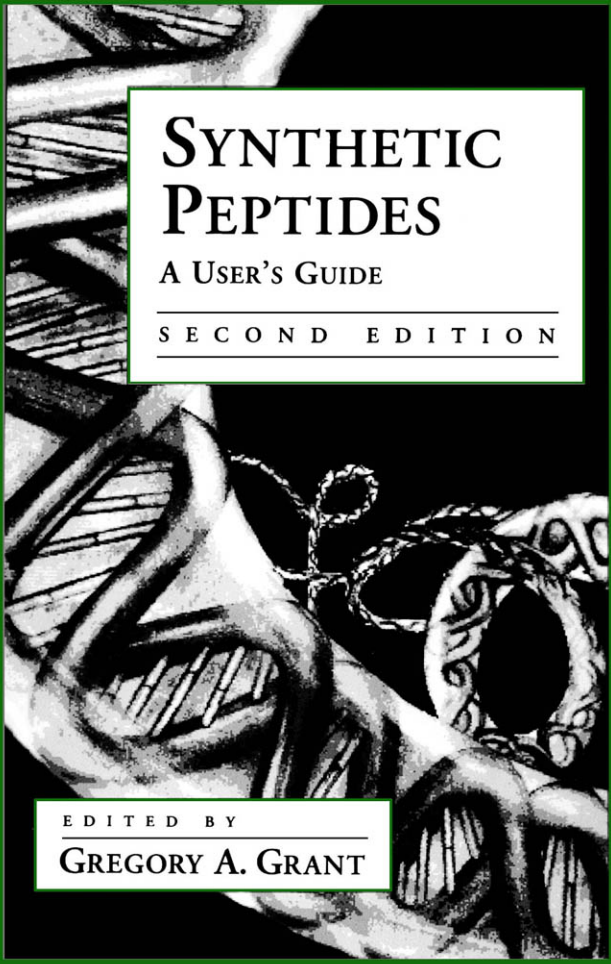

SYNTHETIC
PEPTIDES
A USER'S GUIDE
SECOND EDITION
EDITED BY
GREGORY A. GRANT

SYNTHETIC PEPTIDES

Advances in Molecular Biology

The Yeast Two-Hybrid System, edited by Paul L. Bartel and Stanley Fields

A Comparative Methods Approach to the Study of Oocytes and Embryos, edited by Joel D. Richter

Synthetic Peptides: A User's Guide, Second Edition, edited by Gregory A. Grant

SYNTHETIC PEPTIDES

A User's Guide

Second Edition

Edited by
Gregory A. Grant

OXFORD
UNIVERSITY PRESS
2002

OXFORD
UNIVERSITY PRESS

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

and an associated company in Berlin

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press

Library of Congress Cataloging-in-Publication Data
Synthetic peptides: a user's guide / edited by Gregory A. Grant.—2nd ed.
p. cm. — (Advances in molecular biology)
Includes bibliographical references and index.
ISBN 0-19-513261-0
1. Peptides—Synthesis. 2. Protein engineering. I. Grant, Gregory A., 1949– II. Series.
QP552.P4 S963 2002
572'.6545– dc21 2001036643

9 8 7 6 5 4 3 2 1
Printed in the United States of America
on acid free paper

For Marie

This page intentionally left blank

Contents

This page intentionally left blank

Contributors

George Barany
Departments of Chemistry and Laboratory Medicine and Pathology
University of Minnesota
Minneapolis, Minnesota 55455

Gregg B. Fields
Department of Chemistry and Biochemistry and the Center for Molecular Biology and Biotechnology
Florida Atlantic University
Boca Raton, Florida 33431

Gregory A. Grant
Department of Molecular Biology and Pharmacology
Washington University School of Medicine
St. Louis, Missouri 63110

Victor J. Hruby
Department of Chemistry
University of Arizona
Tucson, Arizona 85721

Janelle L. Lauer-Fields
Department of Chemistry and Biochemistry and the Center for Molecular Biology and Biotechnology
Florida Atlantic University
Boca Raton, Florida 33431

Rong-qiang Liu
Departments of Chemistry and Laboratory Medicine and Pathology
University of Minnesota
Minneapolis, Minnesota 55455

Terry O. Matsunaga
Department of Pharmacology and Toxicology
School of Pharmacy
University of Arizona
Tucson, Arizona 85721

Michael L. Moore
Smith-Kline Beecham Pharmaceuticals
King of Prussia, Pennsylvania 19406

SYNTHETIC PEPTIDES

This page intentionally left blank

1

Synthetic Peptides

Beginning the Twenty-first Century

Gregory A. Grant

This second edition of *Synthetic Peptides* is being published at the beginning of the twenty-first century and marks nearly 100 years since the beginnings of the chemical synthesis of peptides. From the first decade of the twentieth century up to the present time, the evolution of the development, analysis, and use of synthetic peptides has been steady and remarkable.

It has been about 10 years since the first edition of this book was published. Much remains unchanged, such as the basic principles of peptide structure, the basic chemistry for assembling a peptide chain, and many of the techniques used to evaluate synthetic peptides. However, during that time we have seen a switch from primarily the use of Boc chemistry for routine synthesis to that of Fmoc chemistry. Mass spectrometry has also matured with the development of more user-friendly and affordable instrumentation to the point that it is now the premier analytic method for synthetic peptides. Methods for the production of very long peptides, such as chemoselective ligation, are maturing, although they are still not an everyday thing for most peptide chemists, and better chemistries for producing peptides with "post-translational modifications," such as phosphates, sugars, and specific disulfide bonds, are now well within reach. As a result, you will find many sections of this book largely unchanged, but you will also find

many new sections that document the developments of the last ten years with the inclusion of new information and methodologies. However, a good feeling for what the beginning of the twenty-first century offers can best be appreciated by considering the developments that have led to this point.

Emil Fischer introduced the concept of peptides and polypeptides and presented protocols for their synthesis in the early 1900s (Fischer, 1902, 1903, 1906). Although others also made contributions in those days, most notably Theodor Curtius, the work of Fischer and his colleagues stands out, and he is generally regarded as the father of peptide chemistry. An excellent account of the history of peptide synthesis which treats the subject much more comprehensively than can be attempted here can be found in a book by Wieland and Bodanszky (1991). Fischer's place in the history of synthetic peptides is eloquently summed up in one sentence from that book which simply states, "To Emil Fischer we owe the systematic attack on a field of natural substances that had previously been avoided by chemists." In those early days, the chemistries developed by Fischer and others led to the production of molecules containing as many as 18 amino acids, such as Leucyl (triglycyl) leucyl (triglycyl) leucyl (octaglycyl) glycine (Fischer, 1907). Nonetheless, the syntheses performed were difficult and limited to simple amino acids, and progress was slow for many years. Then, the discovery of an easily removable protecting group, the carbobenzoxy group, by Bergmann and Zervas in 1932 (Bergmann and Zervas, 1932) provided new impetus by opening the way to the use of polyfunctional amino acids. As a result, the synthesis of naturally occurring small peptides such as carnosine (Sifferd and du Vigneaud, 1935) and glutathione (Harington and Mead, 1935) were soon achieved. Almost 20 years later, the synthesis of an active peptide hormone, the octapeptide oxytocin, by du Vigneaud (du Vigneaud et al., 1953) was acclaimed as a major accomplishment and spurred the advancement of peptide synthesis once again. The synthesis of the 39 residue porcine adrenocorticotropic hormone in 1963 (Schwyzer and Sieber, 1963) by solution-phase segment condensation methods was viewed as no less than sensational at the time; and in 1967, Bodanszky and colleagues succeeded in synthesizing the 27 residue secretin peptide by solution phase stepwise addition methods (Bodanszky and Williams, 1967; Bodanszky et al., 1967). These accomplishments in peptide synthesis were considered to be monumental at the time in that they were exceptionally difficult undertakings that pushed the prevailing technology to its limits.

In 1963, Bruce Merrifield published a landmark paper (Merrifield, 1963) describing the development of solid-phase peptide synthesis. This technique, for which he was awarded the Nobel Prize in

Chemistry in 1984, was responsible, more than anything else, for opening the way to the widespread use of synthetic peptides as reagents in chemical and biomedical investigations. Theodor Wieland (Wieland, 1981) once described the advances in peptide chemistry in the following way: "[T]he synthesis of glutathione opened the door to peptide synthesis a crack, and the synthesis of oxytocin pushed the door wide open." To extend that analogy, the introduction of the solid-phase method blew the door off its hinges. Since that time, many other developments, such as improved synthesis chemistry, automated instrumentation for the unattended production of peptides, and improved purification and analytical methods have contributed to our ability to exploit the potential of synthetic peptides. Not only has there been an explosion in the number of synthetic peptides being produced, but reports of the synthesis of larger and larger peptides (Gutte and Merrifield, 1969, 1971; Clark-Lewis et al., 1986; Nutt et al., 1988; Schneider and Kent, 1988, Muir et al., 1997), some exceeding 100 residues in length, are becoming more commonplace.

Now, and for the last 10 years or so, synthetic peptides are available not only to those actively involved in developing the field, but to virtually every investigator in any field who perceives a need for them and wishes to pursue their use.

This book is intended primarily for those investigators who wish to utilize synthetic peptides in their research but who themselves are not already intimately involved in the field. This group would encompass researchers who either simply want access to synthetic peptides as tools, or who wish to become actively involved in the production of the peptides, themselves. As such, it strives to provide practical background information, answer basic questions, and address common problems relating to all aspects of synthetic peptides. At the same time, however, it is intended to be current, comprehensive, and sophisticated in its treatment of the subject and should be of general interest to researchers and educators at all levels.

The process of obtaining a synthetic peptide for use in biochemical or biomedical research involves several discrete steps which can be represented diagrammatically as shown in figure 1-1. The process starts with the design of the peptide and follows through with the chemical synthesis, evaluation, and purification of the product, and finishes with application or the actual use of the peptide in an experimental situation. As indicated in the diagram, the process can be viewed as pivoting on the evaluation step which actually is part of a more comprehensive evaluation–purification cycle. It is at this point that the product is characterized and its disposition determined. That disposition can be either (1) use of the peptide as is, (2) additional purification and subsequent evaluation, eventually leading to experimental usage, or (3) changes in either the design or synthesis

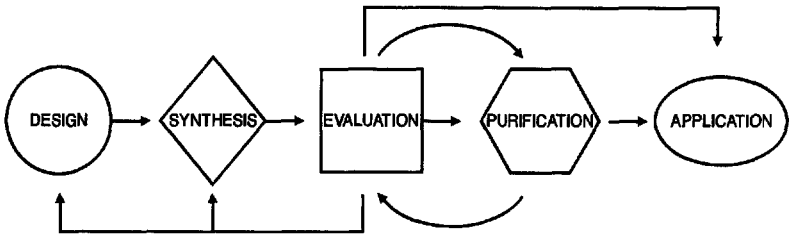


Figure 1-1. Diagram of the steps involved in obtaining a synthetic peptide.

protocols, or both, followed by reiteration of the latter steps of the process.

Each subsequent chapter of this book deals with one major aspect of this process, that is, design, synthesis, evaluation/purification, and application.

The book begins in chapter 2 with a discussion of the fundamental aspects of peptide structure and uses that as a foundation to build into both basic and more advanced considerations in the design of synthetic peptides. The design process is a constantly evolving one which is something of an art in itself. In many cases, there are no hard-and-fast rules and function usually dictates design. Perhaps the most commonly asked question from general users of synthetic peptides is how to choose a sequence for the production of antibodies. This chapter contains an expanded section on this aspect that draws from what has been learned from the last 10 years. In addition, a section on combinatorial peptide libraries has been added. This is an area that presents great promise for screening very large numbers of peptides for de novo discovery as a complement to rational design and for optimizing structure–activity relationships.

Chapter 3 deals with the synthesis chemistry itself. It presents a logical and straightforward explanation of the solid-phase method of peptide synthesis and discusses the overall state of the chemistry today. This chapter places particular emphasis on the Boc and Fmoc protection strategies as the two most useful approaches available and provides specific recommendations for undertaking routine synthesis. It is immediately evident that the chemistry is at the same time very sophisticated and complicated, yet relatively straightforward in the overall approach and amenable to routine procedures. Yet, one must always be mindful that, although routine synthesis using a uniform set of reagents is commonly performed successfully, a vast array of problems can occur. At the same time, there are a large number of alternatives that can be utilized to prevent or circumvent these problems. This chapter therefore presents both recommended "recipes" for routine

synthesis and more advanced chemical strategies and considerations for the maturing practitioner. This chapter also contains a useful section on the automation of solid-phase synthesis. It gives an informative overview of the range of instruments available today and illustrates the choice of capabilities, from single peptide synthesis to as many as 96 at one time and from microgram to gram quantities.

Chapter 4 deals with the evaluation of the finished peptide. It presents a discussion of the latest methods available for the routine characterization of the synthetic product and illustrates their utilization through examples of the evaluation of actual peptides. The message that should be very clear from this chapter is that although the chemistry can in many cases be preprogrammed on modern peptide synthesizers, and these machines do most of the work that used to be done laboriously on the benchtop, the process is by no means foolproof or trivial. Many problems from a large variety of sources can be manifest even in a "routine" synthesis and the final product can never be taken for granted. Every peptide produced must be rigorously evaluated with the ultimate goal being the proof that the peptide obtained is the one intended. In addition, problematic residues and peptide solubility are very important considerations in the evaluation and use of the peptides, which are very often ignored or overlooked. This chapter also contains a more in-depth description of mass spectrometry as an evaluation tool and highlights its use in determining the sequence of a peptide.

Chapter 5 presents an excellent survey of the use and application of synthetic peptides. It not only illustrates the diversity in the use and applications of synthetic peptides in modern research but, perhaps, serves as a basis from which new applications and ideas may develop. These include the antigenic and immunogenic use of synthetic peptides, the use of peptides as enzyme inhibitors, structure/function studies involving synthetic peptides, peptide-based vaccines, antisense peptides, and the use of peptides for affinity labeling of receptors or "acceptors" and their use in structure/function studies of receptors. Undoubtedly, some uses for synthetic peptides have not been included, but the areas that are discussed are among the more exciting and successful applications being investigated today. Furthermore, many of the concepts and approaches discussed can be easily adapted to other systems and areas of exploration. Also, since the use very often dictates the design, many aspects of peptide design are discussed in this chapter and, as such, it complements chapter 2 very nicely.

Peptide synthesis has come a long way from the beginning of the twentieth century to the present day. It has spanned the twentieth century and has enhanced the passing of that period. As we start the new millennium, it is tempting to speculate on what may lie ahead. However, history has told us that such attempts invariably fall far

short of the eventual reality. After all, even Merrifield admitted in a 1986 article (Merrifield, 1986) that he did not foresee the impact of his technique: "From the accumulated data presented, I conclude that the solid phase synthesis of peptides up to 50 or somewhat more residues can be readily achieved in good yield and purity; this is a far better situation than I could have expected when this technique was first proposed" (p. 345).

Undoubtedly, in the future great things will be accomplished in peptide synthesis and in the use of the peptides themselves. It is, perhaps, sufficient to have had the privilege of taking part in a small bit of that process.

References

Bergman, M., and Zervas, L. (1932). Über ein allgemeines Verfahren der Peptid-Synthese. Ber. Dtsch. Chem. Ges. 65:1192–1201.

Bodanszky, M., and Ondetti, M. A. (1966). *Peptide Synthesis*, New York, John Wiley and Sons.

Bodanszky, M., and Williams, N. J. (1967). Synthesis of secretin I. The protected tetradecapeptide corresponding to sequence 14-27. J. Am. Chem. Soc. 89:685–689.

Bodanszky, M., Ondetti, M. A., Levine, S. D., and Williams, N. J. (1967). Synthesis of secretin II. The stepwise approach. J. Am. Chem. Soc. 89:6753–6757.

Clark-Lewis, I., Aebersold, R., Ziltener, H., Schrader, J. W., Hood, L. E., and Kent, S. B. H. (1986). Automated chemical synthesis of a protein growth factor for hemopoietic cells, Interleukin-3. Science 231:134–139.

Fischer, E. (1902). Ueber einige Derivate des Glykocolls Alanins und Leucins. Ber. Dtsch. Chem. Ges. 35:1095–1106.

Fischer, E. (1903). Synthese von Derivaten der Polypeptide. Ber. Dtsch. Chem. Ges. 36:2094–2106.

Fischer, E. (1906). Untersuchungen über Aminosäuren, Polypeptide, und Proteïne. Ber. Dtsch. Chem. Ges. 39:530–610.

Fischer, E. (1907). Synthese von Polypeptiden XVII. Ber. Dtsch. Chem. Ges. 40:1754–1767.

Gutte, B., and Merrifield, R. B. (1969). The total synthesis of an enzyme with ribonuclease A activity. J. Am. Chem. Soc. 91:501.

Gutte, B., and Merrifield, R. B. (1971). The synthesis of ribonuclease A. J. Biol. Chem. 246:1922.

Harington, C. R., and Mead, T. H. (1935). Synthesis of glutathione. Biochem. J. 29:1602–1611.

Merrifield, R. B. (1963). Solid phase peptide synthesis I. The synthesis of a tetrapeptide. J. Am. Chem. Soc. 85:2149–2154.

Merrifield, R. B. (1986). Solid phase synthesis. Science 232:341–347.

Muir, T. W., Dawson, P. E., and Kent, S. B. H. (1997). Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. Meth. Enzymol. 289:266–298.

Nutt, R. F., Brady, S. F., Darke, P. L., Ciccarone, T. M., Colton, C. D., Nutt, E. M., Rodkey, J. A., Bennett, C. D., Waxman, L. H., Sigal, I. S., Anderson, P. S., and Veber, D. F. (1988). Chemical synthesis and enzymatic activity of a 99 residue peptide with a sequence proposed for the human immunodeficiency virus protease. Proc. Natl. Acad. Sci. U.S.A. 85:7129–7133.

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

Schwyzer, R., and Sieber, P. (1963). Total synthesis of adrenocorticotrophic hormone. Nature 199:172–174.

Sifferd, R. H., and du Vigneaud, V. (1935). A new synthesis of carnosine, with some observations on the splitting of the benzyl group from carbobenzoxy derivatives and from benzylthio ethers. J. Biol. Chem. 108:753–761.

Vigneaud, V. du, Ressler, C., Swan, J. M., Roberts, C. W., Katsoyannis, P. G., and Gordon, S. (1953). The synthesis of an octapeptide amide with the hormonal activity of oxytocin. J. Am. Chem. Soc. 75:4879–4880.

Wieland, T. (1981). From glycylglycine to ribonuclease: 100 years of peptide chemistry. In *Perspectives in Peptide Chemistry*, A. Eberle, R. Geiger, and T. Wieland, eds., Basel, Karger, pp. 1–13.

Wieland, T., and Bodanszky, M. (1991). *The World of Peptides: A Brief History of Peptide Chemistry*, Berlin, Springer-Verlag.

2

Peptide Design Considerations

Michael L. Moore
Gregory A. Grant

Peptides have become an increasingly important class of molecules in biochemistry, medicinal chemistry, and physiology. Many naturally occurring, physiologically relevant peptides function as hormones, neurotransmitters, cytokines, and growth factors. Peptide analogs that possess agonist or antagonist activity are useful as tools to study the biochemistry, physiology, and pharmacology of these peptides, to characterize their receptor(s), and to study their biosynthesis, metabolism, and degradation. Radiolabeled analogs and analogs bearing affinity labels have been used for receptor characterization and isolation. Peptide substrates of proteases, kinases, phosphatases, and aminoacyl or glycosyl transferases are used to study enzyme kinetics, mechanism of action, and biochemical and physiological roles and to aid in the isolation of enzymes and in the design of inhibitors. Peptides are also used as synthetic antigens for the preparation of polyclonal or monoclonal antibodies targeted to specific sequences. Epitope mapping with synthetic peptides can be used to identify specific antigenic peptides for the preparation of synthetic vaccines, to determine protein sequence regions that are important for biological action, and to design small peptide mimetics of protein structure or function.

A number of peptide hormones or analogs thereof, including arginine vasopressin, oxytocin, luteinizing hormone releasing hormone

(LHRH), adrenocorticotropic hormone (ACTH), and calcitonin, have already found use as therapeutic agents, and many more are being investigated actively. Peptide-based inhibitors of proteolytic enzymes, such as angiotensin converting enzyme (ACE) and human immunodeficiency virus (HIV) protease, have widespread clinical use, and inhibitors of renin and elastase are also being investigated for therapeutic use. Finally, peptides designed to block the interaction of protein molecules by mimicking the combining site of one of the proteins, such as the fibrinogen receptor antagonists, show great therapeutic potential as well.

With the development of solid-phase peptide synthesis by Bruce Merrifield (1963) and the optimization of supports, protecting groups, and coupling and deprotection chemistries by a large number of researchers, it has become possible to obtain useful amounts of peptides on a more or less routine basis. With the increasing ease of synthesizing peptides, it has become all the more important to understand the underlying principles of peptide structure and physical chemistry that govern solubility, aggregation, proteolytic resistance or susceptibility, secondary structure stabilization or mimicry, and interaction with nonpolar environments like lipid membranes.

The design of any specific peptide depends primarily on the use for which it is intended as well as on synthetic considerations. A number of aspects of the peptide sequence and structure can be manipulated to affect solubility, proteolytic resistance, or stability, and the predilection to adopt specific secondary structures to produce peptides with specific properties. It is this topic that will be addressed in this chapter.

The distinction between what constitutes a peptide and what constitutes a protein becomes increasingly fuzzy as peptides increase in length. An operational definition in the context of this chapter might be that a peptide is any sequence that the researcher can conveniently synthesize chemically. Although such proteins as ribonuclease A (124 amino acids) (Hirschmann et al., 1969; Yajima and Fujii, 1981), acyl carrier protein (74 amino acids) (Hancock et al., 1972), and the HIV protease (99 amino acids) (Nutt et al., 1988; Wlodawer et al., 1989) have been chemically synthesized, most peptide synthesis generally involves peptides of 30 amino acids or less. The underlying principles applied to peptide design will be dependent in some degree on peptide length, especially those relating to peptide secondary structure. For example, shorter peptides do not tend to exhibit preferred solution conformations and possess a large amount of segmental flexibility. If secondary structure is important in a small peptide's design, it must be approached by chemical modifications designed to decrease conformational flexibility. As peptides increase in length, they have a greater tendency to exhibit elements of secondary structure, with a consequent

decrease in segmental flexibility. In such cases, secondary structure can often be induced by optimizing the peptide sequence.

Protein and Amino Acid Chemistry

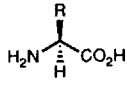
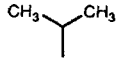
The physical and chemical properties of proteins and peptides are determined by the nature of the constituent amino acid side chains and by the polyamide peptide backbone itself. Twenty protein amino acids are coded for by DNA, which are translationally incorporated into proteins. Amino acids can be modified in the protein post-translationally to yield new amino acids. Also, many peptides are synthesized enzymatically rather than ribosomally, especially in lower eukaryotes, and those peptides often contain highly unusual amino acids.

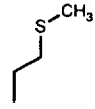
Protein Amino Acids

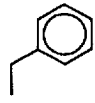
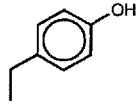
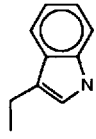
The structures of the 20 primary protein amino acids (those coded for by DNA) are given in table 2-1 along with their three-letter abbreviations, one-letter codes, and a general grouping by physical properties. Amino acids generally can be divided into hydrophobic and hydrophilic residues. The hydrophobic residues include those with aliphatic side chains, such as alanine, valine, isoleucine, leucine, and methionine, and those with aromatic side chains, such as phenylalanine, tyrosine, and tryptophan. The hydrophilic residues include amino acids with (1) neutral, polar side chains, such as serine, threonine, asparagine, and glutamine; (2) those with acidic side chains, such as aspartic acid and glutamic acid; and (3) those with basic side chains such as histidine, lysine, and arginine. It can be appreciated that these categories are not entirely exclusive. Alanine, with its small aliphatic side chain, and glycine can be found in hydrophilic regions of peptides and proteins. Conversely, the long alkyl chains of lysine and arginine can give those residues an overall hydrophobic character with just the terminal charged group being hydrophilic.

Two amino acids, cysteine and proline, have special properties that set them apart. Cysteine contains a thiol moiety that can be oxidatively coupled to another cysteine thiol to form a disulfide linkage. Disulfides are the principal entities by which peptide chains are covalently linked together to stabilize secondary or tertiary structure or to hold two different peptide chains together. Although the disulfide form is the most stable form under normal aerobic conditions, free thiols are also present in some proteins, where they often serve as ligands for metal chelation, as nucleophiles in proteolytic enzymes, such as papain, or as carboxyl activators in acyl transferases. The secondary amino acid proline has specific conformational effects on the peptide or protein

Table 2-1 Amino acid structures and properties

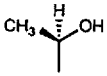
		
		
Glycine Gly G	Alanine Ala A	Valine Val V
Neutral, hydrophobic, aliphatic		

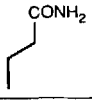
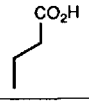
		
Isoleucine Ile I	Leucine Leu L	Methionine Met M
Neutral, hydrophobic, aliphatic		

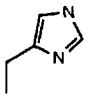
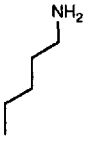
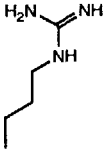
		
Phenylalanine Phe F	Tyrosine Tyr Y	Tryptophan Trp W
Neutral, hydrophobic, aromatic		

(*continued*)

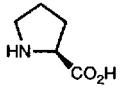
Table 2-1 (*Continued*)

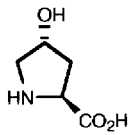
		
Serine Ser S	Threonine Thr T	Asparagine Asn N
Neutral, hydrophilic		

		
Glutamine Gln Q	Aspartic Acid Asp D	Glutamic Acid Glu E
Neutral, hydrophilic	Acidic, hydrophilic	

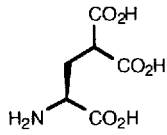
		
Histidine His H	Lysine Lys K	Arginine Arg R
Basic, hydrophilic		


Cysteine Cys C
Thiol-containing

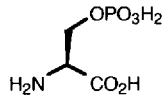

Proline Pro P
Imino acid



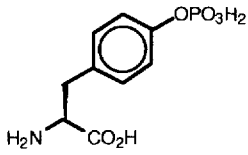
trans-4-Hydroxyproline



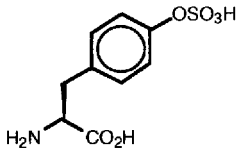
γ-Carboxyglutamic Acid



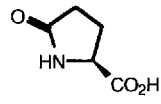
Serine O-Phosphate



Tyrosine O-Phosphate



Tyrosine O-Sulfate



Pyroglutamic Acid

Figure 2-1. Some amino acids formed by post-translational modification.

backbone both because of its cyclic structure and because of the alkylation of the amino group. It often plays an important role in stabilizing or influencing the secondary structure of proteins, as will be discussed later.

Some amino acids can be modified enzymatically after incorporation into certain proteins to give rise to new amino acids, some of which are shown in figure 2-1. In collagen, for example, proline can be hydroxylated to yield *trans*-4-hydroxyproline (Hyp). Glutamic acid is carboxylated by a vitamin K-dependent carboxylase to yield γ-carboxyglutamic acid (Gla) in a number of proteins involved in blood coagulation, where the malonyl moiety of the Gla residue is thought to be important in providing a bidentate ligand for calcium ions. The hydroxyl functions of tyrosine, serine, and threonine can be reversibly phosphorylated by kinases and phosphatases, a process thought to be an important regulator of biological activity in the target proteins. The hydroxyl of tyrosine may also be sulfated in peptide hormones such as gastrin and cholecystokinin.

One important nonenzymatic transformation of a protein amino acid occurs with glutamine, which is chemically unstable at the amino terminus of a peptide or protein in aqueous solution. Glutamine will spontaneously cyclize to form pyroglutamic acid (pyrrolidone carboxylic acid, also shown in figure 2-1). This transformation typically occurs when an internal glutamine residue is exposed by proteolysis of an X-Gln bond (where X represents any amino acid) or during protein sequencing.

All amino acids (except for glycine) are chiral molecules with an asymmetric center at the α-carbon. The protein amino acids all have

the L absolute configuration at this center, as shown in the structure at the top of table 2-1. The L does not refer to the direction of optical rotation but, rather, to having the same stereochemical arrangement as L(-)-glyceraldehyde. In the more unambiguous Cahn–Ingold–Prelog convention, L-amino acids have the *S* absolute configuration (except cysteine, for which this configuration is defined as *R*).

Two amino acids, threonine and isoleucine, have a second asymmetric center at the β-carbon, as shown in table 2-1. Threonine has the *R* absolute configuration at the β-carbon; that is, it is 2-(*S*)-amino-3-(*R*)hydroxybutanoic acid. Isoleucine has the *S* absolute configuration at the β-carbon; that is, it is 2-(*S*)-amino-3-(*S*)-methylpentanoic acid. Because there are two asymmetric centers in the molecule, four stereoisomers exist for each of these amino acids. The D-amino acid has the opposite configuration at both asymmetric centers; D-isoleucine is the 2-(*R*)-3-(*R*) isomer, for example. The trivial prefix allo is used to denote inversion at only one asymmetric center. Thus D-allo-isoleucine is the 2-(*R*)-3-(*S*) isomer. D-allo-Isoleucine occasionally forms as an artifact during hydrolysis of isoleucine-containing peptides or proteins. Because it is a diastereomer of isoleucine rather than an enantiomer, it will appear as a separate peak in amino acid analyses.

Nonprotein Amino Acids

Literally hundreds of naturally occurring amino acids exist that are not found in proteins. Some of them, like γ-aminobutyric acid, have important functions as neurotransmitters. Others, like ornithine, appear as intermediates in metabolic pathways or, like dihydroxyphenylalanine, are precursors to amino acid-derived products, catecholamines in this case.

By far the largest number of nonprotein amino acids are found in prokaryotes and lower eukaryotes, especially in algae, sponges, yeasts, and fungi, although peptides with unusual amino acids have been isolated from chordates, such as tunicates, as well. These amino acids are incorporated into peptides by enzymatic synthesis rather than ribosomally. More than 700 of these nonprotein amino acids are known, and the structural variations are immense (Hunt, 1985). Those amino acids and the peptides containing them are usually the products of secondary metabolism, and their function in the producing organism is often obscure. Those unusual amino acids can confer unusual biological activities on peptides that contain them. Such peptides have been the basis either of biochemical tool molecules or of antibiotics, immunomodulators or antineoplastic agents that have therapeutic utility.

Nonprotein amino acids can be roughly categorized into a few basic structural types, representative members of which are depicted in

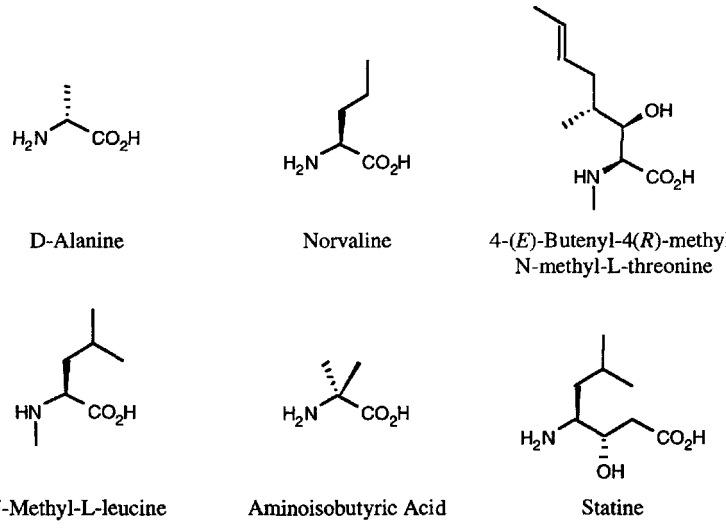


Figure 2-2. Some nonprotein amino acids.

figure 2-2. Some nonprotein amino acids are simply the enantiomeric D-amino acid analog of a protein L-amino acid, for example, D-alanine and D-glutamic acid, which are important constituents of the proteoglycan bacterial cell wall. Others have a normal α-amino acid structure but with a novel side chain. The side chain can be a simple alkyl group, such as in norvaline, or it can be quite unusual, such as in 4-(*E*)-butenyl-4(*R*)-methyl-*N*-methyl-L-threonine (MeBmt), a critical constituent of the immunosuppressive peptide cyclosporin A (Rüegger et al., 1976). Some nonprotein amino acids deviate from the normal α-amino acid structure and are methylated on their amine function, such as the *N*-methylleucine (MeLeu) of the tunicate-derivated antineoplastic peptide didemnin B (Rinehart et al., 1981), or on their α-carbon, such as the aminoisobutyric acid (α-methylalanine, Aib) of the ionophoretic antibiotic peptide alamethicin (Payne et al., 1970; Pandey et al., 1977). In addition, there are large numbers of amino acids in which the amino group is not on the α-carbon but at some other position in the molecule. Sometimes several of these features are combined in one amino acid, such as statine [3-(*S*)hydroxy-4-(*S*)-amino-6-methylheptanoic acid, Sta], the critical residue in the fungal protease inhibitor pepstatin (Morishima et al., 1970).

Side-Chain Interactions

Side chains interact with each other, with the amide peptide backbone, with bulk solvent, and through noncovalent interactions, such as

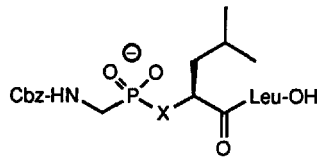
hydrogen bonds, salt bridges, and hydrophobic interactions. Cysteine also participates in a covalent interaction, disulfide bond formation.

Hydrogen Bonds and Salt Bridges

In proteins, polar side chains tend to be extensively solvated. Acidic (Asp and Glu) and basic (Lys, Arg, and His) residues generally are found on the protein surface with the charged ends of the side chains projecting into the bulk solvent, although the alkyl portion of the Lys and Arg side chains is usually buried. Internal charged residues are almost invariably involved in salt bridges, where acidic and basic side chains are either directly bonded ionically to each other or connected by a single intermediary water molecule (Baker and Hubbard, 1984). Nonionic polar residues (Ser, Thr, Asn, Gln, and Tyr at the phenolic hydroxyl) are also extensively hydrogen-bonded, either to bulk solvent or to backbone, other side-chain groups, or to specifically bound water molecules (Thanki et al., 1990). In helices, the side chains of Ser, Thr, and Asn often make specific hydrogen bonds to the carbonyl oxygen of the third or fourth residue earlier in the sequence, which may help to stabilize helical segments (Gray and Matthews, 1984). In shorter peptides, side-chain hydration occurs mostly through the bulk solvent, although the ability to form low-energy intramolecular hydrogen-bonded structures or salt bridges may be an important factor in the association of peptides with macromolecular targets or receptors.

Consideration of solvation and desolvation effects is a potentially critical, but often overlooked, aspect of peptide design and structure–activity relationships. One case will serve to illustrate the importance of solvation in interpreting structure–activity studies. Bartlett and coworkers prepared a series of thermolysin inhibitors that were based on the peptide substrate sequence Cbz-Gly-Leu-Leu-OH but contained a phosphoryl moiety in place of the Gly carbonyl function (Morgan et al., 1991). The phosphoryl moiety was designed to mimic the tetrahedral transition state for amide bond hydrolysis and was suggested by a naturally occurring glycopeptide inhibitor of thermolysin, phosphoramidon, which contains a similar phosphoryl group. The inhibitors that were prepared are shown in table 2-2. Compound 1, which contains a phosphoramidate linkage (PO_2NH), was found to be a potent inhibitor. X-ray crystal structure analysis of the enzyme-inhibitor complex showed a hydrogen bond between the phosphoramidate NH and a backbone carbonyl in the enzyme. Compound 2 contained a phosphonate linkage (PO_2O). It was unable to form the same hydrogen bond as compound 1 because it lacks the corresponding hydrogen. It was found to be three orders of magnitude less potent than compound 1, but X-ray crystal analysis showed that it bound to the enzyme in an identical fashion (except for the absence of the

Table 2-2 Phosphorus-containing thermolysin inhibitors



Compound	*X*	*Ki(nM)*
1	NH	9.1
2	O	9000
3	CH_2	10.6

hydrogen bond). Compound 3 contained a phosphinate linkage (PO_2CH_2). Like compound 2, it could not form that same hydrogen bond. Unlike compound 2, however, it was essentially equipotent with the phosphoramidate 1. The explanation comes from consideration of solvation and desolvation effects. Both the phosphoramidate (1) and the phosphonate (2) will be solvated in aqueous solution and require desolvation to bind to the enzyme active site. The phosphoramidate (1) can recover some of the energy required for desolvation by forming a hydrogen bond with the enzyme, but the phosphonate (2) cannot. The phosphinate (3) is somewhat less polar due to the presence of a methylene group instead of an amine or oxygen moiety. It is correspondingly less solvated in aqueous solution, and, therefore, it requires less desolvation when it binds to the enzyme. The net change in free energy in going from solution to enzyme-bound states is roughly comparable for compound 1 (with the PO_2X group going from solvated to solvated states) and compound 3 (going from unsolvated to unsolvated states) and is more favorable than for compound 2 (going from solvated to unsolvated states).

Hydrophobic Interactions

Just as hydrophilic residues tend to be solvated, hydrophobic side chains have an equally strong tendency to avoid exposure to the aqueous environment. This effect is largely entropic, reflecting the unfavorable free energy of forming a water–hydrocarbon interface, where the side chain would penetrate the aqueous solvent (Tanford, 1973; Burley and Petsko, 1988). It has long been recognized that the interiors of soluble proteins are highly hydrophobic (Kauzmann, 1959)

and that proteins fold in such a way as to minimize the exposure of hydrophobic side chains on the protein surface. It has been proposed that protein folding proceeds first through the formation of hydrophobic clusters which direct further folding of the peptide chain into the various low-energy secondary structures, such as helices and β-structure (Rose and Roy, 1980). The importance of hydrophobic interactions to protein stability has been studied by comparing the susceptibility to denaturation of a series of proteins in which site-specific mutagenesis was used to modify single residues in the protein sequence (Kellis et al., 1988). It was found that the absence of even a single methyl group (Ile to Val substitution) destabilized the protein by 1.1 kcal/mol, underscoring the important cumulative effect of hydrophobic interactions on overall protein structure and stability.

Although the sequestering of nonpolar residues away from the aqueous environment is largely an entropy-driven process, specific interactions of hydrophobic side chains occur as well. These are typically induced dipole-induced dipole interactions, known as van der Waals or London interactions (Burley and Petsko, 1988). Although they are much weaker than the salt bridges and hydrogen bonds involving polar residues, they can be important in local secondary structure and protein interactions. For example, a number of DNA-binding proteins that function biologically as dimers share an unusual sequence, featuring a leucine every seventh residue in a 30-residue segment and having a relatively high probability of helical structure. This sequence would create a helix in which the leucine residues occupied every other turn of the same side of the helix. The structure was termed a "leucine zipper" based on the hypothesis that the leucines from such a helix in two monomers could interdigitate like a zipper, holding the monomers together (Landschutz et al., 1988). Although it was subsequently found that the Leu residues do not interdigitate, but rather align themselves parallel to each other in pairs along the helical interface (O'Shea et al., 1989), the term leucine zipper has remained, and it is an important structural motif in hydrophobic protein-protein interaction.

Aromatic residues have an inherent dipole with the electron-rich π-cloud lying parallel to and above and below the plane of the ring and with positively polarized hydrogen atoms in the plane of the ring (Burley and Petsko, 1988). Although aromatic rings in proteins do interact with each other, the arrangement is not typified by the parallel stacking of base pairs in DNA helices. Rather, the edge of one ring interacts with the face of the other in a roughly perpendicular arrangement (Burley and Petsko, 1985), allowing a favorable interaction between the positively polarized hydrogens and the negatively polarized π-cloud.

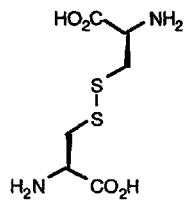
Peptides, in general, are not long enough to allow the hydrophobic residues to arrange themselves in such a way that they are totally

shielded from solvent. This occurrence undoubtedly contributes to the poor aqueous solubility of many peptides compared to proteins. Many peptides require the presence of strong organic co-solvents like dimethylsulfoxide (DMSO), dimethylformamide (DMF), or ethanol to achieve sufficient solubility for biological testing, which may be a limiting factor in some biological test systems. Peptides that have substantial hydrophobic character also tend to aggregate with increasing concentration. Solubility generally increases with peptide length because of the peptide's consequent ability to adopt stable secondary structures and to segregate nonpolar residues.

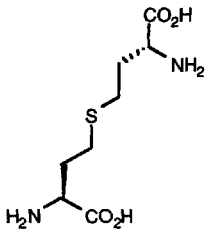
Disulfide Bonds, Thioesters, and Thioethers

Unlike side-chain interactions of other residues that are noncovalent in nature, cysteine residues form a number of covalent linkages with other amino acid side chains. The most common of these is the disulfide, which involves oxidative coupling of two cysteine thiol groups to form cystine (figure 2-3). Like the amide bond, the sulfur-sulfur bond in a disulfide is not freely rotatable. Rather, it exists in one of two rotamers, with torsional angles in the vicinity of either $+90°$ or $-90°$. The entire disulfide moiety, CH-S-S-CH_2, can rotate as a unit by simultaneous rotation about the side-chain angles $\chi 1$ (NH-CαH-CH_2-S) and $\chi 2$ (CαH-CH_2-S-S). In proteins, disulfide bonds are important in the stabilization of tertiary structure. In disulfide-containing peptides like oxytocin, vasopressin, and somatostatin (see figure 2-4), the disulfide has proportionally greater importance in maintaining a biologically active conformation because the peptides are too small to maintain a stable conformation otherwise. Replacement of the cysteine residues with isosteric alanine residues results in a dramatic loss in biological activity (Walter et al., 1967; Polàcek et al., 1970; Sarantakis et al., 1973). A great deal of conformational flexibility still remains in cyclic disulfide-containing peptides like vasopressin, and it does not exhibit a single preferred conformation (Hagler et al., 1985), but the disulfide does constrain the peptide to folded conformations that otherwise would have a low probability for linear peptides.

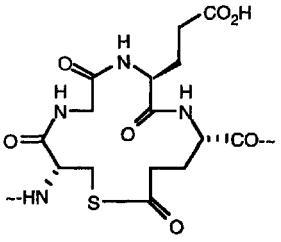
Cysteine residues can also form covalent thioester and thioether linkages. The thioester is found as the activated form of acyl groups in acyl transferases and as an acyl enzyme intermediate in thiol proteases such as papain. Thioesters also are found in the reactive binding sites of complement C3b and α2-macroglobulin, where the tetrapeptide sequence Cys-Gly-Glu-Glu contains a thioester involving the Cys-1 thiol and the Glu-4 side-chain carboxylate (figure 2-3) (Sothrup-Jensen et al., 1980; Tack et al., 1980; Howard, 1981). Cysteine forms aliphatic thioether linkages, such as in lanthionine, a



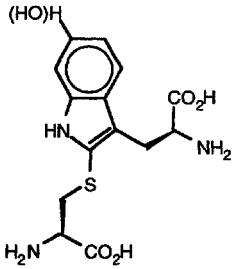
Cystine



meso-Lanthionine

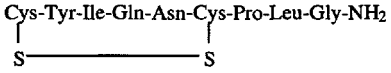


Cys-Gly-Glu-Glu thioester

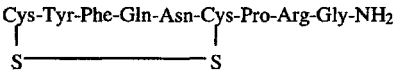


Cys-Trp thioether of phalloidin
(or amanitin)

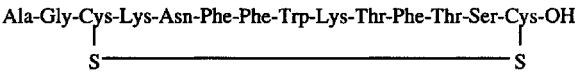
Figure 2-3. Disulfide, thioester, and thioether linkages.



Oxytocin



Arginine Vasopressin



Somatostatin

Figure 2-4. Disulfide-containing peptides.

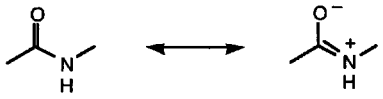
constituent of the peptide antibiotics subtilin (Alderton and Fevold, 1951) and nisin (Berridge et al., 1952). Aromatic thioether linkages also occur, such as in the Cys-Trp conjugate found in mushroom toxins such as amanitin and phalloidin (Wieland, 1968). A novel thioether Cys-Tyr conjugate recently has been identified in the enzyme galactose oxidase (Ito et al., 1991).

The Amide Bond

The polyamide peptide backbone is also an important contributor to overall protein and peptide structure. A substantial double-bond character is found in the carbon-nitrogen peptide bond due to the resonance structure shown in figure 2-5(a). This structure gives the amide bond several characteristics that are important in peptide and protein structure. The amide bond is flat, with the carbonyl carbon, oxygen, nitrogen, and amide hydrogen all lying in the same plane. No free rotation occurs about the carbon-nitrogen bond because of its partial double-bond character (the barrier to rotation is about 25 kcal/mol). The torsional angle of that bond, ω, is defined by the peptide backbone atoms Cα-C(O)-N-Cα. Because of the partial double-bond character, there are two rotational isomers for the peptide bond: *trans* ($\omega = 180°$) and *cis* ($\omega = 0°$), as shown in figure 2-5(b). The lower-energy isomer is the *trans* peptide bond, which is the isomer generally found for all peptide bonds not involving proline. In the case of amide bonds involving proline, the energy of the *trans* X-Pro bond is somewhat elevated, and both the difference in energy between *cis* and *trans* isomers and the barrier to rotation is lowered. Proline-containing peptides thus will often exhibit *cis-trans* isomerism about the X-Pro bond. This can be detected by nuclear magnetic resonance (NMR) studies because the chemical shifts of some hydrogens, especially those on the proline δ-carbon, are often different. The equilibrium ratio of *cis* to *tans* isomers and the rate of isomerization are highly dependent on the exact peptide sequence. Empirically, it has been found that the *cis* content of X-Pro peptides generally increases when X is a bulky, hydrophobic amino acid (Harrison and Stein, 1990).

It has been appreciated only recently that *cis* X-Pro peptide bonds may be important in proteins as well. The X-ray crystal structure of several proteins has revealed specific *cis* X-Pro peptide bonds (Frömmel and Preissner, 1990). Proline isomerization has also been implicated as a slow step in protein refolding (Kim and Baldwin, 1982).

The resonance forms of the amide bond give it one other characteristic that is extremely important in peptide and protein structure. The



(a) Resonance forms of amide bond

(b) *trans* and *cis* amide bonds

Figure 2-5. Characteristics of amide bonds.

amide bond is quite polar and has a significant dipole moment, which makes the amide carbonyl oxygen a particularly good hydrogen-bond acceptor and the amide NH a particularly good hydrogen-bond donor. Hydrogen bonds involving the peptide backbone are an important stabilizing factor in protein secondary structures. Peptide bonds have a strong tendency to be solvated, by either bulk solvent or specifically bound water molecules, or by internal hydrogen bonds. This is especially true in the hydrophobic interior of soluble protein molecules. Peptide bonds of regular secondary structures, such as helices and β-sheets, are internally solvated by the hydrogen bonds that stabilize those structures. The peptide bonds of the so-called random coil, or irregular structures, also participate in an extensive network of hydrogen bonds involving internal polar side chains, bound water molecules, and backbone interactions.

The polarity of the amide bond also can impart a net dipole to regular structures containing peptide bonds, such as helices. The overall dipole points from the amino terminus (positive partial charge) to the carboxyl terminus (negative partial charge) in a helix, which can be an important factor in protein tertiary structure as well as a contributor to catalytic activity in enzymes. The dipole can be used to stabilize interactions with substrates or to modify the pK_a, or nucleophilicity, of catalytically active residues (Knowles, 1991).

Protein Structure

Protein structure is organized in several hierarchical levels of increasing structural complexity. The most basic level is the primary structure, or amino acid sequence, of the protein. Secondary structure deals with the folding up of short segments of the peptide, or protein chain, into regular structures such as α-helices, β-sheets, and turns. Tertiary structure describes how the secondary structural elements of a single protein chain interact with each other to fold into the native protein structure. Quaternary structure involves the interaction of individual protein subunits to form a multimeric complex.

In the design of peptides, we are concerned mainly with the primary and secondary structures. The primary structure, or sequence, of a protein, is readily obtainable by chemical sequence analysis of the isolated, purified protein or by translation of the corresponding DNA coding sequence. Since it is easier to sequence DNA than to sequence protein, most new protein sequences are now obtained in this way. Computer programs are available, which will be discussed later (see below, Prediction of Protein/Peptide Structure), to help analyze and compare protein primary structures. Protein secondary structure has also become increasingly well understood, to the point that it can be predicted or optimized on the basis of the primary sequence reasonably well by using computer algorithms. In addition, several techniques have been applied to either induce or stabilize secondary structure in peptides, which will be discussed in more detail later (see below, Conformational Design and Constraint).

Protein Secondary Structure

Chains of amino acids can fold into several types of regular structures that are stabilized by intrachain or interchain hydrogen bonds in the amide backbone. Helices and turns are formed from continuous regions of protein sequence and are stabilized by intrachain hydrogen bonds. Sheets are formed from two or more chains, which are separated by intervening sequences (regions of secondary structure as well), and are stabilized by interchain hydrogen bonds. The parts of a protein sequence that do not appear in helices, sheets, or turns are said to be in random coil, which is meant to imply only that no regular, repeating structure exists. It should not be taken to mean that the conformation of the protein is random in those regions, because in any individual protein, the random coil sequences are just as highly ordered and reproducibly formed as other regions of secondary structure.

The conformation of the peptide backbone can be described by three torsional angles: ϕ, which is the angle defined by C(O)-N-Cα-C(O); ψ, which is defined by N-Cα-C(O)-N; and the amide torsional angle ω,

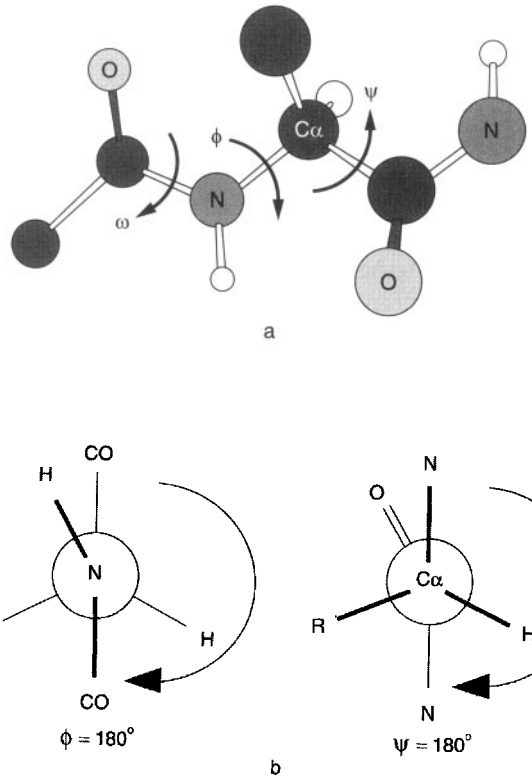


Figure 2-6. Torsional angles of the peptide backbone. (a) Backbone torsional angles ϕ, ψ, and ω. (b) Newman projections of torsional angles ϕ and ψ.

which was discussed previously (see above, The Amide Bond). These angles are shown graphically in figure 2-6(a). The convention in protein chemistry is that these angles are defined with respect to the peptide backbone. The angle ϕ is 180° when the two carbonyl carbons are *trans* to each other, and the angle ψ is 180° when the two amide nitrogens are *trans* to each other as in figure 2-6(b). The angle ψ, therefore, is +180° compared to the usual chemical definition. A fully extended peptide chain, however, would have backbone torsional angles all of 180°, just like a fully extended carbon chain.

Helices

One of the most common types of secondary structure in proteins is the helix in which amino acid residues are wrapped around a central

Table 2-3 Torsional angles for helices

	ϕ	ψ
α-Helix (right-handed)	−57°	−47°
α-Helix (left-handed)	57°	47°
3_{10} Helix	−60°	−30°
Collagen helix	−51°	153°
	−76°	127°
	−45°	148°
Polyproline	−78°	149°

Angles taken from IUPAC-IUB Commission on Biochemical Nomenclature (1970).

axis in a regular pattern. Given planar, trans peptide bonds, helices will be uniquely defined by their ϕ and ψ angles. Several types of helices are found in protein structures, the α-helix being the most common. These helices and their characteristic torsional angles are listed in table 2-3. Helices are also characterized by the number of residues per turn and the hydrogen bonding pattern, specifically the number of atoms in the cyclic structure formed by the hydrogen bond. The α-helix is a 3.6_{13} helix, meaning there are 3.6 residues per turn of the helix, and 13 atoms are included in each repetitive hydrogen-bonding structure. Likewise, a 3_{10} helix has three residues per turn and 10 atoms in the cyclic hydrogen-bonded structure. These are the two most relevant helices for protein and peptide structure. There is also a handedness to helical structure, which is defined by the screw sense of the helix. Looking down the helical axis from the amino terminal end, the peptide backbone can be traced from amino terminus to carboxyl terminus in a clockwise sense in a right-handed helix, which is the form that generally occurs.

Helical structures are stabilized by intrachain hydrogen bonds. In a helix, the carbonyl bonds and amide NH bonds lie parallel to the helix's axis with the carbonyls pointing downward, in the direction of the carboxyl terminus of the helix, and the NHs pointing upward, in the direction of the amino terminus. In an α-helix, the carbonyl of any residue i is hydrogen-bonded to the amide NH of the $i + 4$ residue; that is, the residue four amino acids farther down along the peptide chain. In a 3_{10} helix, the ith residue carbonyl is hydrogen-bonded to the NH of the $i + 3$ residue. These relationships are shown graphically in figures 2-7 and 2-8 for six-residue segments of α and 3_{10} helices as viewed from the side, perpendicular to the helical axis. For clarity, only the peptide backbone atoms are shown.

The helix forms a polyamide cylinder. The uniform arrangement of carbonyl and NH groups imparts a strong dipole moment to the helix,

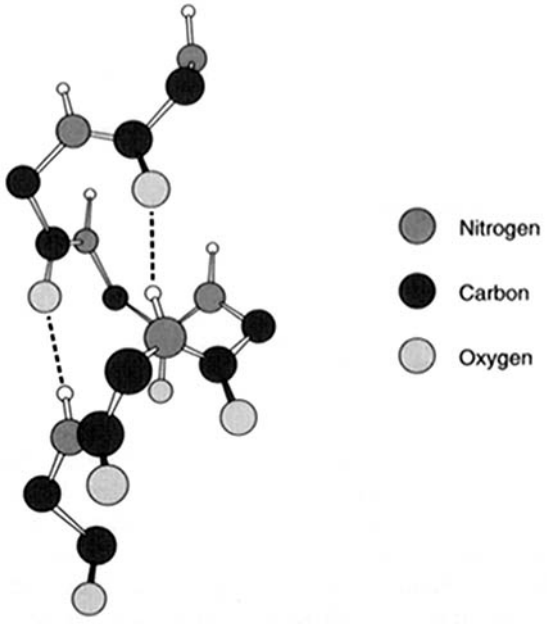
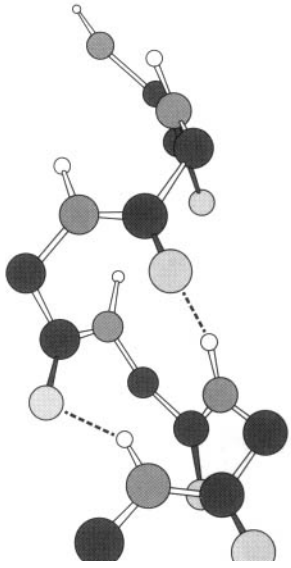
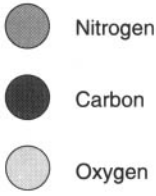
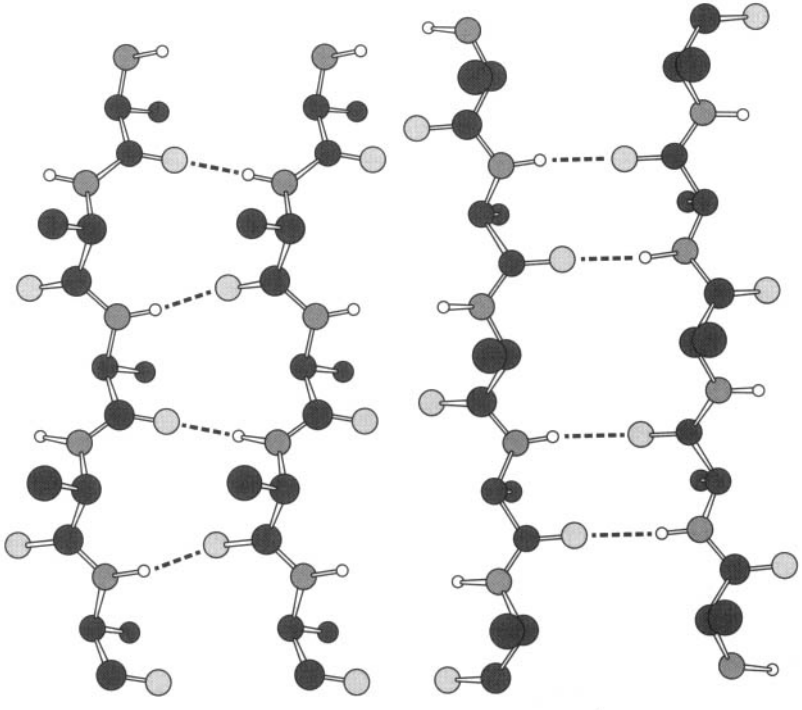


Figure 2-7. Backbone structure of an α-helix. Note that the NH and carbonyl groups are aligned parallel to the axis of the helix. Each hydrogen-bonded segment contains 13 atoms.

with the positive end at the amino terminus and the negative end at the carboxyl terminus. The side chains of the residues are arranged radially outward from the helix, looking down the helix axis (figure 2-9). Because all the backbone amide groups are involved in intrachain hydrogen bonds, the interactions of helices with other peptide chains or small molecules occurs predominantly through side-chain interactions (hydrophobic interactions, salt bridges, and hydrogen-bonding interactions; see above, Side-Chain Interactions) and interactions with the helix dipole itself. Because of this arrangement of intrachain hydrogen bonds, the amide bonds of the helix can be thought of as being internally solvated, resulting in the helix being more readily accommodated in a nonpolar environment, such as a lipid bilayer. The membrane-spanning regions of transmembrane proteins are generally thought to be helical, with the axis of the helix perpendicular to the plane of the membrane (Eisenberg, 1984; Wickner and Lodish, 1985). The membrane pore-forming peptide antibiotics, such as alamethicin, also have a highly helical structure, and a high degree of helix potential exists in the







opposite directions. Both can be found in protein structures. These two arrangements are shown diagrammatically in figure 2-10.

The β-structure can be thought of as forming a surface or sheet, although there is a slight twist to it because the peptide backbone is

Figure 2-9. View of an α-helix looking down the helical axis. The residue side chains, shown in black, all project radially outward from the helix.

not fully extended (typical angles are given in table 2-4). The side chains extend above and below the rough plane of the sheet, with every other side chain on one surface. Because the interchain hydrogen bonds between two adjacent chains involve only every other residue, there is the possibility of larger structures forming that involve many strands. Such extended β-structures are common features in proteins. The twist of the β-structure also allows sheets to fold into cylindrical structures, called β-barrels, which is another common structural motif.

Turns

Strands of α-helix or β-structure do not extend indefinitely in proteins but rather fold back on themselves. Several regular structures, called turns, are involved in changing the direction of the peptide chain. Turns are classified by the number of residues that are involved in the regular structure; β-turns contain four amino acid residues, while γ-turns contain three residues. Each structure is stabilized by a hydrogen bond extending across the turn, in effect holding the two ends together. The first residue of a turn is usually designated as i. In a β-turn, the hydrogen bond is between the carbonyl of the i residue and the NH of the $i + 3$ residue, giving the equivalent of a 10-membered ring. There