the turn fragment (MeCO-Xaa-Gly-NHMe) (Xaa = Pro, 4S-F-Pro, 4R-F-Pro). The proline envelope conformations were placed in conformational equilibrium by virtue of a flip of the five-membered pyrrolidine ring of the proline residue between the *exo* and *endo* orientations. For ease of discussion here and in connection with Table 1, the (4R)- pair of conformers have been designated *exo*-F_{ax} and *endo*-F_{eq}, in which the *exo*-F_{ax} implies a C^γ-*exo* orientation of the proline ring with an axial disposition of the C-F bond. The (4S)-pair of conformers, *exo*-F_{eq} and *endo*-F_{ax}, are similarly characterized by equilibrating pyrrolidine ring pucker isomers. The three parent turns and the six 4-fluorinated turn structures were optimized in C^γ-*exo* and C^γ-*endo* conformations with both double-ξ and enhanced triple-ξ basis sets (6-31G* and 6-311+G*, respectively) in order to define the computational requirements for adequately describing the various turns. Supplemental energy differences were obtained using the triple-ξ basis set with extensive polarization (i.e., 6-311+G(2d,p)). The Beck3LYP/6-31G* optimizations (Table 2-4) were performed both to obtain preliminary structures for the 6-311+G* refinements and to evaluate whether they might be sufficiently predictive

for this class of problem. As it concerns molecular geometry, the Beck3LYP/6-31G* spread of ϕ, ψ angles is satisfactory, though in some cases a bit broader than derived at the 6-311+G* basis set level. Energetically, when the 6-31G* optimized structures are re-evaluated with the Beck3LYP/6-311+G(2d,p) model, conformational energy differences are routinely within 0.2 kcal/mol by comparison with structures optimized at 6-311+G* (Table 2-4). For both parent and fluorinated prolines, the conformational forms for (MeCO-Xaa-Gly-NHMe) (Xaa = Pro, 4S-F-Pro, 4R-F-Pro) positive ϕ and ψ torsional angles for the proline and glycine residues that are within 30° of the ideal ($\phi_1\psi_1\phi_2\psi_2$) values corresponding to the (i+1) and (i+2) residues of the respective turn type (Table 1).[48,49] However, the relative energetic stabilities associated with turn formation display striking differences between ring pucker conformations for the fluoroproline derivatives of the model peptide segment (Table 1). The DFT 6-311+G(2d,p) energies suggest that the type I β-turn structures corresponding to $endo$-F_{ax} and exo-F_{ax} conformers are the more stable species for (4S)- and (4R)- fluoroproline derivatives, respectively (Table 1 and Figure 6). Taken at face value, the observed energy differences of 1.8 and 1.6 kcal/mol for the $endo$-F_{ax} and exo-F_{ax} conformers vis-à-vis their conformational partners suggest 95% and 94% populations, respectively, for these low energy forms at 298 K. The observed conformational preferences of the (4S)- and (4R)-fluoroproline derivatives

β_I turn	total E, au[a]	ΔE	Pro ϕ and ψ[b]	Gly ϕ and ψ[b]	δ (N–H)[c]
exo	−781.530 155	0.4	−66.8, −18.8	−97.5, 10.2	95.1,[d] 105.9[e]
endo	−781.530 807	0.0	−78.3, −3.7	−99.8, 7.8	102.1,[d] 113.8[e]
exo-F_{eq}	−880.798 527	1.8	−68.1, −18.1	−97.3, 10.3	95.4,[d] 106.0[e]
endo-F_{ax}	−880.801 461	0.0	−74.1, −8.9	−103.1, 10.4	100.5,[d] 111.5[e]
exo-F_{ax}	−880.801 367	0.0	−70.5, −15.1	−97.5, 9.5	96.4,[d] 107.5[e]
endo-F_{eq}	−880.799 024	1.6	−78.5, −2.8	−100.3, 7.8	95.1,[d] 105.9[e]

β_{II} turn	total E, au[a]	ΔE	Pro ϕ and ψ[b]	Gly ϕ and ψ[b]	δ (N–H)[c]
exo	−781.533 453	0.0	−57.7, 127.5	99.8, −14.2	111.7,[d] 122.7[e]
endo	−781.533 325	0.08	−65.1, 126.3	97.2, −10.0	110.8,[d] 121.8[e]
exo-F_{eq}	−880.801 957	0.0	−58.8, 127.1	100.1, −13.7	112.2,[d] 123.2[e]
endo-F_{ax}	−880.799 121	1.8	−61.0, 133.1	97.7, −15.2	109.1,[d] 120.6[e]
exo-F_{ax}	−880.805 750	0.0	−60.1, 124.9	100.6, −12.9	113.1,[d] 123.9[e]
endo-F_{eq}	−880.803 256	1.6	−65.5, 125.3	99.9, −10.9	112.1,[d] 123.1[e]

i-γ turn	total E, au[a]	ΔE	Pro ϕ and ψ[b]	Gly ϕ and ψ[b]	δ (N–H)[c]
exo	−781.530 860	1.0	−82.0, 76.6	−122.1, 14.4	111.8,[d] 125.5[e]
endo	−781.532 457	0.0	−83.7, 72.7	−121.5, 14.8	110.5,[d] 124.0[e]
exo-F_{eq}	−880.799 711	0.1	−82.4, 76.8	−115.5, 9.8	113.5,[d] 125.0[e]
endo-F_{ax}	−880.799 937	0.0	−81.6, 54.3	−107.6, 8.7	107.7,[d] 121.3[e]
exo-F_{ax}	−880.803 764	0.0	−83.5, 76.6	−116.0, 9.3	111.6,[d] 125.0[e]
endo-F_{eq}	−880.801 745	1.3	−84.0, 76.1	−113.4, 7.8	111.2,[d] 124.4[e]

Table 1. Calculated values for energy differences (ΔE, kcal/mol), ϕ, ψ angles (deg), and ^{15}N chemical shifts (ppm) for type I β-turns, type II β-turns, and inverse γ-turns in the truncated turn model (MeCO-Xaa-Gly-NHMe) (Xaa = Pro, 4*S*-F-Pro, 4*R*-F-Pro)

a Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-311+G*. *b* ϕ angle (O)C-NC-C(O); ψ angle N-C-C(O)-N. *c* Calculated ^{15}N chemical shifts for N-H bonds of the Gly residue of (MeCO-Xaa-Gly-NHMe) in ppm relative to liquid NH_3 at 25 °C (absolute shielding 244.6 ppm is set to 0 ppm); see Experimental Section for details. *d* MPW1PW91/6-311G*//Beck3LYP/6-311+G*. *e* PBE/6-311+G(2d,p)//Beck3LYP/6-311+G*.

Beck3LYP/6-31G* opt[a]							
	Total E, au	ΔE_1	6-311+G(2d,p)[b]	ΔE_2	Pro ϕ and ψ[c]	Gly ϕ and ψ[c]	δ (N-H)[d,e]
exo	-781.279 423	0.6	-781.529 489	0.4	-66.0 -21.8	-103.2 16.9	96.2
endo	-781.280 313	0.0	-781.530 159	0.0	-79.8 -1.8	-105.5 13.2	104.2
exo-F$_{eq}$	-880.507 427	2.0	-880.797 713	1.7	-66.6 -21.7	-103.9 17.4	96.5
endo-F$_{ax}$	-880.510 664	0.0	-880.800 406	0.0	-75.9 -6.8	-107.6 14.9	102.6
exo-F$_{ax}$	-880.509 888	0.0	-880.800 559	0.0	-67.8 -20.7	-103.2 16.8	97.1
endo-F$_{eq}$	-880.508 173	1.1	-880.798 268	1.4	-80.4 0.2	-106.0 12.5	103.6
Beck3LYP/6-311+G* opt[f]							
	Total E, au	ΔE_3	6-311+G(2d,p)[g]	ΔE_4	Pro ϕ and ψ[c]	Gly ϕ and ψ[c]	δ (N-H)[d]
exo	-781.489 586	0.4	-781.530 155	0.4	-66.8 -18.8	-97.5 10.2	95.1[h] 105.9[i]
endo	-781.490 207	0.0	-781.530 807	0.0	-78.3 -3.7	-99.8 7.8	102.1[h] 113.8[i]
exo-F$_{eq}$	-880.757 107	1.8	-880.798 527	1.8	-68.1 -18.1	-97.3 10.3	95.4[h] 106.0[i]
endo-F$_{ax}$	-880.759 936	0.0	-880.801 461	0.0	-74.1 -8.9	-103.1 10.4	100.5[h] 111.5[i]
exo-F$_{ax}$	-880.759 852	0.0	-880.801 367	0.0	-70.5 -15.1	-97.5 9.5	96.4[h] 107.5[i]
endo-F$_{eq}$	-880.757 524	1.5	-880.799 024	1.6	-78.5 -2.8	-100.3 7.8	95.1[h] 105.9[i]

Table 2. Type I β-turns in the VPGVG truncated peptide; energy differences (ΔE, kcal/mol), ϕ, ψ angles (deg) and ^{15}N chemical shifts (ppm).

a Beck3LYP/6-31G*//Beck3LYP/6-31G*

b Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-31G*

c ϕ angle (O)C-N-C-C(O); ψ angle N-C-C(O)-N

d Calculated ^{15}N chemical shifts for N-H bonds of the Gly residue of (MeCO-Xaa-Gly-NHMe) in ppm relative to liquid NH$_3$ at 25 °C (absolute shielding 244.6 ppm is set to zero ppm); see Experimental Methods for details.

e MPW1PW91/6-311G*//Beck3LYP/6-31G*

f Beck3LYP/6-311+G*//Beck3LYP/6-311+G*

g Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-311+G*

h MPW1PW91/6-311G*//Beck3LYP/6-311+G*

i PBE/6-311+G(2d,p)//Beck3LYP/6-311+G*

Beck3LYP/6-31G* opt[a]	Total E, au	ΔE_1	6-311+G(2d,p)[b]	ΔE_2	Pro φ and ψ[c]		Gly φ and ψ[c]		δ (N̲-H)[d,e]	
exo	-781.282 827	0.3	-781.532 882	0.0	-60.5	123.4	104.6	-16.1	111.7	
endo	-781.283 337	0.0	-781.532 497	0.2	-70.7	113.5	108.6	-12.0	116.1	
exo-F_{eq}	-880.511 004	0.0	-880.801 349	0.0	-60.2	124.8	104.6	-16.5	115.0	
endo-F_{ax}	-880.508 684	1.5	-880.798 090	2.1	-62.1	133.7	100.6	-18.3	111.0	
exo-F_{ax}	-880.514 147	0.0	-880.805 096	0.0	-61.9	121.2	105.7	-15.6	116.1	
endo-F_{eq}	-880.512 803	0.9	-880.802 235	1.8	-71.2	112.1	109.7	-11.3	116.7	
Beck3LYP/6-311+G* opt[f]	Total E, au	ΔE_3	6-311+G(2d,p)[g]	ΔE_4	Pro φ and ψ[c]		Gly φ and ψ[c]		δ (N̲-H)[d]	
exo	-781.492 951	0.0	-781.533 453	0.0	-57.7	127.5	99.8	-14.2	111.7[h]	122.7[i]
endo	-781.492 830	0.08	-781.533 325	0.08	-65.1	126.3	97.2	-10.0	110.8[h]	121.8[i]
exo-F_{eq}	-880.760 646	0.0	-880.801 957	0.0	-58.8	127.1	100.1	-13.7	112.2[h]	123.2[i]
endo-F_{ax}	-880.757 554	0.8	-880.799 121	1.8	-61.0	133.1	97.7	-15.2	109.1[h]	120.6[i]
exo-F_{ax}	-880.764 367	0.0	-880.805 750	0.0	-60.1	124.9	100.6	-12.9	113.1[h]	123.9[i]
endo-F_{eq}	-880.761 920	1.5	-880.803 256	1.6	-65.5	125.3	99.9	-10.9	112.1[h]	123.1[i]

Table 3. Type II β-turns in the VPGVG truncated peptide; energy differences (ΔE, kcal/mol), φ, ψ angles (deg) and ^{15}N chemical shifts (ppm).

a Beck3LYP/6-31G*//Beck3LYP/6-31G*
b Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-31G*
c φ angle (O)C-N-C-C(O); ψ angle N-C-C(O)-N
d Calculated ^{15}N chemical shifts for N-H bonds of the Gly residue of (MeCO-Xaa-Gly-NHMe) in ppm relative to liquid NH_3 at 25 °C (absolute shielding 244.6 ppm is set to zero ppm); see Experimental Methods for details.
e MPW1PW91/6-311G*//Beck3LYP/6-31G*
f Beck3LYP/6-311+G*//Beck3LYP/6-311+G*
g Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-311+G*
h MPW1PW91/6-311G*//Beck3LYP/6-311+G*
i PBE/6-311+G(2d,p)//Beck3LYP/6-311+G*

Beck3LYP/6-31G* opt[a]							
	Total E, au	ΔE_1	6-311+G(2d,p)[b]	ΔE_2	Pro ϕ and ψ[c]	Gly ϕ and ψ[c]	δ (N-H)[d,e]
exo	-781.280 667	1.2	-781.530 318	1.0	-82.0 76.6	-122.1 14.4	113.6
endo	-781.282 518	0.0	-781.531 852	0.0	-83.7 72.7	-121.5 14.8	112.7
exo-F_{eq}	-880.509 150	0.3	-880.799 060	0.0	-82.1 76.1	-121.9 14.5	113.5
endo-F_{ax}	-880.509 689	0.0	-880.798 999	0.04	-82.4 58.6	-117.1 14.6	107.6
exo-F_{ax}	-880.512 361	0.0	-880.803 026	0.0	-83.6 75.7	-123.3 14.8	113.6
endo-F_{eq}	-880.511 499	0.5	-880.801 107	1.2	-83.7 75.2	-121.4 14.2	111.1
Beck3LYP/6-311+G* opt[f]							
	Total E, au	ΔE_3	6-311+G(2d,p)[g]	ΔE_4	Pro ϕ and ψ[c]	Gly ϕ and ψ[c]	δ (N-H)[d]
exo	-781.489 832	1.0	-781.530 860	1.0	-82.0 76.6	-122.1 14.4	111.8[h] 125.5[i]
endo	-781.491 480	0.0	-781.532 457	0.0	-83.7 72.7	-121.5 14.8	110.5[h] 124.0[i]
exo-F_{eq}	-880.757 870	0.0	-880.799 711	0.1	-82.4 76.8	-115.5 9.8	113.5[h] 125.0[i]
endo-F_{ax}	-880.757 723	0.1	-880.799 937	0.0	-81.6 54.3	-107.6 8.7	107.7[h] 121.3[i]
exo-F_{ax}	-880.761 878	0.0	-880.803 764	0.0	-83.5 76.6	-116.0 9.3	111.6[h] 125.0[i]
endo-F_{eq}	-880.760 043	1.2	-880.801 745	1.3	-84.0 76.1	-113.4 7.8	111.2[h] 124.4[i]

Table 4. Gamma turns in the VPGVG truncated peptide; energy differences (ΔE, kcal/mol), ϕ, ψ angles (deg) and ^{15}N chemical shifts (ppm).

a Beck3LYP/6-31G*//Beck3LYP/6-31G*
b Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-31G*
c ϕ angle (O)C-N-C-C(O); ψ angle N-C-C(O)-N
d Calculated ^{15}N chemical shifts for N-H bonds of the Gly residue of (MeCO-Xaa-Gly-NHMe) in ppm relative to liquid NH_3 at 25 °C (absolute shielding 244.6 ppm is set to zero ppm); see Experimental Methods for details.
e MPW1PW91/6-311G*//Beck3LYP/6-31G*
f Beck3LYP/6-311+G*//Beck3LYP/6-311+G*
g Beck3LYP/6-311+G(2d,p)//Beck3LYP/6-311+G*
h MPW1PW91/6-311G*//Beck3LYP/6-311+G*
i PBE/6-311+G(2d,p)//Beck3LYP/6-311+G*

Figure 6. Calculated structures of conformer pairs for type I β-turn structures derived from the model peptide segment (MeCO-Xaa-Gly-NHMe) incorporating (2*S*,4*S*)-4-fluoroproline (top) and (2*S*,4*R*)-4-fluoroproline (bottom). The labels indicate the ring pucker associated with the C^γ position of the pyrrolidine ring and the position of the fluorine substituent. Geometries were derived from structural optimization at the Beck3LYP/6-311+G(2d,p) level of theory.

Figure 7. Calculated structures of conformer pairs for type II β-turn structures derived from the model peptide segment (MeCO-Xaa-Gly-NHMe) incorporating (2*S*,4*S*)-4-fluoroproline (left) and (2*S*,4*R*)-4-fluoroproline (right). The labels indicate the ring pucker associated with the C$^\gamma$ position of the pyrrolidine ring and the position of the fluorine substituent. The contact distance between axially disposed C-F and C-O bond vectors is depicted for the *endo*-F$_{ax}$ isomer of (MeCO-4*S*-F-Pro-Gly-NHMe). Geometries were derived from structural optimization at the Beck3LYP/6-311+G(2d,p) level of theory.
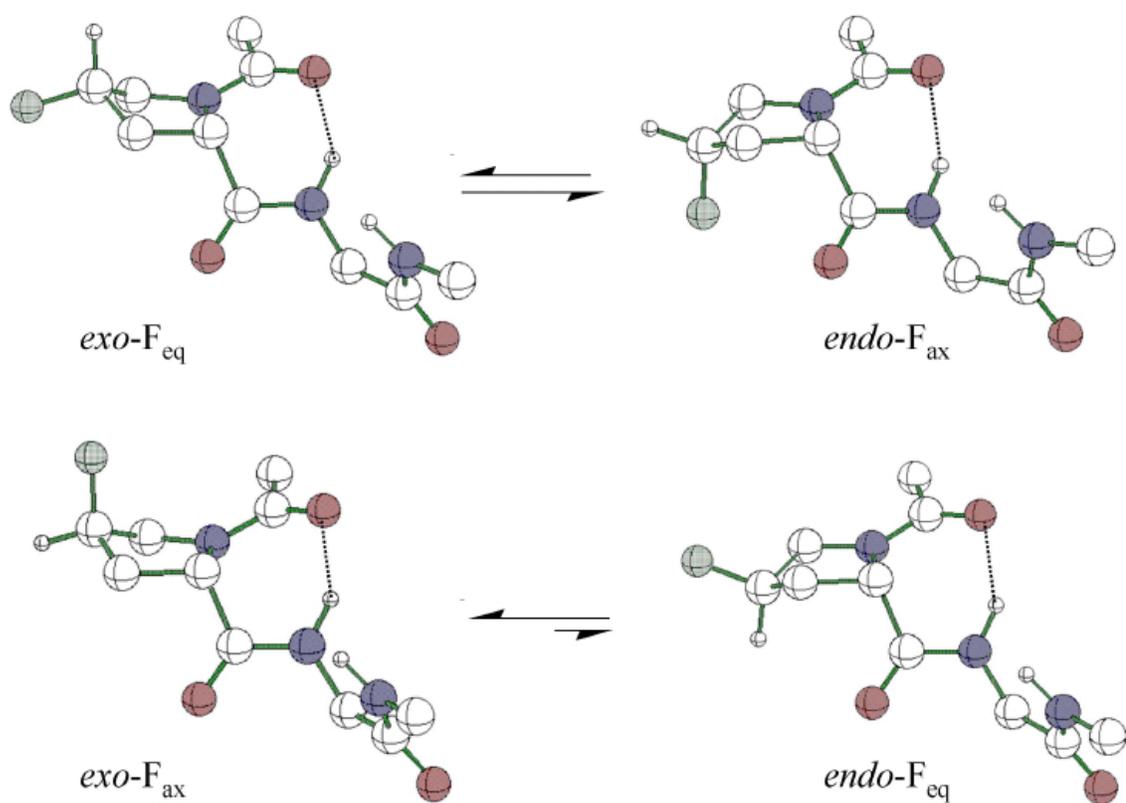
Figure 8. Calculated structures of conformer pairs for inverse γ-turn structures derived from the model peptide segment (MeCO-Pro-Gly-NHMe) incorporating (2*S*,4*S*)-4-fluoroproline (top) and (2*S*,4*R*)-4-fluoroproline (bottom). The labels indicate the ring pucker associated with C$^\gamma$ position of the pyrrolidine ring and the position of the fluorine substituent. Geometries were derived from structural optimization at the Beck3LYP/6-311+G(2d,p) level of theory.

in these type I β-turn structures reflect the conformational preferences for the C^γ-*endo* and C^γ-*exo* isomers, respectively, that had been previously reported from structural studies of the corresponding substituted proline derivatives (vide supra). In both preferred ring pucker conformations, the C-F bond has an axial orientation with respect to the pyrrolidine ring. Previous experimental studies of fluorinated proline derivatives suggested that the primary driving force for this puckering outcome could be ascribed to the fluorine-amide gauche effect and to favorable hyperconjugation between *trans* disposed C-F and C-H bonds.[1,2,20,21] These structural criteria are applicable for rationalizing the stability of the *endo*-F_{ax} and *exo*-F_{ax} conformations within the type I β-turns of (4*S*)- and (4*R*)-fluoroproline-substituted (MeCO-Pro-Gly-NHMe) units, respectively, as the preferred conformer in each case incorporates both of these structural features.

The results obtained for the type I β-turn structures can be compared to the corresponding features calculated for the type II β-turns (Table 1 and Figure 7). Energetically, the C^γ-*endo* and C^γ-*exo* conformers of the proline-containing parent turns derived from the (MeCO-Pro-Gly-NHMe) unit are within 0.1 kcal/mol (54% (*exo*)/46% (*endo*) populations, 298 K). Note that the very slight energetic difference between the two ring puckers for the unsubstituted (MeCO-Pro-Gly-NHMe) unit in a type II β-turn

structure stands in direct contrast to the single ring pucker conformation observed for

proline residues in the crystal structure of *cyclo*-(VPGVG)$_3$.[24] This discrepancy suggests

that the observed C$^\gamma$-*exo* preference in the latter structure may have arisen from a

combination of nonlocal structural considerations distinct from turn formation, including,

but not limited to, crystal packing forces, turn catenation, and peptide backbone

cyclization. In contrast with the computational results obtained for the type I β-turn

system, fluoroproline substitution within type II β-turn structures favors the C$^\gamma$-*exo*

proline conformers, *exo*-F$_{eq}$ and *exo*-F$_{ax}$, for the (4*S*)- and (4*R*)-fluoroproline derivatives,

respectively, with populations predicted to be in the 94-95% range according to the

Beck3LYP/6-311+G(2d,p) calculations. The stability of the (4*R*)-*exo*-F$_{ax}$ form can be

attributed to the gauche effect and C-H/C-F hyperconjugation as described above;

however, the (4*S*)-*exo*-F$_{eq}$ conformation can take advantage of neither of these effects.

We suggest that the source of the relative energies between ring pucker conformers

associated with (4*S*)-fluoroproline lies in the energetic instability of the *endo*-F$_{ax}$

conformational partner. As depicted in Figure 5, the C-F bond is apposed to the central

amide's C=O bond in the *endo*-F$_{ax}$ conformer with the F and O atoms separated by 3.25

Å. While this distance does not violate the sum of the van der Waals boundaries (2.99

Å),[50] it is sufficiently short to bring the nonbonding electron lone pairs from each atom

into juxtaposition. We presume that the accompanying lone-pair/lone-pair repulsion is sufficient to override both gauche and hyperconjugation effects for the axial fluorine in (4S)-$endo$-F_{ax} conformation, thereby raising its energy relative to (4S)-exo-F_{eq}. The distortion of the (4S)-$endo$-Fax conformer in a type II β-turn structure can be inferred from deviations in the values of ($\phi_1 \psi_1 \phi_2 \psi_2$) dihedral angles associated with the (MeCO-4S-F-Pro-Gly-NHMe) unit, particularly ϕ_1 of proline (133.1°), in comparison to the corresponding values for the C^{γ}-$endo$ conformers of the other proline derivatives (Table 1). In this context, it is noteworthy that, while (4S)-exo-F_{eq} is predicted to be thermochemically more stable than the (4S)-$endo$-F_{ax} conformer, the equatorial structure is calculated to be 2.3 kcal/mol *less stable* than the (4R)-exo-F_{ax} epimer associated with (4R)-fluoroproline in the type II β-turn structures predicted for MeCO-4R-F-Pro-Gly-NHMe. For a comparison of relative energies between the various turn conformations within the model peptide segment (MeCOXaa-Gly-NHMe) (Xaa = Pro, 4S-F-Pro, and 4R-F-Pro), we consider the values observed for the most stable forms recorded under ΔE in Table 1. For the unsubstituted parent peptide, the type II β-turn conformation is the most stable, with a very slight preference for the C^{γ}-exo conformation. Similarly, for the model peptide segment (MeCO-4R-F-Pro-Gly-NHMe), the exo-F_{ax} conformer of the type II β-turn is predicted to be lowest in energy and, indeed, is the most energetically stable

conformation by 2.4-3.6 kcal/mol among all of the turn sequences under consideration in this study. Notably, the C^{γ}-*exo* isomers comprise the most stable conformations for all turns involving (4*R*)-fluoroproline, which provides a dramatic example of the influence of stereoelectronic control within this structural context. In addition, the order of relative stabilities among the turn sequences is similar for the parent peptide and the (4*R*)-fluoroproline derivative in that the energies of the most stable conformers are ordered such that $\beta_{II} < \gamma < \beta_{I}$. For the (MeCO-4*R*-F-Pro-Gly-NHMe) model peptide system, the computational data indicate a greater stabilization of the type II β-turn structure with respect to alternative turns (Table 1) than for proline within the model peptide system, which provides support for the hypothesis that experimental differences observed between **elastin-1** and **elastin-3** may arise as a consequence of stereoelectronic influences within the (VPGVG) structural context. However, given the assumptions in truncating VPGVG, the absence of solvent in the calculations, and, most importantly, the absence of any elastin protein environment, it is reasonable to assume that all of the turn forms are energetically accessible to the elastin polypeptides under experimentally relevant conditions, as is suggested from spectroscopic investigations[46] and molecular dynamics simulations[16] of elastin-mimetic model peptides.

In contrast to the stereoelectronic control observed for turns involving the (4*R*)-

fluoroproline epimer, we find that the C$^{\gamma}$-*endo* isomer of (4*S*)-fluoroproline, which should

be preferred on the basis of stereoelectronic considerations associated with fluorine

substitution, is not necessarily the most stable conformation in turn sequences derived

from the model peptide segment (MeCO-4*S*-F-Pro-Gly-NHMe). For the type II β-turn

structure, the C$^{\gamma}$-*exo* isomer corresponds to the energetically most favorable pucker of the

pyrrolidine ring even though this conformational arrangement cannot gain stabilization

from either a stereoelectronic gauche effect or C-H/C-F hyperconjugation.[2,8] However,

as mentioned above, significant energetic destabilization of the C$^{\gamma}$-*endo* isomer may arise

vis-à-vis the C$^{\gamma}$-*exo* isomer due to lone-pair/lone-pair repulsion between the axial F and O

atoms of the pyrrolidine ring within the (4*S*)-fluoroproline (Figure 7). A computational

analysis of the conformational energetics of the *N*-acetyl, methyl ester derivative of (4*S*)-

fluoroproline suggested that significant energetic destabilization of the C$^{\gamma}$-*endo* isomer

was associated with particular values of the dihedral angle ψ.[1] We note from these

results that the energetic destabilization of the C$^{\gamma}$-*endo* isomer of a (4*S*)-fluoroproline-

derived structural fragment was pronounced at ψ values near the preferred ψ range for

the (*i* + 1) residue of type II β-turn structures that would be occupied by the substituted

proline. A corresponding energetic destabilization was not observed for C$^{\gamma}$-*exo* isomer of

(4*S*)-fluoroproline at similar values of the dihedral angle ψ. Taken together, these

computational data suggest that certain secondary structures, e.g., the type II β-turn, that place the value of the ψ angle for the (4S)-fluoroproline residue within this unfavorable region would energetically destabilize the C$^\gamma$-*endo* isomer versus the C$^\gamma$-*exo* isomer, thus negating the stereoelectronic influence of the fluorine substituent. The presence of (4S)-fluoroproline effectively raises the energy of the type II β-turn structures for the model peptide (MeCO-4S-F-Pro-Gly-NHMe) such that other turns, particularly the type I β-turn structure, become more similar in energy (Table 1). Note that the computational data indicate a trend in which the difference in energy between the β_{II} and β_I turns ($\Delta E_{II\text{-}I}$) decreases in the model peptide series (MeCO-Xaa-Gly-NHMe) from Xaa = 4R-F-Pro (2.8 kcal mol^{-1}) to Pro (1.7 kcal mol^{-1}) to 4S-F-Pro (0.3 kcal mol^{-1}). Thus, computational analyses of the model peptide segment (MeCO-Xaa-Gly-NHMe) suggest that the differences in spectroscopic and calorimetric behavior between the corresponding elastin derivatives, in particular, the discrepancy in transition temperature and variation in type II β-turn content observed between **elastin-3** vis-à-vis **elastin-2**, can be rationalized on the basis of differential stereoelectronic stabilization of the incipient type II β-turn structure as the temperature approaches the phase transition. While the calculated energetic differences between turn structures in the model system are relatively small for all three proline derivatives, this situation does not preclude the possibility that the relative

populations of the various turn structures under consideration might be constituted in a manner that can account for the observed differences in physical characteristics between **elastins-1**, **-2**, and **-3**.

To draw a closer correlation between the computational data and the experimental results derived from structural characterization of the series of elastin-mimetic peptides, we computed the values of the ^{15}N chemical shifts for the structurally sensitive Gly amide groups (cf. Experimental Section) of the model peptide segment (MeCO-Xaa-Gly-NHMe) (Xaa = Pro, 4S-F-Pro, and 4R-F-Pro) in different turn structures and ring pucker conformations (Table 1). While differences are observed between the calculated ^{15}N-Gly chemical shifts for the (MeCO-Xaa-Gly-NHMe) units and the experimental chemical shifts for structurally analogous Gly^3 of the elastin-mimetic polypeptides (due to presumed differences in the environment of the peptides, gas phase and aqueous solution, respectively), nonetheless the trends are consistent with predictions based on comparison of the calculated chemical shifts with empirical correlations for the amide groups of polypeptides.[43,44] Significantly, the computational data indicate that the type II β-turn structure of the (MeCO-Xaa-Gly-NHMe) unit typically results in a downfield shift of the amide ^{15}N signals for the (i + 2) Gly residue in comparison to the corresponding type I β-turn structure regardless of the ring pucker conformation (Table 1).

The MPW1PW91/6-311G* ^{15}N chemical shifts calculated for (MeCO-4R-F-Gly-NHMe) and (MeCO-Pro-Gly-NHMe), both presumed to sustain a type-II β-turn in analogy to **elastin-3** and **elastin-1**, are 113.1 (*exo*-F_{ax}) and 111.3$_{avg}$ ppm, respectively (Table 1). As the C$^{\gamma}$-*exo* and C$^{\gamma}$-*endo* conformers of (MeCO-Pro-Gly-NHMe) are isoenergetic, the latter value is an average of these two conformers. The predicted shifts are not only within 4-5 ppm of the measured values, but both the ordering and differences ($\Delta\delta(^{15}N)$ = 1.7(expt) and 1.8(calcd) ppm, respectively) are accurately modeled. The PBE functional predicts the values to be low-field shifted by an additional 10 ppm, but once again both the ordering and difference ($\Delta\delta(^{15}N)$ = 1.1 ppm) matches experiment. Both chemical shift estimates support the assignment of type II β-turn as the preferred structure for the VPGVG units of **elastin-1** and **elastin-3**.

In contrast, the computational data indicate that the significant upfield shift of the amide ^{15}N signals of the Gly^3 residue of **elastin-2** may occur as a consequence of a potentially higher population of alternative turn structures, in particular, the type I β-turn, that arise at the expense of the type II β-turn population. MPW1PW91/6-311G* calculations for the more stable (4S)-*endo*-F_{ax} conformer of the type-I β-turn predicted a sizable upfield ^{15}N shift of 100.5 ppm (Table 1, $\Delta\delta(^{15}N)$ = 11.0 (calcd) ppm) with respect to the corresponding type II β-turn structure. Although the calculated shift difference for

(MeCO-4S-F-Pro-Gly-NHMe) is more substantial than that observed experimentally between the elastin-mimetic polypeptides, the former assumes that the model structures define uniform populations of turn conformations. Computational analyses of the relative energetics of the turn structures for (MeCO-4S-F-Pro-Gly-NHMe) suggest that the type II β-turn becomes destabilized relative to alternative turn structures in comparison to the other proline derivatives, although it remains the most stable turn under the constraints of the analysis (Table 1). The calculations imply an altered equilibrium population of turn species for **elastin-2** relative to **elastin-1** and **elastin-3**, which may manifest itself in the upfield chemical shift of Gly^3 ^{15}N of **elastin-2** with respect to the other elastin derivatives. As suggested in the discussion of the ^{1}H-^{15}N HSQC NMR spectra of the elastin derivatives, the trend in observed Gly^3 ^{15}N NMR spectral shifts for the elastin-mimetic polypeptides indicates a decrease in type II β-turn content from **elastin-3** (108.8 ppm) to **elastin-1** (107.1 ppm) to **elastin-2** (104.8 ppm), which coincides qualitatively with the trend in relative energies between β_{II} and β_I turns calculated for the corresponding model peptides.

Conclusion

The experimental data derived from calorimetric and spectroscopic analyses of **elastin-1**, **elastin-2**, and **elastin-3** provide evidence that stereoelectronic effects may alter the self-assembly of elastin-mimetic polypeptides, and, by inference, native elastin, through their influence on the local conformation parameters of the β-turn structures that develop in the structural repeats above the phase transition. Moreover, the observed differences in macromolecular behavior between the elastin derivatives reinforce the hypothesis that type II β-turn formation plays an important role in elastin assembly, as previously postulated from spectroscopic and computational analyses of elastin-mimetic polypeptides.[12-17] For the elastin-mimetic polypeptides under consideration in this study, structural factors that increase the stability of the type II β-turn structure of the (VPGVG) unit, i.e., stereoelectronic stabilization due to incorporation of (4*R*)-fluoroproline into **elastin-3**, resulted in a lower transition temperature for elastin assembly. Conversely, structural factors that decrease the stability of the type II β-turn structure of the (VPGVG) unit, such as incorporation of (4*S*)-fluoroproline into **elastin-2**, raised the transition temperature of elastin assembly.

Computational analyses of the conformational energetics associated with various

turn structures (β_I, β_{II}, inverse γ) of the model peptide (MeCO-Xaa-Gly-NHMe) unit (Xaa

= Pro, 4S-F-Pro, 4R-F-Pro) suggest that the relative stability of the conformers depends

on the effect of the proline derivative on the metrical parameters associated with the local

polypeptide chain conformation. The (C$^\gamma$-exo/C$^\gamma$-$endo$) ring pucker equilibria for the

various turn structures are strongly influenced by stereoelectronic effects due to fluorine

substitution and by steric interactions between the pyrrolidine ring substituents. The

greatest energetic stabilization is observed under conditions in which the stereoelectronic

interaction reinforces the preferred peptide chain conformation, i.e., the C$^\gamma$-exo pucker

within type II β-turns of the model peptide (MeCO-4R-F-Pro-Gly-NHMe).

Significant energetic destabilization occurs under conditions in which the

stereoelectronically preferred ring pucker opposes the "preferred" peptide chain

conformation, i.e., the C$^\gamma$-$endo$ pucker within type II β-turns of the model peptide

(MeCO-4S-F-Pro-Gly-NHMe). Similar results have been observed for 4-fluoroproline

substitution in collagen-mimetic peptides in which structural stabilization was observed

under conditions in which the ring pucker conformation of the fluoroproline epimer

reinforced the local conformation of the polypeptide backbone of the collagen peptide in

the triple helix structure.[2-5]

Our results indicate that fluoroproline substitution may be employed as a

mechanism to interrogate local conformational effects that arise due to the presence of

proline residues in polypeptide sequences, particularly in situations in which preferences

have been observed between particular ring pucker conformations and secondary

structure elements.[25] This investigation suggests that turn structures may be included

among local peptide structural elements that can be influenced through stereoelectronic

control. As β-turn structures can exert a powerful influence to mediate protein folding

events, stereoelectronic differences between proline analogues may be employed to

address structural questions regarding the specific roles of proline-containing turn

sequences in native protein conformations. For example, the "Pro-Gly" motif is a

recurrent if not dominant structural feature associated with a number of native protein

materials displaying elastomeric behavior.[51-54] The molecular architectures of these

protein-based elastomers differ from those of conventional synthetic elastomers, which

suggests that different structural mechanisms may underlie their elastomeric behavior.

Moreover, mechanical studies of the native protein materials have indicated that the

viscoelastic properties associated with the respective polypeptide sequences have been

evolutionarily optimized for distinct biological function.[55] An understanding of the

structural factors that determine the differences in mechanical behavior among native

protein elastomers should provide criteria for the design of novel elastomeric materials in

which the sequences can be tailored for specific technological applications. In addition,

the mechanism of elasticity of elastin remains a subject of scientific disagreement,

particularly as regards the origin of the elastomeric restoring force.[56,57] The elastin

structural variants reported in this study may be useful materials to discriminate between

the relative contributions of main-chain versus solvation entropy to the elastomeric

restoring force and provide information relevant to the debate over the mechanism of

elasticity.

References

[1]   DeRider, M.L.; Wilkens, S.J.; Waddell, M.J.; Bretscher, L.E.; Weinhold, F.; Raines, R.T.; Markley, J.L. *J. Am. Chem. Soc*. **2002**, *124*, 2497-2505.

[2]   a) Hodges, J.A.; Raines, R.T. *J. Am. Chem. Soc*. **2003**, *125*, 9262-9263. b) Bretscher, L.E.; Jenkins, C.L.; Taylor, K.M.; DeRider, M.L.; Raines, R.T. *J. Am. Chem. Soc*. **2001**, *123*, 777-778. c) Holmgren, S.K.; Bretscher, L.E.; Taylor, K.M.; Raines, R.T. *Chem. Biol*. **1999**, *6*, 63-70. d) Holmgren, S.K.; Taylor, K.M.; Bretscher, L.E.; Raines, R.T. *Nature*. **1998**, *392*, 666-667.

[3]   a) Barth, D.; Milbradt, A.G.; Renner, C.; Moroder, L. *Chembiochem*. **2004**, *5*, 79-86. b) Barth, D.; Musiol, H.M.; Schutt, M.; Fiori, S.; Milbradt, A.G.; Renner, C.; Moroder, L. *Chemistry* **2003**, *9*, 3692-702.

[4]   Persikov, A.V.; Ramshaw, J.A.; Kirkpatrick, A.; Brodsky, B. *J. Am. Chem. Soc*. **2003**, *125*, 11500-11501.

[5]   Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *J Am Chem Soc*. **2003**, *125*, 9922-3.

[6]   a) Umashankara, M.; Babu, I.R.; Ganesh, K.N. *Chem. Commun*. **2003**, 2606-2607. b) Babu, I.R.; Ganesh, K.N. *J. Am. Chem. Soc*. **2001**, *123*, 2079-2080.

[7] Berg, R.A.; Prockop, D.J *Biochem. Biophys. Res. Commun*. **1973**, *52*, 115-120.

[8] Renner, C.; Alefelder, S.; Bae, J.H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem. Int. Ed. Engl*. **2001**, *40*, 923-925.

[9] Rosenbloom, J.; Abrams, W.R.; Mecham, R. *FASEB J*. **1993**, *7*, 1208-1218.

[10] Langer, R.; Tirrell, D.A. *Nature* **2004**, *428*, 487-492. 35

[11] Urry, D.W.; in *Protein-Based Materials*. (Eds. K.P. McGrath, D. Kaplan), Birkhauser, Boston, **1997**, pp 133-177.

[12] Urry, D.W.; Shaw, R.G.; Prasad, K.U. *Biochem. Biophys. Res. Commun*. **1985**, *130*, 50-57.

[13] Thomas, Jr., G.J.; Prescott, B.; Urry, D.W. *Biopolymers* **1987**, *36*, 921-934.

[14] Urry, D.W.; Krishna, N.R.; Huang, D.H.; Trapane, T.L.; Prasad, K.U. *Biopolymers* **1989**, *28*, 819-833.

[15] Reiersen, H.; Clarke, A.R.; Rees, A.R. *J. Mol. Biol*. **1998**, *283*, 255-264.

[16] Li, B.; Alonso, D.O.; Daggett, V. *J. Mol. Biol*. **2001**, *305*, 581-592.

[17] Yao, X.L.; Hong, M.. *J. Am. Chem. Soc*. **2004**, *126*, 4199-4210.

[18] Kim, W.; George, A.; Evans, M.E.; Conticello, V.P. *ChemBioChem* **2004**, *5*, 928-936.

[19] a) Lankin, D. C.; Chandrakumar, N. S.; Rao, S. N.; Spangler, D. P.; Snyder, J. P.

J. Am.Chem. Soc. **1993**, *115*, 3356-3357; b) Snyder, J. P.; Chandrakumar, N. S.; Sato, H.; Lankin, D. C. *J. Am. Chem. Soc.* **2000**, *122*, 544-545; c) Lankin, D. C.; Grunewald, G. L.; Romero, F. A.; Oren, I. Y.; Snyder, J. P. *Org. Lett.* **2002**, *4*, 3557-3560; d) Sun, A.; Lankin, D. C.; Hardcastle, K.; Snyder, J. P. *Chem. Eur. J.* **2005**, *11*, 1579-1591.

[20]   a) O'Hagan, D.; Bilton, C.; Howard, J.A.K.; Knight, L.; Tozer, D.J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605-607. b) Briggs, C.R.S.; O'Hagan, D.; Howard, J.A.K.; Yufit, D.S. *J. Fluorine Chem.* **2003**, *119*, 9-13.

[21]   a) Eberhardt, E.S.; Panasik, Jr., N.; Raines, R.T. *J. Am. Chem. Soc.* **1996**, *118*, 12261- 12266. b) Panasik, Jr., N.; Eberhardt, E.S.; Edison, A.S.; Powell, D.R.; Raines, R.T. *Int.J. Pept. Protein Res.* **1994**, *44*, 262-269.

[22]   a) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568-12577. b) Improta, R.; Mele, F.; Crescenzi, O.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2002**, *124*, 7857-7865. c) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341-350.

[23]   Mooney, S.D.; Kollman, P.A.; Klein, T.E. *Biopolymers*. **2002**, *64*, 63-71.

[24]   Cook, W.J.; Einspahr, H.; Trapane, T.L.; Urry, D.W.; Bugg, C.E. *J. Am. Chem. Soc.* **1980**, *102*, 5502-5505.

[25]   Milner-White, E.J.; Lachlan, H.B.; Maccallum, P.H. *J. Mol. Biol.* **1992**, *228*, 725-734.

[26]   Urry, D.W. *Trends Biotechnol*. **1999**, *17*, 249-257.

[27]   Welsh, E.R.; Tirrell, D.A. *Biomacromolecules* **2000**, *1*, 23-30.

[28]   Keeley, F.W.; Bellingham, C.M.; Woodhouse, K.A. *Philos Trans R Soc Lond B Biol Sci.***2002**, *357*, 185-189.

[29]   Mithieux, S.M.; Rasko, J.E.; Weiss, A.S. *Biomaterials* **2004**, *25*, 4921-4927.

[30]   Wright, E.R.; Conticello, V.P. *Adv Drug Deliv Rev*. **2002**, *54*, 1057-1073.

[31]   Chilkoti, A.; Dreher, M.R.; Meyer, D.E. *Adv Drug Deliv Rev*. **2002**, *54*, 1093-1111.

[32]   Meyer, D.E.; Chilkoti, A. *Nat Biotechnol*. **1999**, *17*, 1112-1115.

[33]   Luan, C.-H.; Harris, R.D.; Prasad, K.U.; Urry, D.W. *Biopolymers* **1990**, *29*, 1699-1706.

[34]   Makhatadze, G.I.. *Biophys. Chem*. **1998**, *71*, 133-156.

[35]   Cooper, A. *Biophys Chem*. **2000**, *85*, 25-39.

[36]   a) Urry, D.W.; Luan, C.-H.; Parker, T.M.; Gowda, D.C.; Prasad, K.U.; Reid, M.C.; Safavy, A. *J. Am. Chem. Soc.* **1991**, *113*, 4346-4347. b) Urry, D.W.; Gowda, D.C.; Parker, T.M.; Luan, C.-H. *Biopolymers* **1992**, *32*, 1243-1250.

[37]   Meyer, D.E.; Chilkoti, A. *Biomacromolecules* **2004**, *5*, 846-851.

[38]   Yamaoka, T.; Tamura, T.; Seto, Y.; Tada, T.; Kunugi, S.; Tirrell, D.A. *Biomacromolecules* **2003**, *4*, 1680-1685.

[39]   Andreotti, A.H. *Biochemistry.* **2003**, *42*, 9515-9524.

[40]   Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J. Mol.Biol.* **1998**, *279*, 449-460.

[41]   Wishart, D.S.; Nip, A.M. *Biochem. Cell. Biol.* **1998**, *76*, 153-163.

[42]   Le, H.; Oldfield, E. *J Biomol NMR* **1994**, *4*, 341-348.

[43]   Provencher, S.W.; Glockner, J. *Biochemistry* **1981**, *20*, 33–37.

[44]   Arad, O.; Goodman, M. *Biopolymers* **1990**, *29*, 1652-1668.

[45]   Chesnut, D. B.; Byrd, E.F.C. *Chem. Phys.* **1996**, *213*, 153-158.

[46]   Wilmot, C.M.; Thornton, J.M. *J Mol Biol.* **1988**, *203*, 221-232.

[47]   Bondi, A. *J. Phys Chem.* **1964**, *68*, 441-451.

[48]   a) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Prot. Chem.* **1985**, *37*, 1-109; b) Milner-White, E. J., Ross, B. M., Ismail, R., Belhadj-Mastefa, K., Poet, R. *J. Mol. Biol.* **1988**, *204*, 777-782.

[49]   a) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liscamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**,

11, 440-467; b) http://www.schrodinger.com/Products/macromodel.html.

[50]   Gaussian 03, Revision C.02, Pople, J. A., et al. Gaussian, Inc., Wallingford CT, **2004**.

[51]   Brender, J. R.; Taylor, D. M.; Ramamoorthy, A. *J. Am. Chem. Soc.* **2001**, *123*, 914-922.

[52]   a) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664-675; b) Cimino, P.; Gomez-Paloma, L.; Duca, D.; Riccio, R.; Bifulco, G. *Magn. Res. Chem.* **2004**, *42*, S26-S33.

[53]   Adamo, C.; Barone, V. *Chem. Phys. Lett.* **1998**, *298*, 113-119.

[54]   a) Wright, E.R.; McMillan, R.A.; Cooper, A.; Apkarian, R.P.; Conticello, V.P. *Adv. Funct. Mater.* **2002**, *2*, 149-154. b) Lee, T.A.T.; Cooper, A.; Apkarian, R.P.; Conticello, V.P. *Advanced Materials* **2000**, *12*, 1105-1110.

CHAPTER 4

Fluoroproline Flip-Flop: Regiochemical Reversal of a

Stereoelectronic Effect on Peptide and Protein Structure.

Introduction

The de novo design of protein structure relies on the ability to uniquely define a polypeptide chain conformation through introduction of specific stabilizing interactions into the amino acid sequence, which act cooperatively to overcome the loss of conformational entropy associated with the protein folding process. Within this concept, the stereoelectronic effect has recently emerged as a means to restrict the local chain conformation of polypeptide sequences and to modulate the thermodynamic stability of secondary and supersecondary structural elements.[1-5]

Substituted proline residues are particularly susceptible to the influence of stereoelectronic effects that alter the conformational energetics of the pyrrolidine ring. Introduction of an electronegative substituent (X = N, O, F) at the C3 or C4 position of the pyrrolidine ring establishes a vicinal N-C-C-X arrangement between the prolyl amide group and the electronegative atom.[6,7] A preference for a *gauche* stereochemical relationship is observed between the two substituents as a result of hyperconjugative delocalization, in which the magnitude of the effect depends on the electronegativity of the substituent. These stereoelectronic interactions strongly influence the equilibrium conformational population of the pyrrolidine ring-pucker isomers, such that epimeric

pairs of substituted proline derivatives have opposite ring-pucker preferences and often display antagonistic effects on protein stability that depend on structural context.[1,8,9] Most comparative analyses have employed the (4*S*)- and (4*R*)-fluoroproline epimeric pair, in which strong conformational preferences are observed for the C$^\gamma$-*endo* and C$^\gamma$-*exo* puckers, respectively. Herein, we report a structural comparison between the (3*S*)- and (3*R*)-fluoroproline epimers, in which the altered regiochemistry results in a reversal of conformational preferences compared to the corresponding 4-fluoroproline epimers, which has structural implications for protein design and engineering.

Experimental Section

Materials and Methods.

All chemical reagents were purchased from either Fisher Scientific, Inc. (Pittsburgh, PA) or Sigma Chemical Co. (St. Louis, MO) unless otherwise noted. *N-tert*-butoxycarbonyl-(2*R*,3*S*)-3-fluoroproline and *N-tert*-butoxycarbonyl-(2*R*,3*S*)-3-fluoroproline were synthesized from (2*S*,3*S)*-3-hydroxyproline (Acros Organics, Inc.) using a modification of the method of Demange et al.[10] The preparation of the *E. coli* expression strain CAG18515(*proA*3096::Tn10Kan)[pWK1] and the plasmid, pAG2, that encodes the elastin-mimetic polypeptide sequence have been previously described.[11] NMM medium was prepared according to the protocol of Budisa et al,[12] with the exception that proline was not added to the medium prior to cell culture. TALON metal affinity resin was purchased from BD Biosciences, Inc. Protein electrophoresis was performed on 10-15% gradient discontinuous SDS polyacrylamide gels on a PhastSystem from Amersham Pharmacia Biotech and was visualized via a silver staining procedure. Aqueous solutions of **elastin-4** and **elastin-5** were prepared from lyophilized specimens of the purified proteins that were dissolved at the appropriate concentration in distilled, deionized water at 4 °C. Quantitative amino acid analysis was performed on aliquots of

protein stock solutions (circa 0.2–0.3 mg/mL) at the W. M. Keck Foundation

Biotechnology Resource Laboratory of Yale University to provide accurate

concentrations for CD spectroscopic experiments and DSC measurements.

Physical and Analytical Measurements.

Suitable crystals of (**1**) and (**2**) were coated with Paratone N oil, suspended in

small fiber loops and placed in a cooled nitrogen gas stream at 173 K on a Bruker D8

APEX II CCD sealed tube diffractometer with graphite monochromated CuK_{α} (1.54178

Å) radiation. Data were measured using a series of combinations of phi and omega scans

with 10 s frame exposures and 0.5° frame widths. Data collection, indexing and initial

cell refinements were all carried out using APEX II[13] software. Frame integration and

final cell refinements were done using SAINT[14] software. The final cell parameters were

determined from least-squares refinement on 3071 and 2727 reflections, respectively. The

structures were solved using Direct methods and difference Fourier techniques

(SHELXTL, V6.12).[15] Hydrogen atoms were placed at their expected chemical positions

using the HFIX command and were included in the final cycles of least squares with

isotropic U_{ij} 's related to the atom's ridden upon. All non-hydrogen atoms were refined

anisotropically. Scattering factors and anomalous dispersion corrections are taken from

the *International Tables for X-ray Crystallography*.[16] Structure solution, refinement, graphics and generation of publication materials were performed by using SHELXTL, V6.12 software. The crystallographic data for (**1**) and (**2**) were deposited to the Cambridge Crystallographic Data Centre (CCDC) with registration numbers 288114 and 288115. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre, 12, Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033. Crystal data for (**1**): $C_8H_{12}FNO_3$, $M = 189.18$; orthorhombic, space group P2(1)2(1)2(1); a = 7.0907(2) Å, b = 9.8318(4) Å, c = 12.9324(5) Å; V = 901.57(6) Å^3; T = 173(2) K; Z = 4; CuKα 1.54178 Å; reflections: total = 4224, independent = 1201 (R_{int} = 0.0198); R_1 = 0.0216, wR_2 = 0.0532 for 1201 observed data [$I > 2\sigma(I)$]. Crystal data for (**2**): $C_8H_{12}FNO_3$, $M = 189.18$, a = 7.555(2) Å, b = 9.550(2) Å, c = 13.080(3) Å; V = 943.8(4) Å^3; T = 173(2) K; Z = 4; CuKα 1.54178 Å; reflections: total = 3638, independent = 1128 (R_{int} = 0.2104); R_1 = 0.0830, wR_2 = 0.2187 for 1128 observed data [$I > 2 \sigma (I)$].

ESI mass spectra were acquired on a JEOL JMS-SX 102/SX 102 A/E mass spectrometer. Molar masses of elastin analogues were determined by MALDI-TOF MS

on an Applied Biosystems Voyager System 428 mass spectrometer in the positive linear mode. The matrices, 2-(4-hydroxyphenylazo)benzoic acid (HABA) or 4-hydroxy-3-methoxycinnamic acid, were used at a concentration of 10 mg/ml in a 50:50 mixture of water and 2-propanol. The protein solution (1 mg/ml in distilled water) was mixed with the matrix solution in a ratio of 1:10 and dried under vacuum or air. Bovine serum albumin was used as a standard for external calibration.

The inverse temperature transitions of the elastin polypeptides were monitored as a function of temperature using an ultra-sensitive differential scanning calorimeter (VP-DSC MicroCal, LLC, Northampton, MA). Proteins samples were dissolved in distilled, deionized water at 4 °C at concentrations ranging from 0.5 to 2 mg/mL, degassed under dynamic vacuum and scanned from 5-60 °C at a rate of 60 deg/hr. DSC data were processed using the program Origin (MicroCal, LLC, Northampton, MA).

Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a PFD-425S Peltier temperature control unit in 1 mm sealed quartz cells. Proteins were dissolved in distilled, deionized water at a concentration of 5.4 μM for **elastin-4** and 9.1 μM for **elastin-5**. Temperature/wavelength

CD-scans were performed within the temperature range from 5 °C to 65 °C with equilibration for 5 min at each temperature. Spectra were obtained from 260 to 190 nm at a resolution of 0.2 nm and at a scanning rate of 50 nm per min. The CD curves represented the average of five measurements and were smoothed using the means-movement method on the interval analysis of the spectral manager program. CD data are reported as mean residue ellipticity ($[\theta]$, deg cm^2 dmol^{-1}) in which the molar masses of the polypeptides, **elastin-4** and **elastin-5**, were calculated on the basis of on complete substitution of the canonical proline residues with the respective amino acid analogue.

NMR spectra were acquired on either a Varian INOVA 400 (^{1}H, 399.94 MHz; ^{13}C, 100.57 MHz), a Varian Unity (^{19}F, 564.044 MHz) or a Varian INOVA 600 (^{1}H, 599.74 MHz; ^{13}C, 150.82 MHz). Chemical shifts for ^{1}H NMR and ^{13}C spectra in organic solvents were reported in ppm and were referenced to the solvent signals and reported relative to tetramethylsilane (0.0 ppm). Chemical shifts for ^{1}H NMR spectra in aqueous solution were reported in ppm and were referenced and reported relative to internal sodium 2,2-dimethyl-2-silapenta-5-sulfonate (0.0 ppm). Protein samples were prepared for NMR analysis by dissolving the polypeptide in a H$_2$O/D$_2$O (70:30) mixture at a concentration of 10 mg/mL. The pH of the specimens was adjusted to 2.7 to retard amide

proton exchange on the NMR time scale. The NMR spectra of polypeptide samples were acquired at 4 ℃. Standard solvent suppression techniques were employed to reduce signal due to the residual protons of H_2O in the ^{1}H NMR of aqueous solutions of the polypeptides. Chemical shifts for the ^{19}F NMR spectra are referenced and reported relative to external sample of aqueous (10 % v/v) trifluoroacetic acid (0.0 ppm). Two-dimensional ^{1}H-^{1}H NOESY NMR spectra were acquired in phase-sensitive mode using the hypercomplex method with a mixing time of 200 ms at a spectral width of 6799.8 Hz. Spectra were collected with 512 t_1 increments and 2048 complex data points with 32 scans. The two-dimensional NMR data were further processed using the program NutsPro (Acorn NMR, Inc.). The equilibrium constants, $K_{trans/cis}$, were calculated from ration of integration of the well-resolved resolved H^β resonances in the ^{1}H NMR spectra of (**1**) and (**2**).

Chemical synthesis of model compounds (1) and (2) and free amino acids

Synthesis of *N-tert*-butoxycarbonyl-(2*R*,3*R*)-3-fluoroproline methyl ester: *N-tert*-butoxycarbonyl-(2*R*,3*R*)-3-fluoroproline (0.80 g, 3.43 mmol) was dissolved in 15 mL of methanol. A solution of diazomethane (16.6 mmol) in diethyl ether was added dropwise

to the reaction mixture, which was stirred for 30 min at ambient temperature. The solvent

was removed by rotary evaporation and crude product was purified by flash

chromatography (ethyl acetate:hexanes, 1:1) to give 0.82 g (96.8 %) of the product as a

colorless oil. ^{1}H NMR (400 MHz, CDCl$_3$): δ 1.44 and 1.49 (s, 9H), 1.96-2.3 (m, 2H),

3.58-3.78 (m, 2H), 3.77 (s, 3H), 4.51 (dd, J = 26.8 Hz, 0.6H), 4.58 (dd, J = 27.6 Hz,

0.4H), 5.29-5.31 and 5.43-5.46 (m, 1H); ESI-MS (m/z): Calc. for C$_{11}$H$_{18}$FNO$_4$, 247.26;

Obs., 248.13 (M+H).

Synthesis of *N-tert*-butoxycarbonyl-(2*R*,3*S*)-3-fluoroproline methyl ester: *N-tert*-

butoxycarbonyl-(2*R*,3*S*)-3-fluoroproline (1.13 g, 4.8 mmol) was converted to the methyl

ester using the procedure described above. Flash chromatography (ethyl acetate:hexanes,

1:1) afforded 1.19 g (99 %) of the product as a yellow oil. ^{1}H NMR (400 MHz, CDCl$_3$):

δ : 1.42 and 1.48 (s, 9H), 2.1-2.24 (m, 2H), 3.48-3.81 (m, 2H), 3.77 (s, 3H), 4.48 (d, J =

23.2 Hz, 0.6H), 4.60 (d, J = 22.8 Hz, 0.4H), 5.1 and 5.23 (br s, 1H); ESI-MS (m/z): Calc.

for C$_{11}$H$_{18}$FNO$_4$, 247.26; Obs., 248.13 (M+H).

Synthesis of (2*R*,3*R*)-3-fluoroproline methyl ester: *N-tert*-butoxycarbonyl-(2*R*,3*R*)-3-

fluoroproline methyl ester (0.82 g, 3.3 mmol) was dissolved in 10 mL of trifluoroacetic

acid/methylene chloride (1:2, v/v), and stirred for 2 h at ambient temperature. The

compound was concentrated by rotary evaporation and dried under high vacuum to afford

the product as a dark red oil (0.84 g, 96.4 %). ^{1}H NMR(400 MHz, CDCl$_3$): δ 2.33-2.57

(m, 2H), 3.64-3.87 (m, 2H), 3.92 (s, 3H), 4.84 (dd, J = 31.6 Hz, 1H), 5.5 and 5.64 (br s,

1H); ESI-MS (m/z): Calc. for C$_6$H$_{10}$FNO$_2$, 147.15; Obs., 148.08 (M+H).

Synthesis of (2R,3S)-3-fluoroproline methyl ester: This compound was synthesized

from N-$tert$-butoxycarbonyl-(2R,3S)-3-fluoroproline methyl ester (1.19 g, 4.8 mmol)

utilizing the procedure described above for (2R,3R)-3-fluoroproline methyl ester. The

product was obtained as a dark red oil (1.22 g, 97 %). ^{1}H NMR (400 MHz, CDCl$_3$): δ

2.08-2.56 (m, 2H), 3.8-3.89 (m, 2H), 3.92 (s, 3H), 4.79 (d, J = 23.2 Hz, 1H), 5.51 (d, J =

52.0 Hz, 1H); ESI-MS (m/z): Calc. for C$_6$H$_{10}$FNO$_2$, 147.15; Obs., 148.08 (M+H)

Synthesis of N-acetyl-(2R,3R)-3-fluoroproline methyl ester (1): (2R,3R)-3-

fluoroproline methyl ester (0.60 g, 4.1 mmol) was dissolved in neat acetic anhydride (10

mL) and the mixture was stirred for 5 h at ambient temperature. The solvent was removed

rotary evaporation to afford a crude product that was purified by flash chromatography

(ethyl acetate:hexane, 2:1, followed by 100 % ethyl acetate). The product was obtained as

a colorless oil (0.55 g, 70.0 %) that crystallized upon drying under high vacuum. Crystals

suitable for X-ray diffraction analysis were grown from ethyl acetate solution by slow evaporation at ambient temperature. ^{1}H NMR (400 MHz, CDCl$_3$): δ 1.89 and 2.08 (s, 3H), 2.02-2.35 (m, 2H), 3.64-3.79 (m, 2H), 3.73 and 3.77 (s, 3H), 4.61 (dd, J = 25.2 Hz, 0.2H), 4.65 (dd, J = 22.0 Hz, 0.8H), 5.26-5.29 and 5.39-5.43 (m, 0.8H), 5.35-5.38 and 5.48-5.52 (m, 0.2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 22.25 (21.92), 32.61 (J_{C-F} = 21.9 Hz) (30.82 (J_{C-F} = 21.5 Hz)), 44.17 (45.76), 52.49 (52.89), 63.56 (J_{C-F} = 22.1 Hz) (64.80 (J_{C-F} = 22.9 Hz)), 91.80 (J_{C-F} = 183.2 Hz) (93.70 (J_{C-F} = 185.2 Hz)), 167.21 (168.0), 169.46 (169.7); ESI-MS (m/z): Calc. for C$_8$H$_{12}$FNO$_3$, 189.18; Obs., 190.09 (M+H); Elemental analysis: Calc. for C$_8$H$_{12}$FNO$_3$, C: 50.79, H: 6.39, N: 7.40, found. C: 50.51, H: 6.41, N: 7.31. $[\alpha]_D^{20°C}$: -180 (c = 1, H$_2$O).

Synthesis of *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester, (2): This compound was synthesized from (2*R*,3*S*)-3-fluoroproline methyl ester (0.65 g, 2.5 mmol) using the procedure described above for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester. The product was obtained as a pale yellow oil (0.40 g, 84 %) that crystallized upon drying under high vacuum. Crystals suitable for X-ray diffraction analysis were grown by slowly cooling a concentrated solution of (**2**) in ethyl acetate to 4 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 2.23 and 2.14 (s, 3H), 2.05-2.39 (m, 2H), 3.68-3.83 (m, 2H), 3.77 and 3.81 (s,

3H), 4.60 (d, J = 19.2 Hz, 0.2H), 4.78 (d, J = 25.2 Hz, 0.8H), 5.19 (dd, J = 51.2 Hz, 0.8H),

5.33 (dd, J = 51.6 Hz, 0.2H); ^{13}C NMR (100 MHz, CDCl$_3$): δ 22.26 (22.48), 31.73 ($J_{C\text{-}F}$ =

21.5 Hz) (29.81 ($J_{C\text{-}F}$ = 21.1 Hz)), 45.53 (44.40), 52.97 (53.34), 65.74 ($J_{C\text{-}F}$ = 24.4 Hz)

(67.19 ($J_{C\text{-}F}$ = 24.4 Hz)), 93.60 ($J_{C\text{-}F}$ = 185.3 Hz) (95.42 ($J_{C\text{-}F}$ = 185.6 Hz)), 169.30

(169.47), 170.02; ESI-MS (m/z): Calc. for C$_8$H$_{12}$FNO$_3$, 189.18; Obs., 190.09 (M+H);

Elemental analysis: Calc. for C$_8$H$_{12}$FNO$_3$, C: 50.79, H: 6.39, N: 7.40, found. C: 50.80, H:

6.36, N: 7.38. $[\alpha]_D^{20^\circ C}$: -86 (c = 1, H$_2$O).

Synthesis of (2R,3R)-3-fluoroproline: N-$tert$-butoxycarbonyl-(2R,3R)-3-fluoroproline

(1.0 g, 4.2 mmol) was dissolved in 24 mL of TFA/CH$_2$Cl$_2$ (1:2, v/v), and stirred for 1 h at

ambient temperature. The progress of the reaction was monitored by thin-layer

chromatography (ethyl acetate:methanol:acetic acid, 99:1:0.01) and visualized by staining

with ninhydrin. After disappearance of the starting material, the mixture was

concentrated by rotary evaporation and triturated with diethyl ether. The residual solid

was filtered and washed with ethyl ether and dried under vacuum. The solid was

recrystallized from ethyl acetate and MeOH affording the product as colorless needles

(0.70 g, 2.8 mmol, 65.8 %) that corresponded to a 3:1 adduct of amino acid to

trifluoroacetic acid on the basis of elemental analysis. ^{1}H NMR (400 MHz, D$_2$O): δ 2.24-

2.54 (m, 2H), 3.53-3.67 (m, 2H), 4.39 (dd, J = 32.8 Hz, 1H), 5.48-5.50 and 5.61-5.63 (m,

1H); ^{13}C NMR (150 MHz,D$_2$O): δ 31.36 (31.49), 43.86, 66.38 (66.54), 92.54 (93.71),

168.49; DIOS MALDI-TOF MS (m/z): Calc. for C$_5$H$_8$FNO$_2$, 133.12; Obs., 133.50

Elemental analysis (3:1, amino acid:trifluoroacetic acid adduct): Calc. for C$_{17}$H$_{25}$F$_6$N$_3$O$_8$,

C: 39.77, H: 4.91, N: 8.18, found. C: 40.07, H: 4.96, N: 8.27. $[\alpha]_D^{20°C}$: -73 (c = 1, H$_2$O).

Synthesis of (2R,3S)-3-fluoroproline: *N-tert*-butoxycarbonyl-(2R,3S)-3-fluoroproline

(0.90 g, 3.8 mmol) was deprotected using the procedure described above. The solid was

recrystallized from ethyl acetate and MeOH affording colorless needles of the product

(0.77 g, 3.1 mmol, 81.6 %). The product was isolated as a 1:1 adduct with a molecule of

trifluoroacetic acid on the basis of elemental analysis. ^{1}H NMR (400MHz, D$_2$O): δ 2.02-

2.48 (m, 2H), 3.52-3.60 (m, 1H), 3.67-3.73 (m, 1H), 4.64 (d, J = 22.0 Hz, 1H), 5.59 (dd, J

= 51.2 Hz, 1H); ^{13}C NMR (150 MHz, D$_2$O): δ 30.23 (30.37), 44.27, 67.15 (67.30), 94.81

(96.01), 169.13; DIOS MALDI-TOF MS (m/z): Calc. for C$_5$H$_8$FNO$_2$, 133.12; Obs.,

133.44; Elemental analysis (1:1, amino acid:trifluoroacetic acid adduct): Calc. for

C$_7$H$_9$F$_4$NO$_4$ C: 34.02, H: 3.67, N: 5.67, found. C: 34.23, H: 3.62, N: 5.67. $[\alpha]_D^{20°C}$: -7 (c

= 1.2, H$_2$O).

Protein Expression and Purification.

The expression vector pAG2–encoding elastin-mimetic protein sequence–was transformed into the auxotrophic *E. coli* strain CAG18515[pWK1] to generate the expression host employed for these studies. Single colonies of the expression strains were inoculated into sterile LB broth (50 mL) supplemented with the appropriate antibiotics (100 μg/mL ampicillin and 34 μg/mL chloramphenicol) as required for plasmid maintenance. The overnight culture was centrifuged at 4000g for 10 min to isolate the cells, which were re-suspended in sterile NMM medium (1 L) supplemented with the appropriate antibiotics. The proline concentration was adjusted to 0.3 mM from a sterile 100× stock solution. The culture was incubated with agitation (225 rpm) at 37 °C until the OD_{600} reached between 0.8 and 1.0 absorbance units, the cells were collected by centrifugation at 4,000g for 10 min. The cell pellet was washed with cold (4 °C), sterile 0.9 % aqueous NaCl twice (2 × 100 mL) and resuspended in sterile NMM containing antibiotics but without proline supplementation. After incubation at 37 °C for 30 min, the proline analogues were added to a final concentration of 0.5 mM from sterile 100× stock solutions. An aliquot of aqueous 1.0 M IPTG was added to the cultures to a final concentration of 1 mM to induce expression of the elastin-mimetic protein. After a 3 h induction period, the cells were harvested by centrifugation at 4,000g and 4 °C for 20 min.

The cell pellets were resuspended in lysis buffer (50 mL, 50 mM sodium phosphate, 300 mM NaCl, pH 7.0) and stored at -80 °C. The frozen cells were lysed by three freeze/thaw cycles. Lysozyme (1 mg/mL), EDTA-free protease inhibitor cocktail, benzonase (25 units/mL), and MgCl$_2$ (1 mM) were added to the lysate and the mixture was incubated with shaking at 4 °C overnight. The cell lysate was centrifuged at 40,000g for 30 min at 4 °C. Supernatant and pellet were separated and analyzed by SDS-PAGE to determine the location of the target protein. **Elastin-4** was localized in the supernatant fraction and was loaded directly onto cobalt charged TALON resin (5 mL) and washed of lysis buffer (50 mL) containing 20 mM imidazole. The target protein was eluted with elution buffer (20 mL, 50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole, pH 7.0) and dialyzed (MWCO = 10 kDa) against distilled deionized water (5 × 4 L). The dialysate was lyophilized to produce a white spongy solid. **Elastin-5** was located in the cell pellet and was solubilized in denaturing lysis buffer (100 mL, 50 mM sodium phosphate, 300 mM NaCl, 6 M urea, pH 7.0). The resulting mixture was centrifuged at 40,000g and 4 °C for 30 min. SDS-PAGE analysis indicated that the majority of the target protein dissolved under these conditions. The soluble fraction was loaded onto cobalt-charged TALON resin (10 mL) that had been previously equilibrated with denaturing lysis buffer. The target protein was washed with denaturing lysis buffer (100 mL) containing 20 mM

imidazole and eluted with denaturing lysis buffer (40 mL) containing 250 mM imidazole.

The eluted target protein was dialyzed (MWCO = 10 kDa) against a diminishing urea

step gradient from 6 M to 1 M, and, subsequently, against distilled water (5 × 4 L).

Lyophilization of the dialysate produced the elastin-mimetic polypeptide as a white

spongy solid.

Results and Discussion

Analysis of *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester (1) and *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester (2)

To evaluate the effect of 3-fluoro substitution on pyrrolidine ring conformation, the epimeric proline derivatives **1** and **2** were synthesized and their structures were determined by single-crystal X-ray diffraction analysis (Figure 1). Prior structural investigations of substituted N-acetylproline methyl ester derivatives[6,17] have established that these compounds provide reliable structural models to assess the conformational preferences of the pyrrolidine ring. The structural data indicated that both **1** and **2** crystallized with a *trans* configuration of the prolyl-peptide bond; however, significant differences were observed between **1** and **2** with respect to the conformation of the substituted pyrrolidine ring. The values of the N-C^α-C^β-F torsion angles for **1** and **2** were 90.7 and -87.48°, respectively, which indicated a *gauche* arrangement of the vicinal fluorine and amide substituents, as expected for the strong fluorine-amide stereoelectronic interaction.[18] As observed for the 4-fluoroproline epimeric pair, the *gauche* interaction enforced opposing pyrrolidine ring puckers between **1** and **2**. The conformation of **1** corresponded to a C^β-*endo*/C^γ-*exo* arrangement (displacements from

C^δ-N-C^α mean plane: C^β, 0.112 Å; C^γ, 0.460 Å), whereas that of **2** corresponded to a C^β-*exo*/C^γ-*endo* arrangement (displacements from C^δ-N-C^α mean plane: C^β, 0.308 Å; C^γ, 0.230 Å). The relatively small displacement of C^β for **1** may result from steric hindrance as a result of close contacts between the fluorine atom and the carbon (C1) and oxygen (O1) atoms of the carbonyl group of the methyl ester, in which the nonbonding distances (d_{F-C} = 2.67 Å and d_{F-O} = 2.94 Å) were less than the sum of the corresponding van der Waals radii (3.17 and 2.99 Å, respectively).

Hodges et al. have reported computational analyses of the conformational energetics of **1** and **2**,[7] which indicate that the preferred ring pucker in each case corresponded to that observed in the respective crystal structure. However, the thermodynamic preference of **2** for the C^γ-*endo* isomer (estimated 97% population) was significantly greater than that of **1** for the C^γ-*exo* isomer (estimated 69% population). The steric congestion surrounding the C^β atom observed in the crystal structure of **1** may account for the weaker conformational preference in the latter case. The persistence of the C^γ-*endo* conformation of **2** in solution may be inferred from the very small value of $^3J_{H\alpha H\beta}$ ($\leq$ 1 Hz) in the ^{1}H NMR spectrum in D_2O, which is consistent with the corresponding value of 0.79 Hz that was calculated from the Karplus equation for the observed H^α-C^α-C^β-H^β dihedral angle of -92.68° in the crystal structure of **2**.

A.
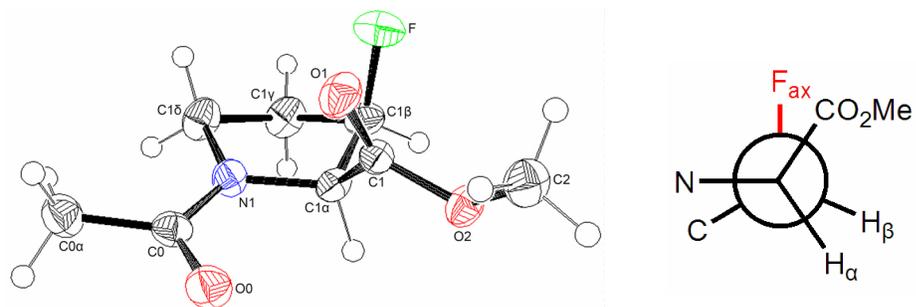
B.
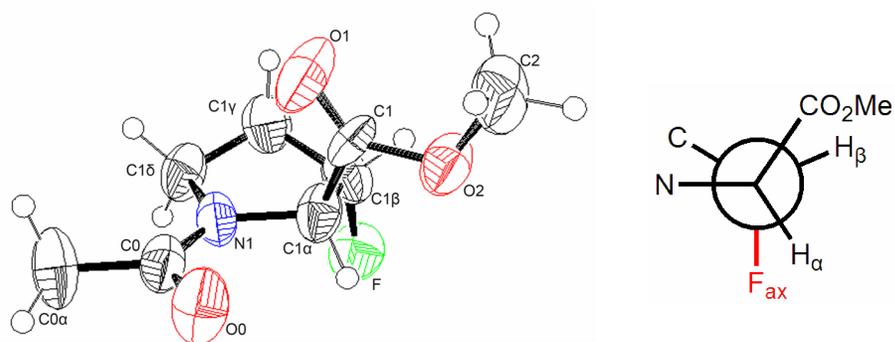
Figure 1. Crystallographically determined structures of (**A**) *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester (**1**) and (**B**) *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester (**2**) in conjunction with the Newman projections along the C^α-C^β bond vector, which depict the *gauche* relation between the amide and fluorine substituents.

The availability of crystallographic and computational data permits a comparison of the structural and thermodynamic parameters of the 3-fluoroproline (3-F-Pro) derivatives with the corresponding 4-F-Pro regioisomers (Table 1). Notably, the ring-pucker conformational preferences were reversed between *N*-Ac-F-Pro-OMe species that had similar *syn* (**1** and **3**) or *anti* (**2** and **4**) orientations of the *fluoro* substituent with respect to the fixed L-configuration of C^α. Thus, an *anti* orientation of the *fluoro* group resulted in a predominant C^γ-*endo* pucker for the 3-fluoroproline derivative **2** and a predominant C^γ-*exo* pucker for the 4-fluoroproline derivative **4**, whereas the opposite was true for the respective *syn*-oriented fluoroproline epimers **1** and **3** (Table 1).

Prior analyses of proteins within the structural database,[19] as well as small-molecule proline analogues,[6-8] suggested that the pyrrolidine ring pucker strongly influenced the main-chain torsion angles ϕ and ψ that define the peptide backbone conformation. Moreover, the main-chain dihedral angles appeared to be correlated with the pyrrolidine ring pucker. Thus, the C^γ-*exo* pucker of **4** displayed a less negative value of the angle ϕ and a smaller value of the angle ψ than did the C^γ-*endo* pucker of **3** (Table 1). A similar, although less pronounced, difference was observed between the corresponding torsional angles for **1** and **2**. The observed values of ϕ and ψ for the C^γ-*exo* ring pucker were rationalized on the basis of an energetic stabilization associated with a

Ac-FPro-OMe	$K_{trans/cis}$[a]	$\Delta E_{exo/endo}$ [kcal mol^{-1}][b]	ϕ [°][c]	ψ [°][c]
(3R)-**1** (*syn,exo*)	8.9	−0.48	−56.4	151.5
(3S)-**2** (*anti,endo*)	4.3	2.13	−71.2	158.4
(4S)-**3** (*syn,endo*)	2.5	0.61	−76.4	172.0
(4R)-**4** (*anti,exo*)	6.7	−0.85	−55.1	140.5

Table 1. Thermodynamic data and main-chain torsion angles for the *N*-acetyl-fluoroproline methyl ester derivatives.

a The values of the equilibrium constants $K_{trans/cis}$ were calculated from integration of the well-resolved H$^\alpha$ peaks in the ^{1}H NMR spectra of **1** and **2** in D$_2$O solution at 25 °C. *b* Energy values were obtained from DFT calculations with B3YLP at 6-311+G(2d,p).[6,7] *c* Dihedral angles were determined from the crystal structures of **1**, **2**, and **4** and from DFT calculations for **3**.[6,7]

non-bonded n→π* interaction between the p-type lone pair of the amide oxygen atom and the antibonding orbital of the ester carbonyl group.[6a] Raines et al. have postulated that the close contact between the O0 and C1 atoms (δ_{BD} = 2.76 Å) and the relatively large value of the O0···C1=O1 angle (98°) that were observed in the crystal structure of **4** resembled the Bürgi-Dunitz (BD) trajectory[20] for approach of a nucleophile to a carbonyl group.[1a,b,6a] Similarly, a close contact (2.81 Å) was observed between O0 and C1 in the crystal structure of **1**, albeit with a more acute value of the O0···C1=O1 angle (91°) than that observed in the crystal structure of **4**. These data imply the presence of an n→π* interaction in the C$^\gamma$-*exo* conformation of **1** and **4**, although the apparently stronger interaction in the latter case reduces the corresponding values of the ϕ and ψ dihedral angles to a greater extent than for **1**. Notably, the nonbonded contacts observed between O0 and C1 for the energetically preferred C$^\gamma$-*endo* conformations of **2** and **3** (δ_{BD} = 3.08 and 3.23 Å, respectively) were beyond the range in which a significant energetic stabilization could be reasonably expected.

The origin of the observed differences in ϕ and ψ between **1** and **4** could conceivably arise from the anticipated steric repulsion between the *syn*-oriented fluorine atom and the carbonyl group of the methyl ester. This repulsive interaction presumably distorts the C$^\gamma$-*exo* conformation of **1** from that observed in the crystal structure of **4**. A

similar repulsive interaction was invoked to rationalize the variance of the ϕ and ψ torsions that was observed in a computational analysis of the conformational energetics of **3**.[6a] In the latter case, a 1,3-diaxial steric interaction between the *syn*-oriented *fluoro* and carbomethoxy substituents significantly distorts values of the ϕ and ψ torsions of the C^γ-*endo* ring pucker in comparison to the corresponding conformations for the unsubstituted *N*-acetylproline methyl ester or the *anti*-substituted derivatives **2** and **4**. These results suggest that lone pair/lone pair repulsion between the *syn*-oriented fluorine atom and the carbonyl oxygen atom (O0) of the ester can significantly alter the peptide torsions of the *syn*-substituted fluoroproline derivatives **1** and **3**. Thus, two potentially conflicting structural considerations can modulate the conformational energetics of fluoroproline derivatives: the stereoelectronic *gauche* effect and the steric repulsion between *syn*-oriented substituents. The interplay between these two factors may determine the energetic effect of fluoroproline substitution into polypeptide sequences.

Biosynthesis of Elastin Derivatives.

The subtle differences in conformational properties between the C^γ-*exo* and C^γ-*endo* ring puckers suggested that fluoroproline substitution may be employed as a mechanism to interrogate local structural effects that arise from the presence of proline

residues in polypeptide sequences.[1,5,7, 21] We recently demonstrated that stereoelectronic effects that result from 4-fluoroproline substitution altered the thermodynamics of self-assembly of an elastin-mimetic polypeptide.[9] The differences in macromolecular properties between the elastin derivatives were interpreted on the basis of differential stabilization of the β-turn structures[22-27] that develop in the VPGVG structural repeats above the phase transition (Scheme 1). The greater stability of the (4R)-fluoroproline-substituted elastin (**elastin-3**) was attributed to the close correspondence between the observed values of the angles ϕ and ψ for the preferred C^γ-*exo* pucker of **3** and the range of ϕ and ψ angles associated with proline residues in the (i + 1) position of a type II β-turn conformation (ideal β_{II} ϕ, ψ: -60°, 120°).[28] In contrast, the deviation of the angles ϕ and ψ for the preferred C^γ-*endo* conformation of **2** from those of the type II β-turn conformation caused a destabilization of the corresponding elastin derivative (**elastin-2**). DFT calculations on model turn segments provided evidence for a relative destabilization of the type II β-turn structure in relation to the type I β-turn structure in the latter situation, which was supported by the conformational analysis of the corresponding elastin-mimetic polypeptides.[9]

To assess the effect of 3-fluoroproline substitution in a well-defined polypeptide model system, the elastin-mimetic proteins **elastin-4** and **elastin-5** were synthesized for

A.   **Elastin-1**: MGH$_{10}$S$_2$GHID$_4$KHM [(VPGVG)$_4$VPGIG]$_{16}$V

B.
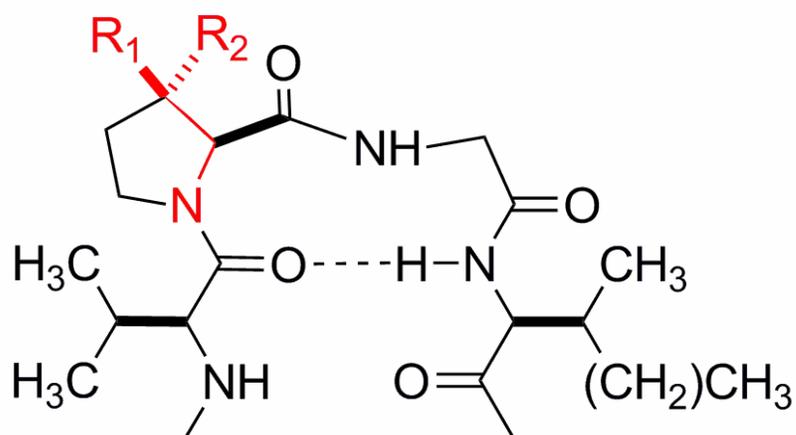
Scheme 1. (**A**) Amino acid sequence of the elastin-mimetic model polypeptide. (**B**) Structural representation of the β-turn unit of the elastin pentapeptide repeat. The bond vectors that geometrically define the stereoelectronic fluorine-amide *gauche* interactions within the substituted proline residues are highlighted in red for **elastin-4** (R^1 = F; R^2 = H), and **elastin-5** (R^1 = H; R^2 = F).

comparison to the 4-fluoroproline derivatives **elastin-2** and **elastin-3** (Scheme 1). On the basis of the observed regiochemical reversal of ring-pucker preferences for **1** and **2** in relation to **3** and **4**, we expected that **elastin-4** and **elastin-5** would display opposite tendencies in conformational and thermodynamic behavior compared to the corresponding 4-fluoroproline derivatives.

The elastin-mimetic polypeptides **elastin-4** and **elastin-5** were prepared through the biosynthetic approach described in chapter 2 for **elastin-2** and **elastin-3**.[11] The purified proteins were obtained in comparable yield to the 4-fluoroproline elastin derivatives (**elastin-4**: 53.5, and **elastin-5**: 48.7 mg/L culture). MALDI-TOF mass spectrometry of the **elastin-4** and **elastin-5** indicated virtually complete substitution of proline with the respective analogue (Figure 2). The ^{1}H-^{1}H NOESY and ^{19}F NMR spectra of **elastin-4** and **elastin-5** revealed that the *trans* configuration of the prolyl-peptide bond was the dominant one in aqueous solution (Figure 3 and 4, respectively). In addition, the values of the coupling constants, $^3J_{H\alpha H\beta}$ and $^3J_{H\alpha F\beta}$, for the substituted proline residues in the elastin-mimetic polypeptides were similar to the corresponding values for **1** and **2**, which indicated that the ring-pucker preferences of the analogues were conserved within the respective polypeptides.
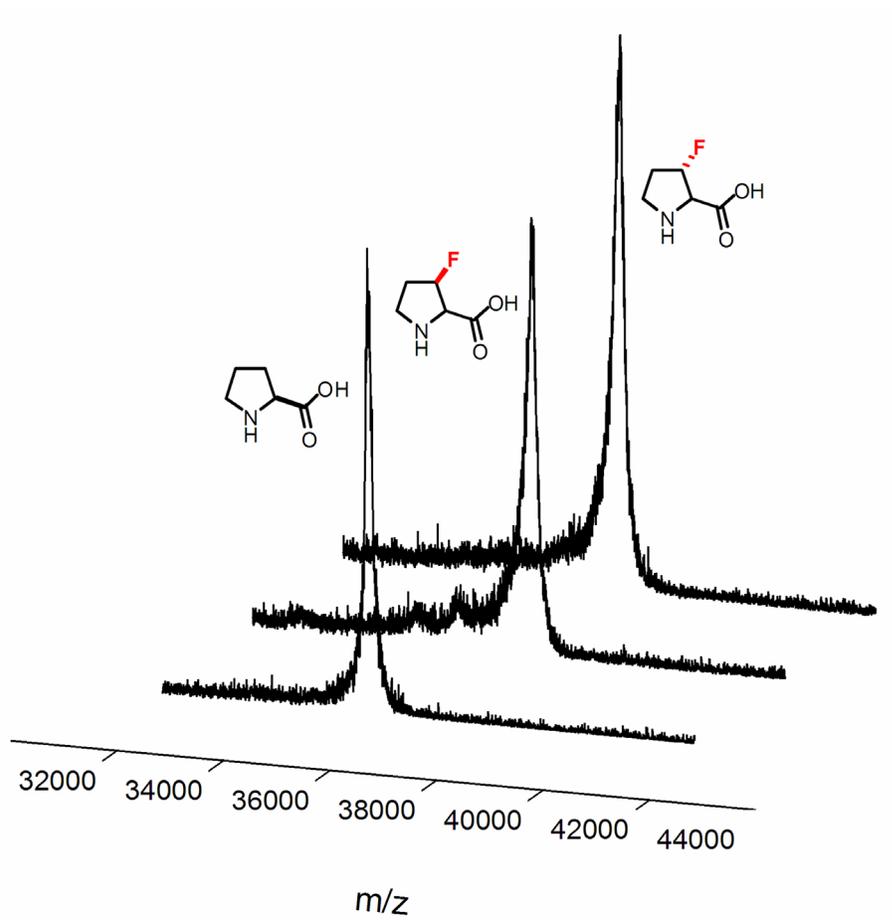
Figure 2. Mass determination for **elastin-4** and **elastin-5** using MALDI-TOF MS spectrometer. The observed mass differences between native elastin and elastin derivatives are proportional to the extent of proline substitution with 3-fluoroprolines.
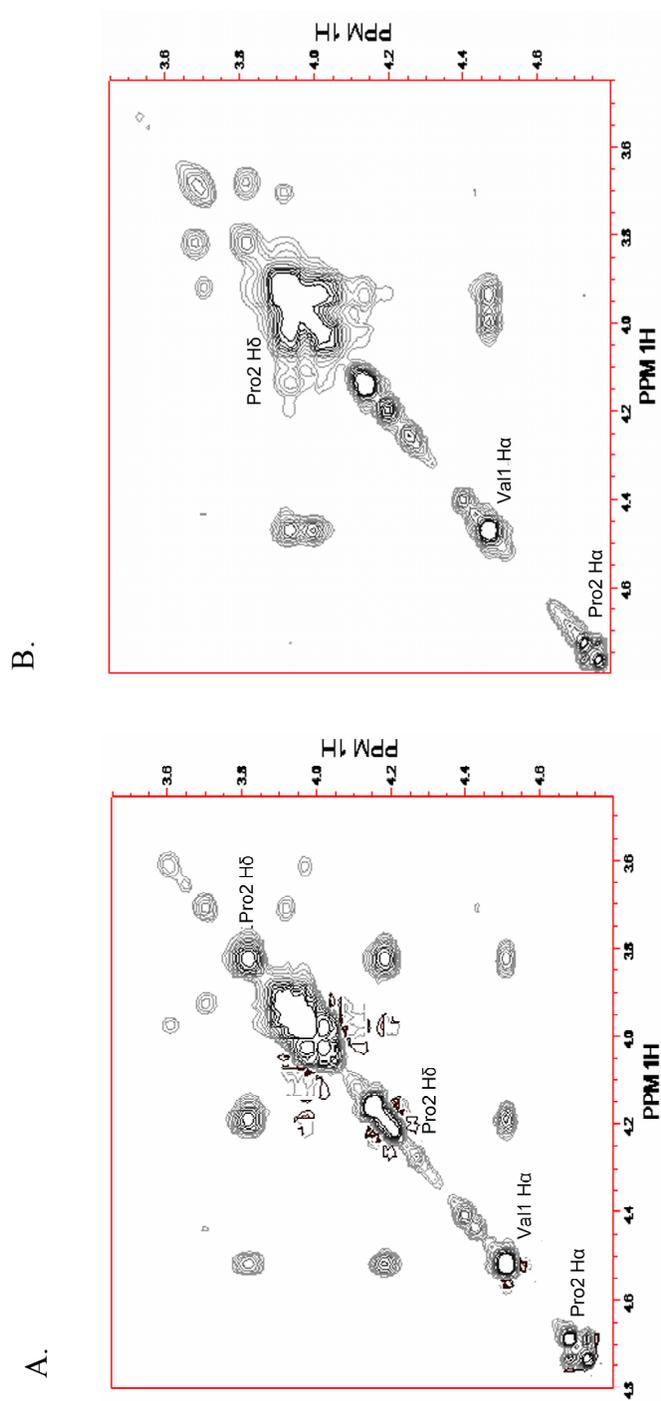
Figure 3. Expansion of the two-dimensional ^{1}H-^{1}H NOESY NMR spectra of **elastin-4**, (**A**), and **elastin-5**, (**B**). The spectral windows depict the Val(αH)-Pro(δH) cross-peaks associated with the *trans* configuration of the Val-Pro peptidyl bond. Note the absence of a strong Val(αH)-Pro(αH) cross-peak corresponding to a *cis* Val-Pro peptidyl bond configuration.
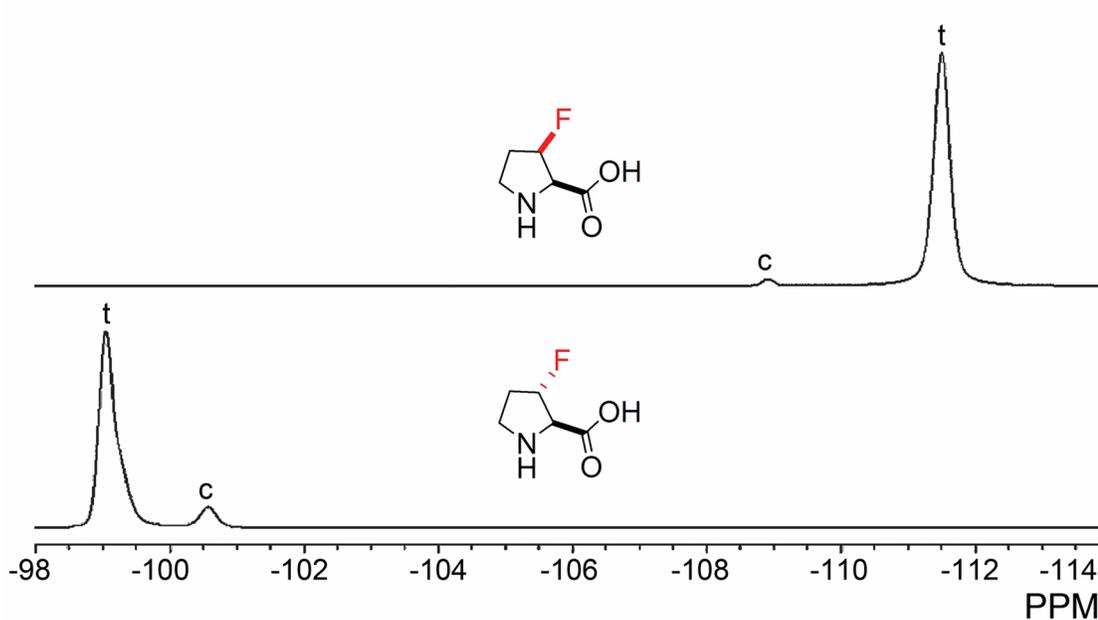
Figure 4. ¹⁹F NMR spectra of **elastin-4**, (top), and **elastin-5**, (bottom), indicating the incorporation of (2*R*,3*R*)-3-fluoroproline and (2*R*,3*S*)-3-fluoroproline, respectively, into the recombinant polypeptide. The symbols **t** and **c** refer to the *trans* and *cis* isomers, respectively, of the Val-Pro peptide bonds within the repeat units. These spectroscopic assignments were based on comparison with ¹⁹F NMR chemical shift values for the corresponding 3-fluoroproline model compounds, (**1**) and (**2**). The greater relative abundance of the *cis* prolyl-peptidyl bond isomer in **elastin-5** versus **elastin-4** reflects the increased thermodynamic preference for this configuration in (2*R*,3*S*)-3-fluoroproline derivatives.

Conformational Analysis of Elastin Polypeptides.

Circular dichroism (CD) studies of elastin-mimetic polypeptides have established that a conformational rearrangement occurs as the temperature increases through the transition point, which corresponds to a conversion of the local secondary structure of the pentapeptide repeats from a random coil conformation to a more ordered type II β-turn conformation. This conformational transition can be detected in the CD spectra of **elastin-4** as a function of temperature, in that the random coil signature (negative ellipticity near 196 nm) is gradually replaced with the type II β-turn signature (positive ellipticity near 208 nm) as the temperature is increased through the transition point (Figure 5).[22,25,29] In contrast, the CD temperature manifold for **elastin-5** displays a different spectroscopic profile in which a strong negative ellipticity (minimum at 223 nm) develops above the transition temperature. This spectroscopic feature is not consistent with the formation of a type II β-turn conformation, but more closely resembles the CD spectrum of a type I β-turn structure.[30] Thus, the CD spectroscopic data indicate a formal reversal of turn preferences for the 3-fluoroproline-substituted elastin sequences with respect to those previously observed for the 4-fluoroproline derivatives. The C$^\gamma$-*exo* pucker of (4*R*)-fluoroproline significantly stabilized the type II β-turn conformation for **elastin-3**, whereas the C$^\gamma$-e*ndo* pucker of (4*S*)-fluoroproline in **elastin-2** appreciably
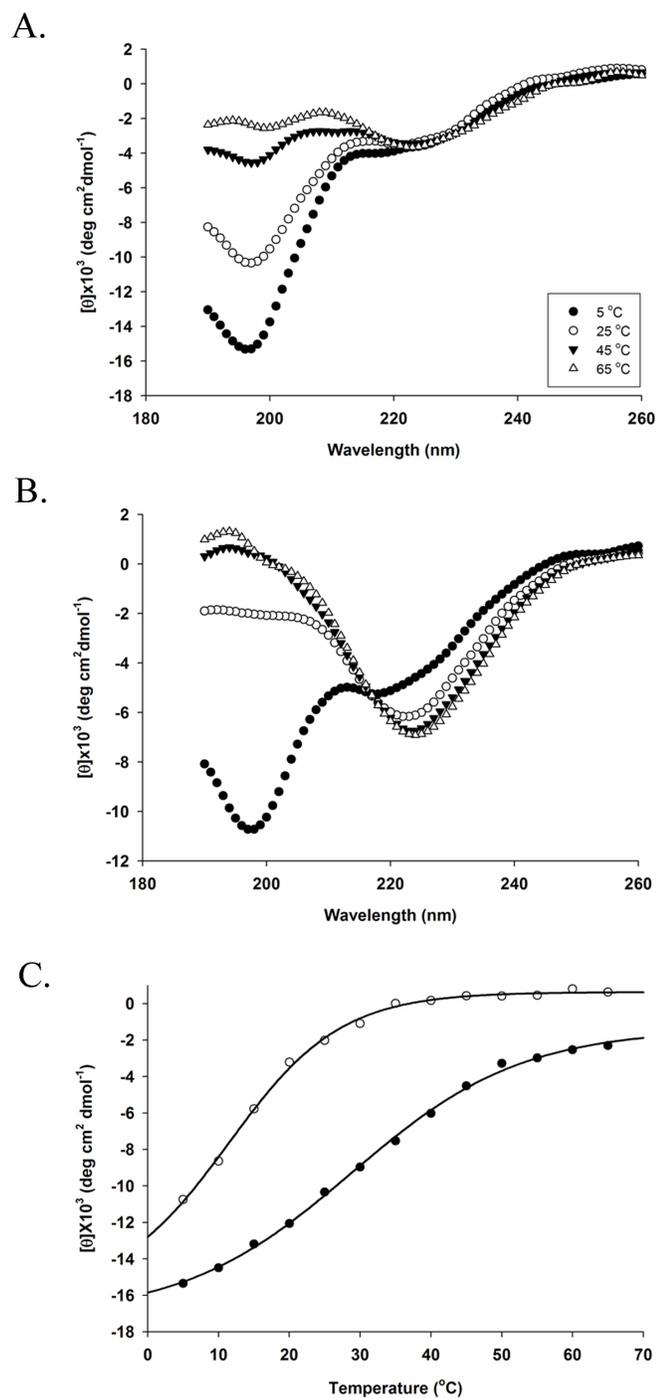
Figure 5. (**A**), (**B**) Temperature dependent CD spectroscopic manifolds for **elastin-4** and **elastin-5**, respectively. (**C**) Thermal transition curves for the disappearance of the random coil CD signal $[\theta]_{198}$ for **elastin-4** (filled circles) and **elastin-5** (open circles).

destabilized the type II β-turn conformation relative to alternative turn structures.[9] The CD spectroscopic data for both sets of fluoroproline-substituted elastin derivatives suggest that the C^{γ}-*exo* pucker is more compatible with a type II β-turn structure, and, conversely, the C^{γ}-*endo* pucker is less compatible with a type II β-turn structure to the point that the alternative type I β-turn conformation becomes energetically competitive, if not dominant, for the ordered state of the polypeptide.

Calorimetric Measurement of the Elastin Phase Transition.

In addition, the *syn* versus *anti* stereochemistry of the proline analogue may strongly influence the thermodynamic parameters associated with the elastin assembly. The CD thermal transition curves indicated that the disappearance of the low-temperature random coil conformation adheres to a quasi-two-state transition for both **elastin-4** and **elastin-5** (Figure 5C).[29] The mathematical fits of these thermal transitions provided estimates for the respective transition temperatures (T_t) of 30.1 and 10.9 °C. Differential scanning calorimetry (DSC) measurements provided T_t values of 35.9 and 13.6 °C for **elastin-4** and **elastin-5**, respectively, which correlate well with those determined from the CD temperature profiles (Figure 6). These T_t values[31] provide an indication of the facility with which the elastin assembly occurs, which, as a disorder-to-order

conformational transition, depends on the thermodynamic stabilization of the turn

structures in relation to the random coil conformation. The DSC data indicate that the

self-assembled state of **elastin-2** is stabilized with respect to that of **elastin-1** by

approximately 22 °C. Similarly, a stabilization of 20 °C was observed for the ordered

state of **elastin-3** versus that of **elastin-2**.[9] In both cases, incorporation of the *syn*-

fluoroproline derivative significantly raises the transition temperature for self-assembly

of the elastin-mimetic polypeptide with respect to that of the corresponding *anti*-

fluoroproline derivative. These data strongly suggest that the steric repulsion that

develops within the *syn*-oriented fluoroproline derivatives can adversely influence the

energetics of elastin assembly, and, more generally, of local turn-folding events

compared to the *anti*-fluoroproline derivatives. Moreover, manifestation of this effect is

independent of the intrinsic ring-pucker preference of the proline derivative and the

observed turn type of the elastin analogue.
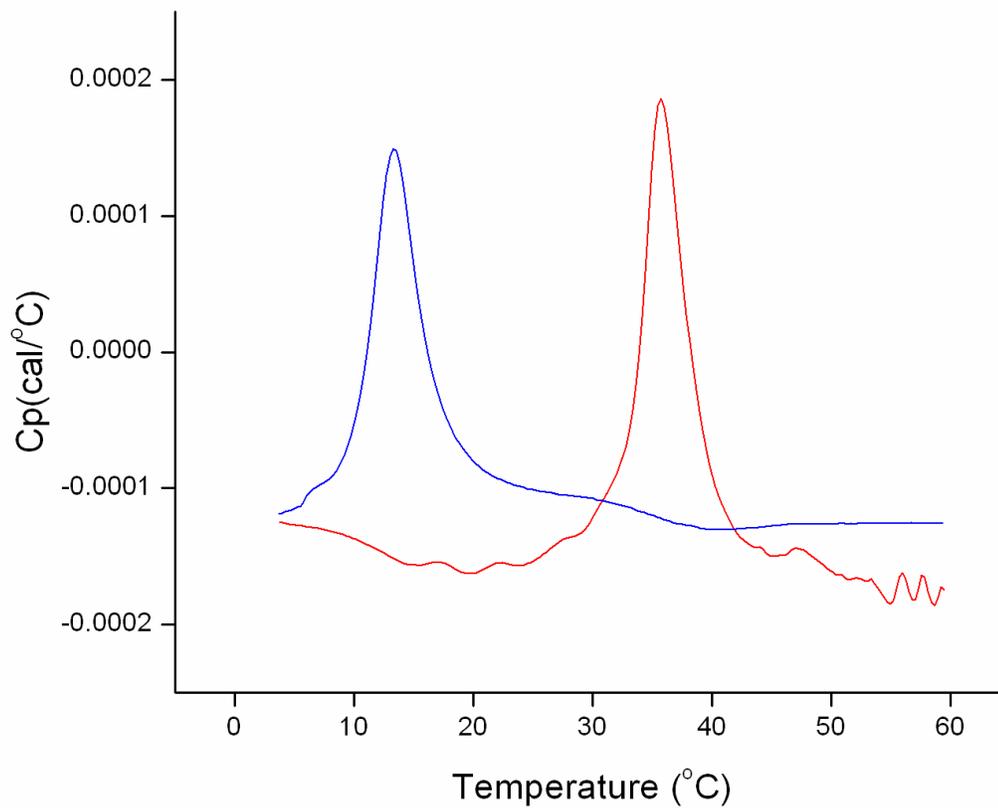
Figure 6. Raw, unadjusted differential scanning calorimetry data for dilute aqueous solutions (1.0 mg/mL) of **elastin-4** (red), and **elastin-5** (blue). The maxima for the endothermic transitions for **elastin-1** and **elastin-2** occur at temperatures of 35.9 ºC and 13.6 ºC, respectively.

Conclusion

In conclusion, we have demonstrated that the conformational energetics of a polypeptide can be altered through a combination of stereoelectronic and steric effects that arise from the presence of substituted proline residues at structurally critical positions. In the specific case under consideration, the assembly of a series of elastin-mimetic polypeptides could be interpreted in terms of the influence of the preferred fluoroproline ring pucker on the conformational energetics of the local β-turn structures. This "proline editing"[32] strategy-based on a set of four related fluoroproline derivatives[1.5,7,9]-may represent a general approach to interrogating the structural and functional role of specific proline residues within a polypeptide sequence, through comparison of the differential effect of fluoroproline substitution on the thermodynamics of folding and assembly among the resulting group of structural variants.

References

[1]   a) J.A. Hodges, R.T. Raines, *J. Am. Chem. Soc*. **2003**, *125*, 9262-9263; b) L. E. Bretscher, C.L. Jenkins, K.M. Taylor, M.L. DeRider, R.T. Raines, *J. Am. Chem. Soc*. **2001**, *123*, 777-778; c) S.K. Holmgren, L.E. Bretscher, K.M. Taylor, R.T. Raines, *Chem. Biol*. **1999**, *6*, 63-70; d) S.K. Holmgren, K.M. Taylor, L.E. Bretscher, R.T. Raines, *Nature*. **1998**, *392*, 666-667.

[2]   a) D. Barth, A.G. Milbradt, C. Renner, L. Moroder, *Chembiochem*. **2004**, *5*, 79-86; b) D. Barth, H.M. Musiol, M. Schutt, S. Fiori, A.G. Milbradt, C. Renner, L. Moroder, *Chem. Eur. J.* **2003**, *9*, 3692-3702.

[3]   A.V. Persikov, J.A. Ramshaw, A. Kirkpatrick, B. Brodsky, *J. Am. Chem. Soc*. **2003**, *125*, 11500-11501.

[4]   M. Doi, Y. Nishi, S. Uchiyama, Y. Nishiuchi, T. Nakazawa, T. Ohkubo, Y Kobayashi, *J. Am. Chem. Soc*. **2003**, *125*, 9922-9923.

[5]   a) M. Umashankara, I.R. Babu, K.N. Ganesh, *Chem. Commun*. **2003**, 2606-2607; b) I.R. Babu, K.N. Ganesh, *J. Am. Chem. Soc*. **2001**, *123*, 2079-2080.

[6]   a) M.L. DeRider, S.J. Wilkens, M.J. Waddell, L.E. Bretscher, L. F. Weinhold, R.T. Raines, J.L. Markley, *J. Am. Chem. Soc*. **2002**, *124*, 2497-2505; b) E.S.

Eberhardt, N. Panasik, Jr., R.T. Raines, *J. Am. Chem. Soc*. **1996**, *118*, 12261-12266; c) N. Panasik, Jr., E.S. Eberhardt, A.S. Edison, D.R. Powell, R.T. Raines, *Int. J. Pept. Protein Res*. **1994**, *44*, 262-269.

[7]   J.A. Hodges, R.T. Raines, *J. Am. Chem. Soc.* **2005**, *127*, 15923-15932.

[8]   C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber, L. Moroder, *Angew. Chem. Int. Ed*. **2001**, 40, 923-925.

[9]   W. Kim, R. A. McMillan, J. P. Snyder, V. P. Conticello, *J. Am. Chem. Soc.* **2005**, *127*, 18121-18132.

[10]   L. Demange, J. Cluzeau, A. Menez, C. Dugave, *Tetrahedron Lett.* **2001**, *42*, 651–653.

[11]   W. Kim, A. George, M. Evans, V. P. Conticello, *ChemBioChem* **2004**, *5*, 928-936.

[12]   N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann, R. Huber, *Eur. J. Biochem*. **1995**, *230*, 788-796.

[13]   APEX II, **2005**, Bruker AXS, Inc., Analytical X-ray Systems, 5465 East Cheryl Parkway, Madison WI 53711-5373.

[14]   SAINT Version 6.45A, **2003**, Bruker AXS, Inc., Analytical X-ray Systems, 5465 East Cheryl Parkway, Madison WI 53711-5373.

[15]   SHELXTL V6.12, **2002**, Bruker AXS, Inc., Analytical X-ray Systems, 5465 East
         Cheryl Parkway, Madison WI 53711-5373.

[16]   A. J. C. Wilson (ed), *International Tables for X-ray Crystallography, Volume C.*
         Kynoch, Academic Publishers, Dordrecht, **1992**, Tables 6.1.1.4 (pp. 500-502)
         and 4.2.6.8 (pp. 219-222).

[17]   C. L. Jenkins, L. E. Bretscher, I. A. Guzei, R. T. Raines, *J. Am. Chem. Soc.* **2003**,
         125, 6422-6427.

[18]   a) D. O'Hagan, C. Bilton, J.A.K. Howard, L. Knight, D.J. Tozer, *J. Chem. Soc.,*
         Perkin Trans. 2 **2000**, 605-607. b) C.R.S. Briggs, D. O'Hagan, J.A.K. Howard,
         D.S. Yufit, *J. Fluorine Chem.* **2003**, *119*, 9-13.

[19]   L. Vitagliano, R. Berisio, L. Mazzarella, A. Zagari, *Biopolymers* **2001**, *58*, 459-
         464.

[20]   a) H.B. Bürgi, J. Dunitz, E. Shefter, *Acta Crystallogr. Sect. B* **1974**, *30*, 1517-
         1527. b) H.B. Bürgi, J.M. Lehn, G. Wipff, *J. Am. Chem. Soc.* **1974**, *96*, 1965-
         1966. c) H.B. Bürgi, J.D. Dunitz, J.M. Lehn, G. Wipff, *Tetrahedron* **1974**, *30*,
         1563-1572. d) H.B. Bürgi, J.D. Dunitz, E. Shefter, *J. Am. Chem. Soc.* **1973**, *95*,
         5065-5067.

[21]   J.C. Horng, R.T.Raines, *Protein Sci.* **2006**, *15*, 74-83.

[22]   D.W. Urry, R.G. Shaw, K.U. Prasad, *Biochem. Biophys. Res. Commun*. **1985**, *130*, 50-57.

[23]   G.J. Thomas, Jr., B. Prescott, D.W. Urry, *Biopolymers* **1987**, *36*, 921-934.

[24]   D.W. Urry, N.R. Krishna, D.H. Huang, T.L. Trapane, K.U. Prasad, *Biopolymers* **1989**, *28*, 819-833.

[25]   H. Reiersen, A.R. Clarke, A.R. Rees, *J. Mol. Biol.* **1998**, *283*, 255-264.

[26]   B. Li, D.O. Alonso, V. Daggett, *J. Mol. Biol*. **2001**, *305*, 581-592.

[27]   X.L. Yao, M. Hong, *J. Am. Chem. Soc.* **2004**, *126*, 4199-4210.

[28]   C.M. Wilmot, J.M. Thornton, *J. Mol. Biol.* **1988**, *203*, 221-232.

[29]   T. Yamaoka, T. Tamura, Y. Seto, T. Tada, S. Kunugi, D.A. Tirrell, *Biomacromolecules* **2003**, *4*, 1680-1685.

[30]   a) A. Perczel, M. Hollosi, P. Sandor, G.D. Fasman, *Int. J. Pept. Protein Res.* **1993**, *41*, 223-236. b) A. Perczel, E. Kollat, M. Hollosi, G.D. Fasman, *Biopolymers* **1993**, *33*, 665-685.

[31]   a) D.W. Urry, C.-H. Luan, T.M. Parker, D.C. Gowda, K.U. Prasad, M.C. Reid, A. Safavy, *J. Am. Chem. Soc.* **1991**, *113*, 4346-4347. b) D.W. Urry, D.C. Gowda, T.M. Parker, C.-H. Luan, *Biopolymers* **1992**, *32*, 1243-1250.

[32]   K. M. Thomas, D. Naduthambi, G. Tririya, N. J. Zondlo, *Org. Lett.* **2005**, 7,

2397-2400

CHAPTER 5

Conclusions

The goal of this work was the investigation of the effect of proline substitution on the macromolecular properties of an elastin-mimetic polypeptide. Three specific aims were described in this process: (1) identification of a pool of proline analogues structurally and functionally distinct from the canonical residue; (2) development of methods for co-translational incorporation of those structurally diverse set of proline analogues into elastin-mimetic polypeptides in proline auxotrophic *Escherichia coli* strains; and (3) characterization of the resulting elastin derivatives using structural and thermodynamic analytical methods.

The co-translational incorporation method was appropriate for multi-site substitution of the relatively large number of proline residues (80 residues per polypeptide chain) with proline analogues into proteins. In this research, three different methods were developed for the substitution of proline residues. After protein purification, it was turned out that the cellular accumulation of the proteins depended on the degree of structural similarity between proline and the proline analogues. For example, the structural analogues with high similarity to proline were efficiently incorporated into the target protein in comparably high yield through use of the endogenous prolyl-tRNA synthetase (ProRS) in proline auxotrophic *E. coli* strains. In contrast, the structurally less similar proline analogues were incorporated into the target

proteins in relatively lower yield. The MALDI-TOF mass spectrometry, amino acid compositional analysis, and NMR spectroscopy were employed to examine the substitution level of proline analogues in respective to proline codons. The results show that virtually complete substitution was observed in the purified elastin analogues.

Incorporation of proline analogues into elastin-mimetic polypeptides had critical effects on the thermodynamic and conformational properties of the material that could be associated with the self-assembling properties. Of the incorporated proline analogues, fluoroprolines were especially interesting as a stereoelectronic structural probe. In order to evaluate the influence of fluoroprolines on pyrrolidine ring conformation, N-acetyl-fluoroproline methyl esters provided good models for local conformation of proline residue within the native protein sequences. The crystal structures and DFT calculation of the small molecules indicated that (3R)- and (4R)-fluoroprolines prefer C^{γ}-exo ring puckers, whereas (3S)- and (4S)-fluoroprolines prefer C^{γ}-$endo$ ring puckers. In addition, computational analysis of the "Pro-Gly" model fragment stressed the importance of steric influence of fluoro substituent. As a result, we suggested that two factors, the stereoelectronic and steric influence of fluoro substituent, contribute to the preference of pyrrolidine ring pucker and the preorganization of main-chain dihedral angles of fluoro-substituted proline residues.

Differential scanning calorimetry (DSC) was employed to examine the thermal transition of the fluoro-elastins. The (3R/4S)- and (3S/4R)-fluoroproline-substituted elastins displayed transition temperatures that were shifted toward higher (35.9/41.1 °C) and lower (13.6/21.6 °C) temperatures, respectively, in comparison with the T_t of the parent polypeptide (33.1 °C). Interestingly, stabilization of approximately 20 degrees was observed for both (3S)- and (4R)-fluoroproline-substituted elastins vis-à-vis the respective diastereoisomeric pair.

Circular dichroism (CD) studies of the polypeptides indicated that a conformational rearrangement occurred as the temperature approached the transition temperature, T_t. A comparative analysis of the CD spectra suggested that (3R/4R)-fluoroproline-substituted elastins stabilized a type II β-turn conformation with analogy with the parent elastin at T_t, whereas (3S/4S)-fluoroprolines destabilized the type II β-turn. The transition temperatures of the elastin derivatives determined from CD temperature profiles matched well with the corresponding ones from the DSC measurements.

The ^{1}H-^{1}H NOESY NMR spectra of the all elastins showed that the *trans* isomer of Val-Pro peptide bonds predominantly occurred in aqueous solution. In addition, the ^{19}F NMR spectra of fluoroproline-substituted elastins provided not only evidence for the predominant occurrence of the *trans* conformation, but also the existence of the minor *cis*

conformation.

We found that those observed differences in secondary structure and transition temperature of the fluoro-elastins were related to conformational features of fluoroprolines. On one hand, the *endo*/*exo* ring pucker of pyrrolidine ring seems to be correlated with the stability of specific turn types of elastin-mimetic polypeptides. The C^γ-*exo* pucker more likely stabilized a type II β-turn conformation, and, conversely, the C^γ-*endo* pucker destabilized the type II β-turn. On the other hand, the *syn*/*anti* orientation of the fluoro substituent seems to be correlated with the transition temperature associated with elastin self-assembly. The *syn* oriented fluoroprolines shifted the transition temperature to much higher temperature in comparison with the corresponding *anti* oriented fluoroprolines.

In summary, the incorporation methods developed in this thesis were effective for global substitution of proline residues, especially for the multi-site substitution in protein-based materials with a repeat sequence motif. Furthermore, we anticipate that usage of noncanonical amino acids provide a powerful tool to understand the structure-function relationships in protein polymers. The stereoelectronic effects of fluoroprolines on structural stabilization have been explored in collagen-mimetic peptides as well as elastin-mimetic polypeptide in this work. Recently, this stereoelectronic effect was also

applied to proline-containing peptide substrate as a stabilizing factor in protein folding and binding. The use of stereoelectronic structural control can emerge as a general protein engineering method for altering the local conformation of specific proline residues at structurally critical positions within polypeptide sequences.

The future studies are envisioned that involve cross-linking of these elastin polymers to create hydrogels, which would enable the investigation of the effect of proline substitution of the mechanical properties of elastin analogues. In addition, morphological studies should be considered for understanding the structure of the non-cross-linked and cross-linked protein polymers. The utility of elastin-mimetic polypeptides in biomedical engineering and materials science has just begun to be explored. The elastin analogues developed in this thesis could provide useful insights into the development of novel elastin matrices for biomaterial applications.

APPENDIX 1

CD and NMR Spectra of Elastin Analogues

Circular dichroism (CD) spectroscopy was also employed to observe secondary structure development of elastin derivatives during the thermal transition in dilute aqueous solution. The conformational transition of **elastin-7** and **elastin-8** (contain (4*S*)- and (4*R*)-hydroxyprolines, respectively) showed similar spectral behavior with those of **elastin-2** and **elastin-3**, respectively. The temperature-dependent CD spectra of **elastin-8** indicated the development of type II β-turn structure from random coil as the temperature increases. However, the CD spectra comparison of **elastin-7** with **elastin-2** suggested that **elastin-7** consisted of alternative turns in addition to type II β-turn. The thermal transition plots for both elastins monitoring the disappearance of the random coil structure were linear, and therefore the transition temperature was undetectable in the CD temperature profiles possibly due to increased hydrophilic property from a high content of hydroxyl group.

The 3,4-dehydroproline-containing elastin (**elastin-6**) generated different signatures in the CD spectra, which were observed as a positive ellipticity near 195 nm, a strong negative minimum at 204 nm, and a weak shoulder at near 220 nm. The negative intensity was decreased and the positive intensity was increased as the temperature increases. The gradual decrease of negative minimum at 204 nm probably indicates the appearance of ordered structure from disordered structure, and, however, the

conformation is not consistent with the type II β-turn structure.

Unexpectedly, the 4,4-difluoroproline-containing elastin (**elastin-9**) already adopted the type II β-turn structure even at lower temperature. The characteristic type II β-turn signature was shown at near 210 nm, which a positive maximum ellipticity was shifted to slightly shorter wavelength and then disappeared as the temperature is increased, while a negative minimum ellipticity was still maintained at approximately 225 nm even at high temperature. The spectroscopic feature at higher temperature was not consistent with the CD spectra of the type II β-turn structure, but more similar to that of **elastin-2**. The random coil conformation was not observed at near 195 nm in the measurable temperature range. The conformational rearrangement corresponds to the conversion between the ordered secondary structures. These CD spectroscopic data suggest that the substitution of proline residue with 4,4-difluoroproline significantly stabilize the type II β-turn conformation in the range of lower temperatures (5 or lower to 20 ^{o}C), and gradually destabilize the type II β-turn, but whereas it stabilizes the alternative turn structure in the range of higher temperatures.

The CD spectra of **elastin-10** (contains (2*S*)-azetidine-2-carboxylic acid) displayed different thermal behavior from those of other elastin analogues. A random coil or disordered structure signature was not observed at lower temperature. Interestingly, the

possible helical signatures were detected even at lower temperatures possessing a negative ellipticity at 208 and 220 nm and a positive ellipticity near 195 nm. Negative signature at 208 nm diminished at higher temperatures, while the longer wavelength feature dominated. A weak positive ellipticity was also observed near 208 nm as a shoulder that might indicate the presence of type II β-turn conformation. However, the conformation developed at higher temperatures is likely composed of such alternative turn structures as those observed for **elastin-2**, **elastin-5** and **elastin-7**.

The CD spectra of **elastin-11** (contains (2*S*)-piperidine-2-carboxylic acid) at lower temperature resembled the random coil signature. The conformational transition was occurred, and similar spectral behavior to **elastin-10** was observed at higher temperatures.

Due to a high insoluble property of **elastin-12** in water and a low concentration of urea even at lower temperatures, it was not possible to obtain the CD spectra.

A.
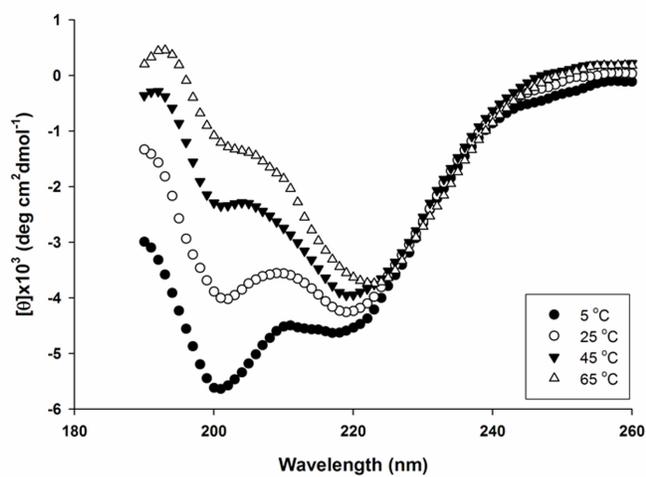
B.
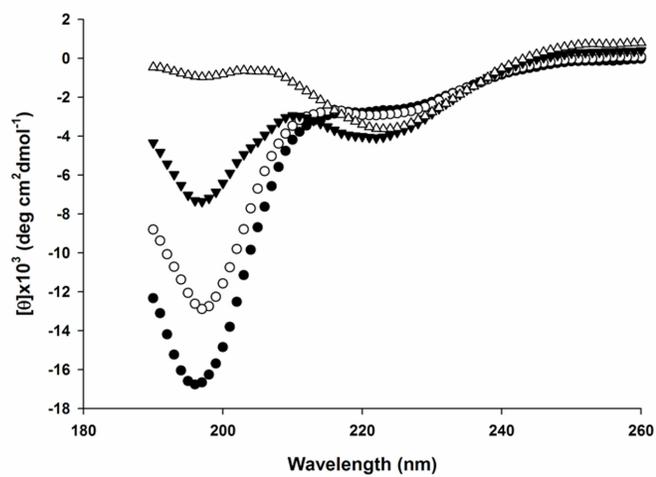
Figure A1-1. Temperature dependent CD spectroscopic manifolds for **elastin-7** (**A**) and **elastin-8** (**B**) at a concentration of 26.9 μM and 21.5 μM, respectively.
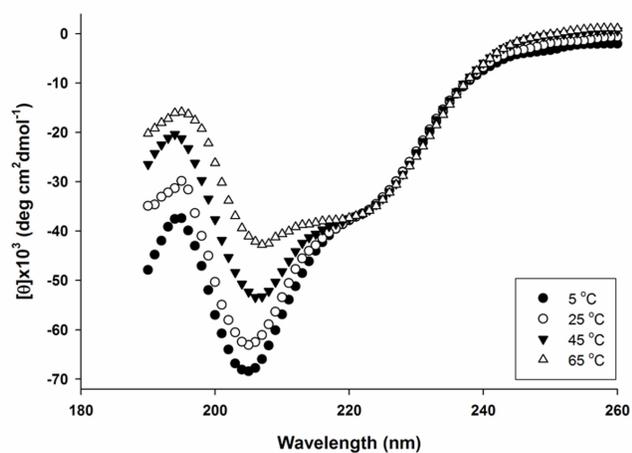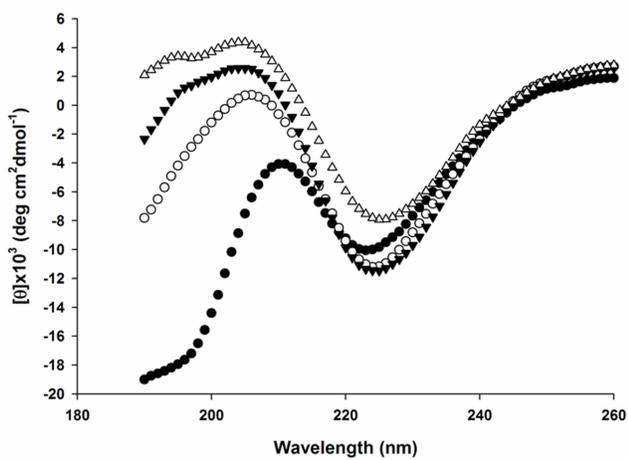
Figure A1-2. Temperature dependent CD spectroscopic manifolds for **elastin-6** (**C**) and **elastin-9** (**D**) at a concentration of 8.4 μM and 5.2 μM, respectively. Concentration was approximately estimated from the lyophilized protein polymers.

E.
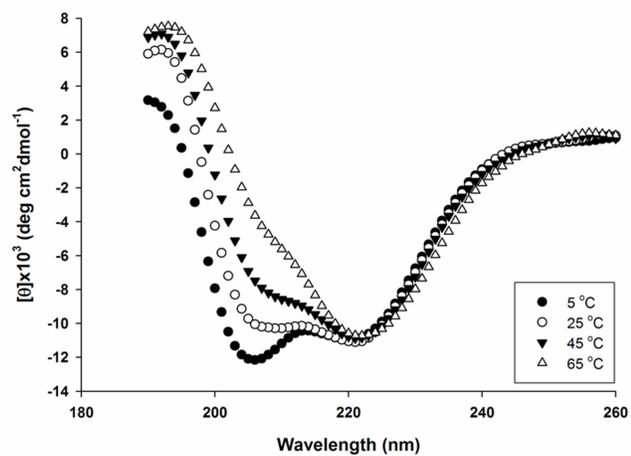
F.
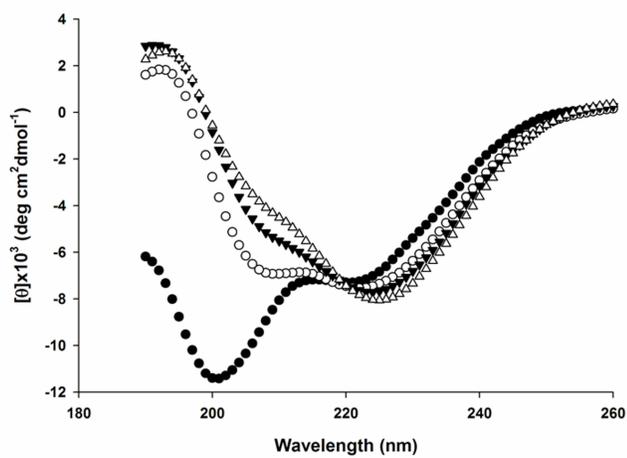
Figure A1-3. Temperature dependent CD spectroscopic manifolds for **elastin-10** (**E**) and **elastin-11** (**F**) at a concentration of 4.6 μM and 5.4 μM, respectively.
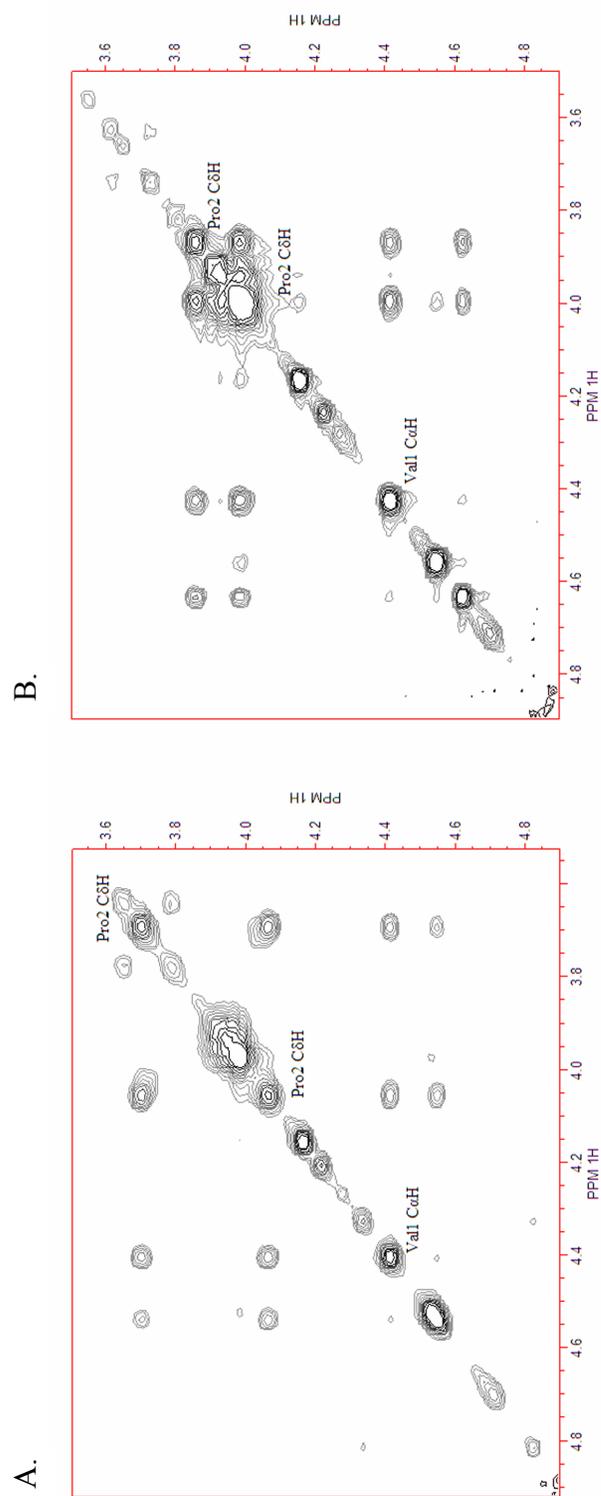
Figure A1-4. Expansion of the two-dimensional ^{1}H-^{1}H NOESY NMR spectra depicting the Val(αH)-Pro(δH) cross-peaks associated with the *trans* configuration of the Val-Pro peptidyl bond for **elastin-7** (**A**), and **elastin-8** (**B**). Note the absence of a strong Val(αH)-Pro(αH) cross-peak corresponding to a *cis* Val-Pro peptidyl bond configuration.
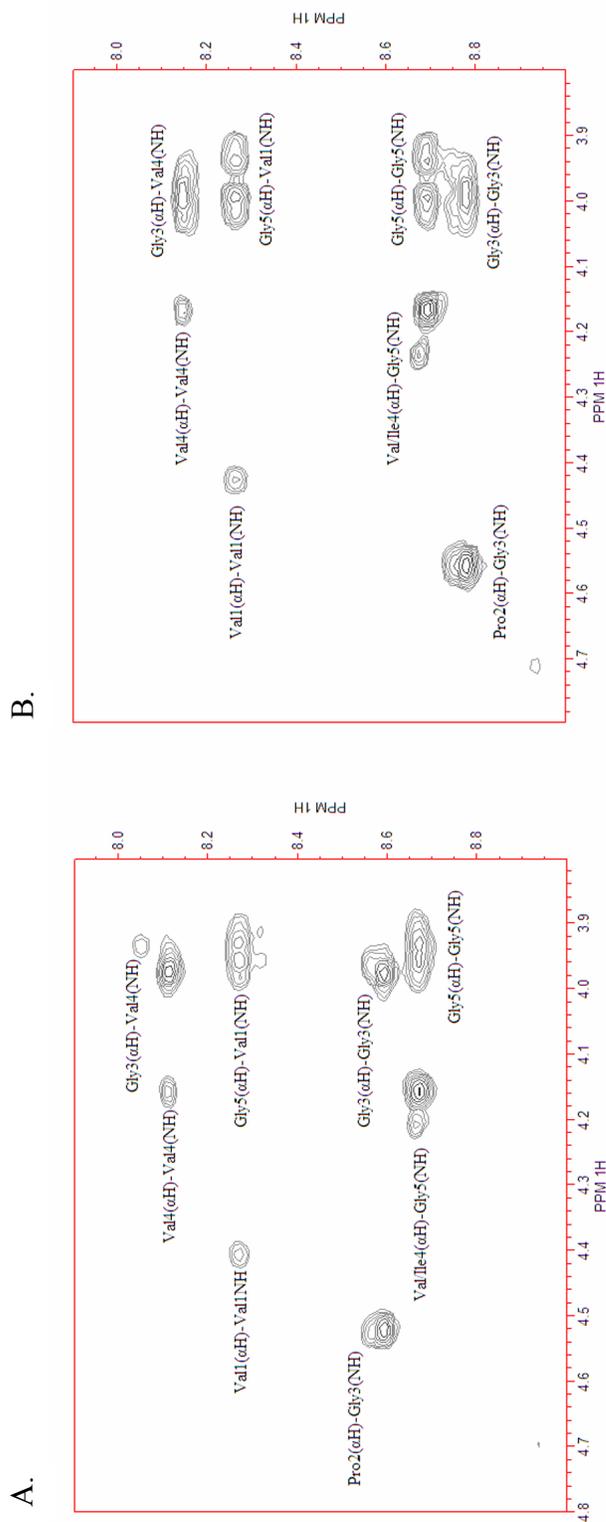
Figure A1-5. Expansion of the two-dimensional ^{1}H-^{1}H NOESY NMR spectra depicting the HN-Hα cross-peaks that were employed to confirm the spectroscopic assignments of the non-proline residues in the pentapeptide repeats of the elastomeric domains of **elastin-7** (**A**), and **elastin-8** (**B**). Similar trend of Gly^3 NH chemical shift was observed in comparison with 4-fluoroproline-elastin epimers

APPENDIX 2

Morphology Studies: Cryoetch-HRSEM and Cryo-TEM

images of elastin-1, elastin-2, and elastin-3

Cryoetch high-resolution scanning electron microscopy (cryoetch-HRSEM) and cryo-transmission electron microscopy (cryo-TEM) were employed to investigate the morphology of coacervate of the non-crosslinked elastin-mimetic derivatives above the transition temperature. In cryoetch-HRSEM images, a network of filaments was observed for **elastin-1** and **-2** at low primary magnification, while the population of protein polymer looked too dense for **elastin-3** to define the microstructure. The filaments of **elastin-1** and **-2** formed fibers which were three-dimensionally interconnected with each other. On the other hand, some filaments were nucleated to be spherical at a certain region of fiber. At high primary magnification, the cryoetch-HRSEM images of **elastin-1** showed a network of fibrils joined with spherical aggregates of the filaments (approximately 70 to 200 nm in diameter), while those of **elastin-2** showed more fibrous structures with smaller sized sphere (approximately 60 to 80 nm in diameter). The cryoetch-HRSEM images of **elastin-3** at high magnification revealed the compact/dense aggregation states of the protein polymer above the transition temperature.

In cryo-TEM images, coacervates of **elastin-2** and **elastin-3** above the transition temperature were comparable with those of tropoelastin, and the elastin-mimetic polypeptide (**Lys25**) containing lysine in place of isoleucine residue. The cryo-TEM images of **elastin-2** and **elastin-3** confirmed the presence of the polydisperse, spherical particles with the diameter ranging from 30 to 80 nm. The prior morphology study of elastin-mimetic polypeptide suggested that spherical particles were overlapped and joined to form beaded filament. Based on cryo-SEM and TEM images, these morphological features of **elastin-1, -2** and **-3** may also indicate that the filamentous structure results from interconnection of spherical particles on the basis of cryoetch-HRSEM images.
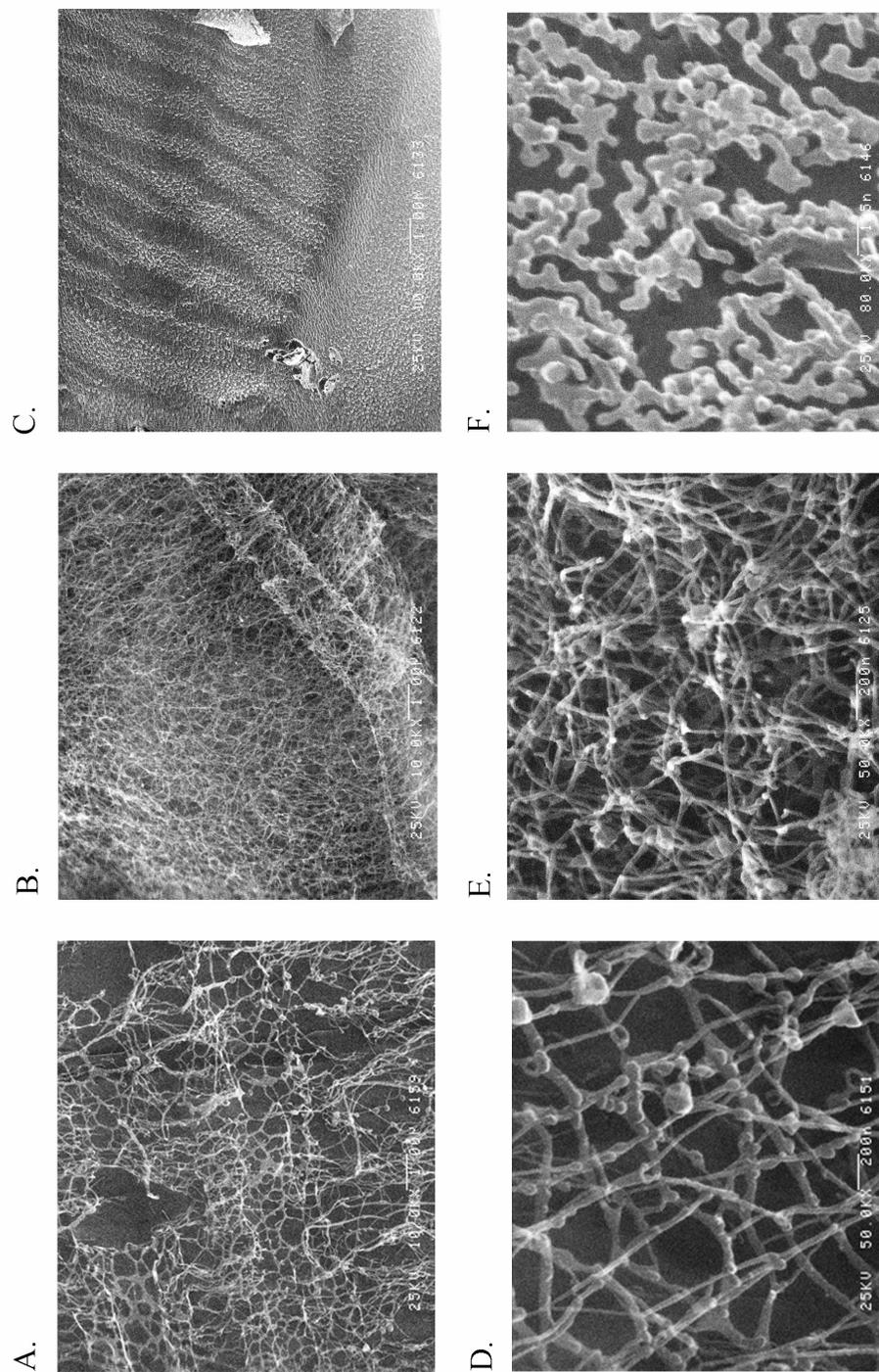
Figure A2-1. Cryo-etch HRSEM images of coacervates of **elastin-1** (**A** and **D**), **elastin-2** (**B** and **E**), and **elastin-3** (**C** and **F**). Solution of **elastin-1**, **elastin-2**, and **elastin-3** (2 mg/mL) in water was heated above the transition temperature, 38, 45, and 25 °C, for 10 minutes to produce the coacervates, respectively. The specimens were plunged into liquid ethane (-183°C) and etched prior to image acquisition.
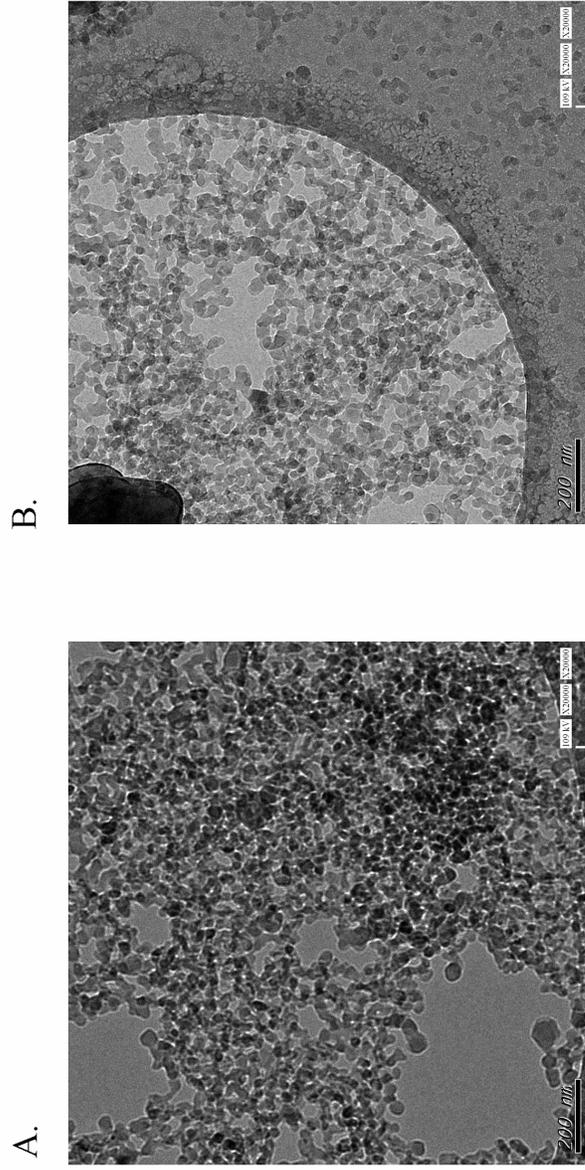
Figure A2-2. Cryo-TEM images of **elastin-2** (**A**) and **elastin-3** (**B**). Solution of **elastin-2** and **elastin-3** (1 mg/mL) in water was heated above the transition temperature. The solution was incubated for 10 minutes at 45 and 25 °C, respectively. The specimens were plunged into liquid ethane (-183°C) prior to image acquisition.

APPENDIX 3

Crystallographic data for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline

methyl ester and *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester

Table A3-1. Crystal data and structure refinement for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester.

Identification code	wkaccf_0m
Empirical formula	$C_8H_{12}FNO_3$
Formula weight	189.19
Temperature	173(2) K
Wavelength	1.54178 Å
Crystal system	Orthorhombic
Space group	P2(1)2(1)2(1)
Unit cell dimensions	a = 7.0907(2) Å $\alpha = 90^{\circ}$.
	b = 9.8318(4) Å $\beta = 90^{\circ}$.
	c = 12.9324(5) Å $\gamma = 90^{\circ}$.
Volume	901.57(6) Å^3
Z	4
Density (calculated)	1.394 Mg/m^3
Absorption coefficient	1.025 mm^{-1}
F(000)	400
Crystal size	0.36 x 0.11 x 0.09 mm^3
Theta range for data collection	5.65 to 59.21°.
Index ranges	-7<=h<=7, -9<=k<=11, -14<=l<=13
Reflections collected	4224
Independent reflections	1201 [R(int) = 0.0198]
Completeness to theta = 59.21°	95.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.9134 and 0.7092
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	1201 / 0 / 120
Goodness-of-fit on F^2	1.109
Final R indices [I>2sigma(I)]	R1 = 0.0216, wR2 = 0.0532
R indices (all data)	R1 = 0.0221, wR2 = 0.0535
Absolute structure parameter	-0.05(17)
Largest diff. peak and hole	0.114 and -0.174 e.Å^{-3}

Table A3-2. Atomic coordinates (x10^4) and equivalent isotropic displacement parameters (Å^{2}x10^3) for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
C(1)	-532(2)	2061(2)	6427(1)	30(1)
C(2)	843(3)	1800(2)	5558(1)	36(1)
C(3)	1959(2)	561(2)	5916(1)	34(1)
C(4)	594(2)	1719(1)	7407(1)	25(1)
C(5)	-643(2)	1105(2)	8250(1)	24(1)
C(6)	3236(2)	201(2)	7731(1)	28(1)
C(7)	4622(2)	-829(2)	7335(1)	36(1)
C(8)	-2540(2)	1644(2)	9698(1)	36(1)
N(1)	2024(2)	756(1)	7047(1)	27(1)
O(1)	-1040(1)	-77(1)	8332(1)	33(1)
O(2)	3168(2)	526(1)	8651(1)	37(1)
O(3)	-1282(1)	2087(1)	8874(1)	28(1)
F(1)	-2052(1)	1147(1)	6334(1)	41(1)

Table A3-3. Bond lengths [Å] and angles [°] for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester.

C(1)-F(1)	1.4091(17)	C(2)-C(1)-H(1)	111.5
C(1)-C(2)	1.510(2)	C(4)-C(1)-H(1)	111.5
C(1)-C(4)	1.535(2)	C(1)-C(2)-C(3)	104.16(12)
C(1)-H(1)	1.0000	C(1)-C(2)-H(2A)	110.9
C(2)-C(3)	1.525(2)	C(3)-C(2)-H(2A)	110.9
C(2)-H(2A)	0.9900	C(1)-C(2)-H(2B)	110.9
C(2)-H(2B)	0.9900	C(3)-C(2)-H(2B)	110.9
C(3)-N(1)	1.476(2)	H(2A)-C(2)-H(2B)	108.9
C(3)-H(3A)	0.9900	N(1)-C(3)-C(2)	102.32(13)
C(3)-H(3B)	0.9900	N(1)-C(3)-H(3A)	111.3
C(4)-N(1)	1.463(2)	C(2)-C(3)-H(3A)	111.3
C(4)-C(5)	1.524(2)	N(1)-C(3)-H(3B)	111.3
C(4)-H(4)	1.0000	C(2)-C(3)-H(3B)	111.3
C(5)-O(1)	1.2000(18)	H(3A)-C(3)-H(3B)	109.2
C(5)-O(3)	1.3376(17)	N(1)-C(4)-C(5)	111.76(11)
C(6)-O(2)	1.2327(18)	N(1)-C(4)-C(1)	103.90(11)
C(6)-N(1)	1.349(2)	C(5)-C(4)-C(1)	112.22(12)
C(6)-C(7)	1.501(2)	N(1)-C(4)-H(4)	109.6
C(7)-H(7A)	0.9800	C(5)-C(4)-H(4)	109.6
C(7)-H(7B)	0.9800	C(1)-C(4)-H(4)	109.6
C(7)-H(7C)	0.9800	O(1)-C(5)-O(3)	124.51(13)
C(7)-H(7D)	0.9831	O(1)-C(5)-C(4)	125.52(13)
C(7)-H(7E)	0.9769	O(3)-C(5)-C(4)	109.93(12)
C(7)-H(7F)	1.0284	O(2)-C(6)-N(1)	120.18(14)
C(8)-O(3)	1.4557(18)	O(2)-C(6)-C(7)	122.01(14)
C(8)-H(8A)	0.9800	N(1)-C(6)-C(7)	117.80(13)
C(8)-H(8B)	0.9800	C(6)-C(7)-H(7A)	109.3
C(8)-H(8C)	0.9800	C(6)-C(7)-H(7B)	109.6
		H(7A)-C(7)-H(7B)	109.5
F(1)-C(1)-C(2)	108.78(12)	C(6)-C(7)-H(7C)	109.6
F(1)-C(1)-C(4)	109.16(11)	H(7A)-C(7)-H(7C)	109.5
C(2)-C(1)-C(4)	103.95(12)	H(7B)-C(7)-H(7C)	109.5
F(1)-C(1)-H(1)	111.5	C(6)-C(7)-H(7D)	115.4

H(7A)-C(7)-H(7D)	52.4		H(7D)-C(7)-H(7F)	103.7
H(7B)-C(7)-H(7D)	58.2		H(7E)-C(7)-H(7F)	102.9
H(7C)-C(7)-H(7D)	134.9		O(3)-C(8)-H(8A)	109.5
C(6)-C(7)-H(7E)	115.3		O(3)-C(8)-H(8B)	109.5
H(7A)-C(7)-H(7E)	135.0		H(8A)-C(8)-H(8B)	109.5
H(7B)-C(7)-H(7E)	60.8		O(3)-C(8)-H(8C)	109.5
H(7C)-C(7)-H(7E)	49.9		H(8A)-C(8)-H(8C)	109.5
H(7D)-C(7)-H(7E)	109.6		H(8B)-C(8)-H(8C)	109.5
C(6)-C(7)-H(7F)	108.5		C(6)-N(1)-C(4)	119.67(12)
H(7A)-C(7)-H(7F)	55.7		C(6)-N(1)-C(3)	128.11(13)
H(7B)-C(7)-H(7F)	141.9		C(4)-N(1)-C(3)	112.18(12)
H(7C)-C(7)-H(7F)	57.2		C(5)-O(3)-C(8)	115.65(12)

Symmetry transformations used to generate equivalent atoms:

Table A3-4. Anisotropic displacement parameters ($\text{Å}^2 \text{x} 10^3$) for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester. The anisotropic displacement factor exponent takes the form: $-2\pi^2 [\, h^2 a^{*2} U^{11} + ... + 2\, h\, k\, a^* b^* U^{12}]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
C(1)	35(1)	23(1)	31(1)	1(1)	-4(1)	-4(1)
C(2)	47(1)	35(1)	25(1)	2(1)	2(1)	2(1)
C(3)	45(1)	36(1)	22(1)	-2(1)	4(1)	5(1)
C(4)	29(1)	22(1)	25(1)	-3(1)	2(1)	-2(1)
C(5)	24(1)	24(1)	25(1)	-2(1)	-3(1)	2(1)
C(6)	26(1)	30(1)	27(1)	1(1)	1(1)	-5(1)
C(7)	28(1)	43(1)	37(1)	-1(1)	1(1)	2(1)
C(8)	37(1)	41(1)	29(1)	-2(1)	9(1)	-4(1)
N(1)	28(1)	31(1)	22(1)	-1(1)	2(1)	2(1)
O(1)	38(1)	22(1)	40(1)	1(1)	7(1)	-2(1)
O(2)	37(1)	47(1)	27(1)	-3(1)	-4(1)	2(1)
O(3)	34(1)	26(1)	26(1)	-2(1)	6(1)	-1(1)
F(1)	40(1)	42(1)	40(1)	5(1)	-11(1)	-11(1)

Table A3-5. Hydrogen coordinates $(x10^4)$ and isotropic displacement parameters $(Å^2 x10^3)$ for N-acetyl-(2R,3R)-3-fluoroproline methyl ester.

	x	y	z	U(eq)
H(1)	-973	3027	6431	36
H(2A)	1685	2592	5454	43
H(2B)	169	1607	4904	43
H(3A)	1304	-295	5733	41
H(3B)	3241	552	5614	41
H(4)	1222	2559	7674	30
H(7A)	5173	-1321	7921	54
H(7B)	5625	-367	6950	54
H(7C)	3975	-1473	6879	54
H(7D)	5909	-747	7611	54
H(7E)	4672	-924	6584	54
H(7F)	4197	-1778	7575	54
H(8A)	-1973	873	10063	54
H(8B)	-3752	1367	9401	54
H(8C)	-2741	2395	10184	54

Table A3-6. Torsion angles [°] for *N*-acetyl-(2*R*,3*R*)-3-fluoroproline methyl ester.

F(1)-C(1)-C(2)-C(3)	-79.57(14)
C(4)-C(1)-C(2)-C(3)	36.64(15)
C(1)-C(2)-C(3)-N(1)	-33.31(16)
F(1)-C(1)-C(4)-N(1)	90.72(13)
C(2)-C(1)-C(4)-N(1)	-25.22(14)
F(1)-C(1)-C(4)-C(5)	-30.18(16)
C(2)-C(1)-C(4)-C(5)	-146.12(12)
N(1)-C(4)-C(5)-O(1)	-30.5(2)
C(1)-C(4)-C(5)-O(1)	85.72(18)
N(1)-C(4)-C(5)-O(3)	151.48(12)
C(1)-C(4)-C(5)-O(3)	-92.26(14)
O(2)-C(6)-N(1)-C(4)	-1.7(2)
C(7)-C(6)-N(1)-C(4)	177.26(13)
O(2)-C(6)-N(1)-C(3)	176.05(15)
C(7)-C(6)-N(1)-C(3)	-5.0(2)
C(5)-C(4)-N(1)-C(6)	-56.43(17)
C(1)-C(4)-N(1)-C(6)	-177.64(13)
C(5)-C(4)-N(1)-C(3)	125.50(14)
C(1)-C(4)-N(1)-C(3)	4.29(16)
C(2)-C(3)-N(1)-C(6)	-159.85(14)
C(2)-C(3)-N(1)-C(4)	18.02(17)
O(1)-C(5)-O(3)-C(8)	0.0(2)
C(4)-C(5)-O(3)-C(8)	178.03(12)

Symmetry transformations used to generate equivalent atoms:

Table A3-7. Crystal data and structure refinement for *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester.

Identification code	wkactf_0m
Empirical formula	$C_8H_{12}FNO_3$
Formula weight	189.19
Temperature	173(2) K
Wavelength	1.54178 Å
Crystal system	Orthorhombic
Space group	P2(1)2(1)2(1)
Unit cell dimensions	a = 7.555(2) Å α= 90°.
	b = 9.550(2) Å β= 90°.
	c = 13.080(3) Å γ= 90°.
Volume	943.8(4) Å^3
Z	4
Density (calculated)	1.331 Mg/m^3
Absorption coefficient	0.979 mm^{-1}
F(000)	400
Crystal size	0.46 x 0.32 x 0.09 mm^3
Theta range for data collection	5.74 to 54.10°.
Index ranges	-7<=h<=6, -8<=k<=10, -13<=l<=13
Reflections collected	3638
Independent reflections	1128 [R(int) = 0.2104]
Completeness to theta = 54.10^o	99.3 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.9171 and 0.6616
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	1128 / 0 / 120
Goodness-of-fit on F^2	1.022
Final R indices [I>2sigma(I)]	R1 = 0.0830, wR2 = 0.2187
R indices (all data)	R1 = 0.1008, wR2 = 0.2610
Absolute structure parameter	1.8(10)
Largest diff. peak and hole	0.333 and -0.351 e.Å^{-3}

Table A3-8. Atomic coordinates (x10^4) and equivalent isotropic displacement parameters (Å^{2}x10^3) for *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
C(1)	10085(11)	4220(10)	1249(6)	59(2)
C(2)	8566(12)	5129(11)	940(7)	67(2)
C(3)	6943(10)	4221(10)	1051(6)	59(2)
C(4)	9322(10)	3298(10)	2097(6)	55(2)
C(5)	9610(10)	3983(8)	3146(6)	49(2)
C(6)	6407(11)	2320(10)	2335(7)	57(2)
F(1)	10530(6)	3338(6)	439(3)	71(2)
N(1)	7460(8)	3216(8)	1840(5)	51(2)
O(1)	8633(9)	4732(9)	3561(5)	89(3)
O(2)	6992(8)	1514(7)	2990(5)	77(2)
C(7)	4499(11)	2282(13)	1983(8)	81(3)
O(3)	11211(7)	3584(7)	3488(4)	65(2)
C(8)	11801(13)	4219(14)	4437(7)	92(4)

Table A3-9. Bond lengths [Å] and angles [°] for *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester.

C(1)-F(1)	1.395(9)	C(1)-C(2)-H(2A)	110.7
C(1)-C(2)	1.495(12)	C(3)-C(2)-H(2A)	110.7
C(1)-C(4)	1.530(12)	C(1)-C(2)-H(2B)	110.7
C(1)-H(1)	1.0000	C(3)-C(2)-H(2B)	110.7
C(2)-C(3)	1.509(12)	H(2A)-C(2)-H(2B)	108.8
C(2)-H(2A)	0.9900	N(1)-C(3)-C(2)	103.2(6)
C(2)-H(2B)	0.9900	N(1)-C(3)-H(3A)	111.1
C(3)-N(1)	1.462(11)	C(2)-C(3)-H(3A)	111.1
C(3)-H(3A)	0.9900	N(1)-C(3)-H(3B)	111.1
C(3)-H(3B)	0.9900	C(2)-C(3)-H(3B)	111.1
C(4)-N(1)	1.448(9)	H(3A)-C(3)-H(3B)	109.1
C(4)-C(5)	1.535(12)	N(1)-C(4)-C(1)	103.2(7)
C(4)-H(4)	1.0000	N(1)-C(4)-C(5)	111.6(7)
C(5)-O(1)	1.162(9)	C(1)-C(4)-C(5)	110.4(7)
C(5)-O(3)	1.345(9)	N(1)-C(4)-H(4)	110.5
C(6)-O(2)	1.233(10)	C(1)-C(4)-H(4)	110.5
C(6)-N(1)	1.336(11)	C(5)-C(4)-H(4)	110.5
C(6)-C(7)	1.513(11)	O(1)-C(5)-O(3)	126.2(8)
C(7)-H(7A)	0.9800	O(1)-C(5)-C(4)	126.1(7)
C(7)-H(7B)	0.9800	O(3)-C(5)-C(4)	107.7(7)
C(7)-H(7C)	0.9800	O(2)-C(6)-N(1)	121.5(8)
O(3)-C(8)	1.452(11)	O(2)-C(6)-C(7)	122.6(9)
C(8)-H(8A)	0.9800	N(1)-C(6)-C(7)	115.8(8)
C(8)-H(8B)	0.9800	C(6)-N(1)-C(4)	120.0(7)
C(8)-H(8C)	0.9800	C(6)-N(1)-C(3)	127.1(7)
		C(4)-N(1)-C(3)	112.8(7)
F(1)-C(1)-C(2)	109.3(7)	C(6)-C(7)-H(7A)	109.5
F(1)-C(1)-C(4)	107.1(7)	C(6)-C(7)-H(7B)	109.5
C(2)-C(1)-C(4)	103.9(7)	H(7A)-C(7)-H(7B)	109.5
F(1)-C(1)-H(1)	112.0	C(6)-C(7)-H(7C)	109.5
C(2)-C(1)-H(1)	112.0	H(7A)-C(7)-H(7C)	109.5
C(4)-C(1)-H(1)	112.0	H(7B)-C(7)-H(7C)	109.5
C(1)-C(2)-C(3)	105.3(7)	C(5)-O(3)-C(8)	116.3(7)

O(3)-C(8)-H(8A)	109.5	O(3)-C(8)-H(8C)	109.5
O(3)-C(8)-H(8B)	109.5	H(8A)-C(8)-H(8C)	109.5
H(8A)-C(8)-H(8B)	109.5	H(8B)-C(8)-H(8C)	109.5

Symmetry transformations used to generate equivalent atoms:

Table A3-10. Anisotropic displacement parameters ($\text{Å}^2 \text{x} 10^3$) for *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester. The anisotropic displacement factor exponent takes the form: $-2\pi^2 [\, h^2 a^{*2} U^{11} + ... + 2\, h\, k\, a^* b^* U^{12}\,]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
C(1)	62(5)	73(6)	43(4)	-14(4)	-5(4)	-10(5)
C(2)	60(5)	80(6)	62(5)	2(5)	-13(5)	-2(5)
C(3)	48(5)	79(6)	50(4)	3(4)	-12(4)	14(4)
C(4)	43(4)	78(6)	43(4)	-6(4)	-2(4)	7(4)
C(5)	43(4)	49(5)	56(4)	8(4)	3(4)	24(4)
C(6)	40(5)	71(6)	59(5)	-4(5)	2(4)	4(5)
F(1)	72(3)	90(4)	51(3)	-11(3)	7(2)	9(3)
N(1)	34(3)	68(4)	52(4)	-11(4)	-2(3)	-4(3)
O(1)	68(4)	135(6)	63(4)	-25(4)	-13(3)	52(5)
O(2)	59(4)	95(5)	78(4)	34(4)	4(3)	1(4)
C(7)	42(5)	128(9)	73(6)	13(6)	7(5)	-16(6)
O(3)	53(3)	86(4)	56(3)	-12(3)	-3(3)	20(3)
C(8)	73(6)	140(11)	63(6)	-33(6)	-32(5)	21(7)

Table A3-11. Hydrogen coordinates $(x10^4)$ and isotropic displacement parameters $(\text{Å}^2 x10^3)$ for N-acetyl-(2R,3S)-3-fluoroproline methyl ester.

	x	y	z	U(eq)
H(1)	11122	4782	1491	71
H(2A)	8704	5448	225	81
H(2B)	8485	5961	1390	81
H(3A)	5908	4780	1271	70
H(3B)	6658	3741	401	70
H(4)	9872	2346	2079	65
H(7A)	3953	1401	2201	121
H(7B)	4457	2351	1236	121
H(7C)	3853	3071	2284	121
H(8A)	11586	5231	4413	138
H(8B)	13069	4045	4528	138
H(8C)	11145	3811	5011	138

Table A3-12. Torsion angles [°] for *N*-acetyl-(2*R*,3*S*)-3-fluoroproline methyl ester.

F(1)-C(1)-C(2)-C(3)	79.5(9)
C(4)-C(1)-C(2)-C(3)	-34.5(9)
C(1)-C(2)-C(3)-N(1)	26.9(9)
F(1)-C(1)-C(4)-N(1)	-87.4(8)
C(2)-C(1)-C(4)-N(1)	28.2(8)
F(1)-C(1)-C(4)-C(5)	153.2(6)
C(2)-C(1)-C(4)-C(5)	-91.2(8)
N(1)-C(4)-C(5)-O(1)	-22.8(13)
C(1)-C(4)-C(5)-O(1)	91.4(11)
N(1)-C(4)-C(5)-O(3)	158.4(8)
C(1)-C(4)-C(5)-O(3)	-87.4(8)
O(2)-C(6)-N(1)-C(4)	-2.6(12)
C(7)-C(6)-N(1)-C(4)	-178.0(8)
O(2)-C(6)-N(1)-C(3)	179.9(8)
C(7)-C(6)-N(1)-C(3)	4.6(12)
C(1)-C(4)-N(1)-C(6)	170.3(7)
C(5)-C(4)-N(1)-C(6)	-71.2(10)
C(1)-C(4)-N(1)-C(3)	-11.9(9)
C(5)-C(4)-N(1)-C(3)	106.6(7)
C(2)-C(3)-N(1)-C(6)	168.6(8)
C(2)-C(3)-N(1)-C(4)	-9.0(9)
O(1)-C(5)-O(3)-C(8)	-3.8(13)
C(4)-C(5)-O(3)-C(8)	175.0(8)

Symmetry transformations used to generate equivalent atoms: