CHAPTER TWELVE

NMR and Drug Discovery

DAVID J. CRAIK
RICHARD J. CLARK
Institute for Molecular Bioscience
Australian Research Council Special Research
Centre for Functional and Applied Genomics
University of Queensland
Brisbane, Australia

Contents

1 Introduction, **508**
- **1.1** Overview of Drug Development, **509**
- **1.2** Scope of Chapter, **510**
- **1.3** Principles of NMR Spectroscopy, **510**
- **1.4** Instrumentation, **514**
- **1.5** Applications of NMR in Drug Design and Discovery, **516**

2 Ligand-Based Design, **517**
- **2.1** Structure Elucidation, **517**
  - **2.1.1** Structure Elucidation of Natural Products, **517**
  - **2.1.2** Structure Determination of Bioactive Peptides, **518**
    - **2.1.2.1** NMR Structure of Ziconotide: A Novel Treatment for Pain, **518**
    - **2.1.2.2** Endothelin as a Lead in Ligand-Based Design, **523**
  - **2.1.3** Instrumental Advances and their Impact on Structure Elucidation, **524**
- **2.2** Conformational Analysis, **525**
- **2.3** Charge State, **526**
- **2.4** Tautomeric Equilibria, **526**
- **2.5** Ligand Dynamics: Line-Shape and Relaxation Data, **528**
- **2.6** Pharmocophore Modeling: Conformations of a Set of Ligands, **531**
- **2.7** Limitations of Analog-Based Design, **532**
- **2.8** Conformation of Bound Ligands: Transferred **NOEs, 532**

3 Receptor-Based Design, **532**
- **3.1** Macromolecular Structure Determination, **533**
  - **3.1.1** Overview of Approach, **533**

Burger's Medicinal Chemistry and Drug Discovery
Sixth Edition, Volume **1:** Drug Discovery
Edited by Donald **J.** Abraham
ISBN 0-471-27090-3 © **2003** John Wiley & Sons, Inc.

3.1.2 Sample Requirements and Assignment Protocols, 534
3.1.3 Recent Developments, 534
3.1.4 Dynamics, 534
3.1.5 Nucleic Acid Structures, 535
3.1.6 Challenges for the Future: Membrane-Bound Proteins, 535
3.2 Macromolecule-Ligand Interactions, 535
3.2.1 Overview, 535
3.2.2 Influence of Kinetics and NMR Timescales, 536
3.2.2.1 Slow Exchange, 538
3.2.2.2 Fast Exchange, 539
3.2.2.3 Intermediate Exchange, 540
3.2.3 NMR Techniques, 543
3.2.3.1 Chemical-Shift Mapping, 543
3.2.3.2 NMR Titrations, 545
3.2.3.3 Isotope Editing and Filtering, 545
3.2.3.4 NOE Docking, 545
3.2.4 Selected Examples, 547
3.2.4.1 DNA-Binding Drugs, 547
3.2.4.2 Immunophilins: Studies of FK506 Analog Binding to FKBP, 552
3.2.4.3 Matrix Metalloproteinases, 555
3.2.4.4 Dihydrofolate Reductase, 557
3.2.4.5 H N Protease, 559
4 NMR Screening, 562
4.1 Methods, 562
4.1.1 Chemical-Shift Perturbation, 562
4.1.2 Magnetization Transfer Experiments, 568
4.1.3 Molecular Diffusion, 570
4.1.4 Relaxation, 571
4.1.5 NOE, 571
4.1.6 Spin Labels, 573
4.2 Practical Considerations, 574
4.2.1 Screening Approach, 574
4.2.2 Library Design, 574
4.2.2.1 Ligand Properties, 574
4.2.2.2 Mixture Design, 576
4.2.3 Hardware and Automation, 576
5 Conclusions, 577
6 Acknowledgments, 577

1 INTRODUCTION

NMR spectroscopy has been widely used as a front-line tool in the pharmaceutical industry for several decades. In the past, the main use of NMR was in the structural characterization of organic molecules synthesized in the course of medicinal chemistry programs. Indeed, medicinal chemists have long regarded NMR as the premier tool to be used in the structure characterization process, to confirm the identity of intermediates or to determine the conformation of lead molecules. Over the last decade major developments in both instrumentation and methods have resulted in this traditional use of NMR in the pharmaceutical industry being augmented by a range of exciting new applications. Two of the most important of these are the use of NMR in structure-based drug design and in screening for drug discovery. Both applications differ from the traditional use of NMR in that now the macromolecular binding partner of the medicinal compound is included in the mixture to be analyzed; that is, contemporary applications of NMR in drug discovery are predominantly focused on the interaction between drug molecules and their macromolecular targets.

The aim of this chapter is to describe how NMR spectroscopy is used in modern drug discovery. The term discovery is used generically throughout to include processes that involve rational drug design as well as those that involve discovery through NMR screening. The latter is a relatively recent development and refers to the use of NMR as a tool to screen a compound library, to identify a molecule or molecules that bind to a chosen macromolecular target. Of course, the distinction between "design" and "discovery" is often quite blurred. This is nowhere more evident than in the recently developed SAR-by-NMR approach (1), in which the discovery of several weakly bound ligands from a screening program is intimately linked to a design process to chemically join them. SAR-by-NMR represents an exciting new technique for lead generation and is described in more detail later in this chapter.

Drug design/discovery represents only the first stage in the whole drug development process. As is clear from the other chapters in this volume, there are many other steps that need to be made once a lead molecule has been designed or discovered. Although other stages of the process, including lead optimization, tox-

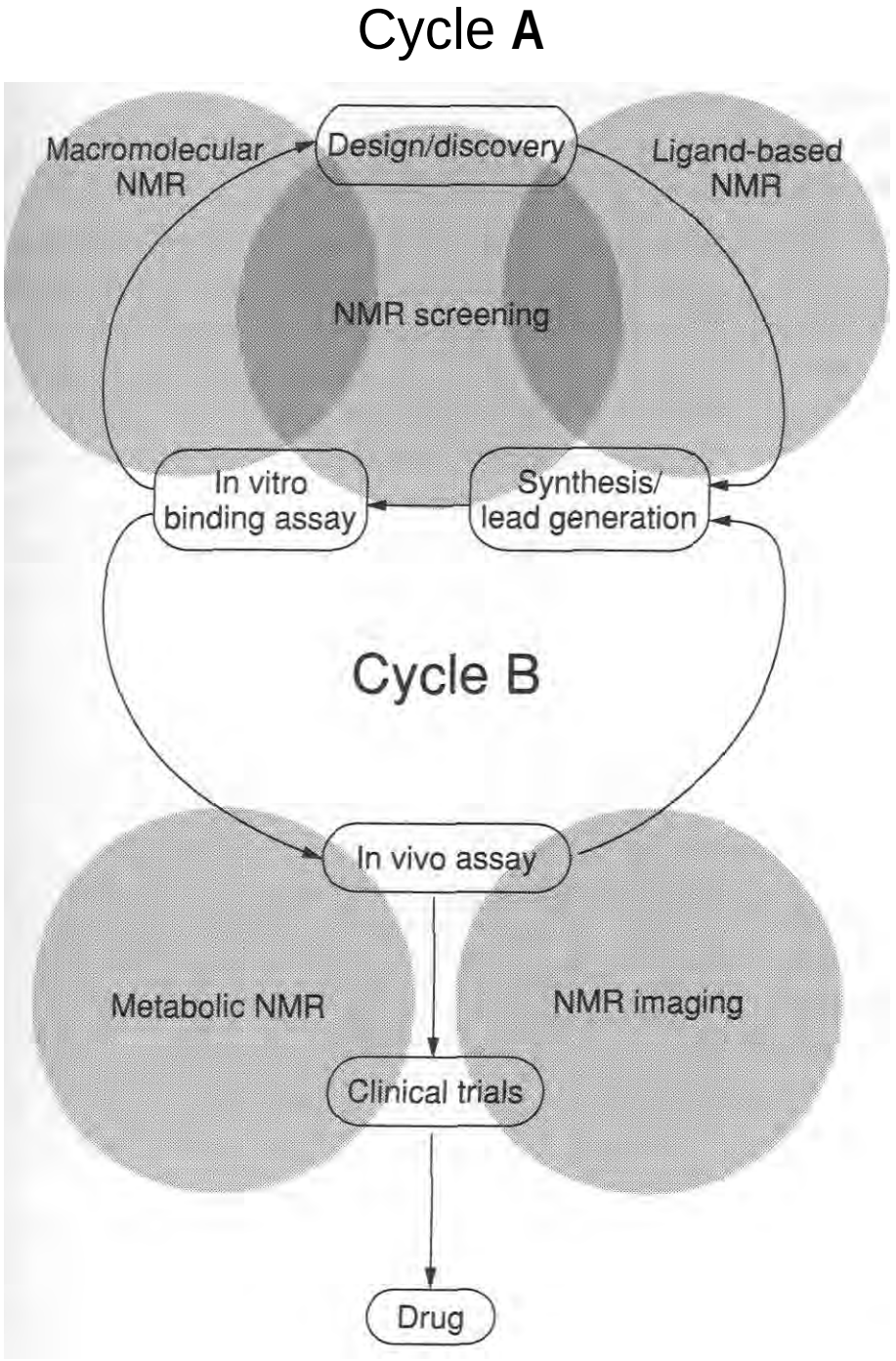


Figure 12.1. Overview of the **drug** development process and summary of various types of NMR experiments that contribute at different stages.

icity studies, preclinical investigations, and clinical monitoring, do not fall within the scope of this chapter, it is worth mentioning that NMR spectroscopy contributes significsntly across the whole spectrum of drug development, right through into the clinical domain. For example, NMR spectroscopy has been applied for the detection of drug metabolites in biological fluids and magnetic resonance imaging, which is based on the fundamental principles of NMR, plays an important role in clinical investigations. It is increasingly being used to monitor the functional outcomes of drug therapy. We briefly address these broader applications of NMR before returning to the main topic of NMR in drug discovery.

1.1 Overview of Drug Development

To give an overview of the breadth of applications of NMR, Fig. 12.1 summarizes the drug development process and indicates the role of NMR at various stages. Drug development is an iterative process and can be simplified by representing it with two interconnected cycles of activity. Cycle A involves the design or discovery of an initial lead followed by its synthesis and bioassay. Based on the initial assay results there may be several loops around this cycle before commencing the *in vivo* studies represented in Cycle B. At this stage consideration of bioavailability, metabolism, and pharmacokinetic profiles must be made and this may involve synthetic modifications of the

lead molecules to improve their druglike properties. Again, several loops around Cycle B may be necessary before one or more development candidates are identified. Ultimately one or two of these development candidates are identified for progression through clinical trials.

As indicated in Fig. 12.1, it is convenient to envisage five broad categories of NMR experiments that may contribute to this overall drug development process.

1. Small molecule, or ligand-based, NMR. This involves studies of drugs and drug leads, typically organic molecules with a molecular weight <500 Da, but also including small proteins of up to a few kDa. These studies may be used to characterize natural products or synthetic drug leads, or to determine their conformation.
2. Macromolecular NMR. This involves studies of the macromolecular targets of drugs, typically to determine their three-dimensional structure and/or the nature of their complexes with ligands.
3. NMR screening. This involves the use of NMR to identify lead molecules that bind to a macromolecular target. These studies typically involve both small molecules and macromolecules and seek to detect the presence of binding interactions between them.
4. Metabolic *NMR*. This involves studies of endogenous molecules whose levels may be modified by drug treatment, or studies of the metabolites of drugs themselves.
5. NMR imaging. Such studies provide anatomical information in an animal model or human patient. This includes, for example, monitoring the size of plaques or tumors in the brains of Alzheimer's or cancer patients, respectively, during drug therapy.

It is clear from these descriptions that NMR covers a wide range of applications in the pharmaceutical industry, although for the remainder of this chapter we will focus on NMR in the drug design/discovery phase of drug development, that is, on categories 1–3 of the preceding list. Together, the studies in categories 1 and 2 may be classified as structure-based design, whereas category 3 relates to drug discovery.

1.2 Scope of Chapter

Our aim is to give a broad overview on the use of NMR as a tool in structure-based design and in screening approaches to drug discovery. The chapter also contains a description of the relevant NMR methods, which are highlighted by illustrative examples. We briefly describe the instrumentation required for such studies and emerging trends in the field are discussed. This includes developments in the field of drug discovery in the postgenomic era that are likely to have an impact on the way in which NMR is used, as seen for example by the recent interest in structural genomics programs. NMR instrument developments are also described. For example, recent advances in cryoprobe technology promise to dramatically increase the sensitivity of NMR spectroscopy and increase its application across the pharmaceutical industry. Finally, a section outlining some of the practical considerations in structure-based design and screening is included. Future directions for the field are mentioned throughout the discussion.

There have been a number of reviews that describe applications of NMR in drug discovery or screening and the reader is referred to these for additional information (2–15). Recent books covering aspects of NMR in drug design are also available (16, 17).

It is assumed that most readers will be familiar with the basic principles of NMR. However, for completeness and to define some of the terms that will be used in this chapter it is useful to give a brief overview of the principles. Excellent texts are available to provide more detail (18, 19).

1.3 Principles of NMR Spectroscopy

The underlying basis of NMR is that when nuclei with a nonzero spin quantum number are placed in a magnetic field they take up one of a discrete number of quantized states. The application of radiofrequency (rf) energy produces transitions between these states. The energy changes associated with these transitions are detected as small voltages induced in

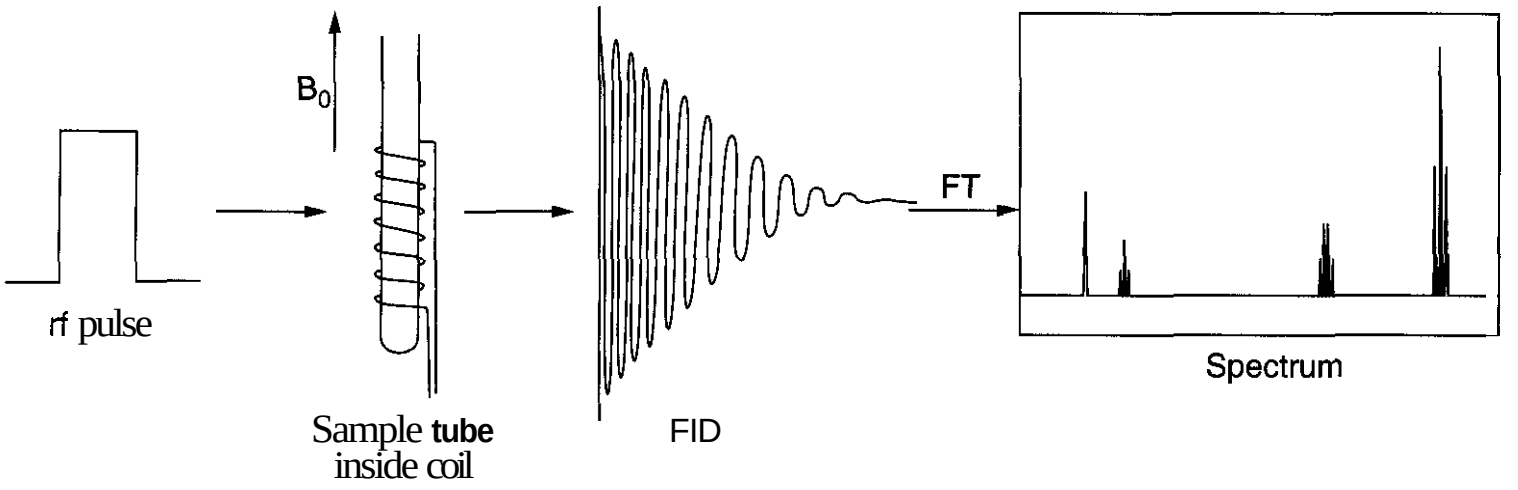


Figure 12.2. Overview of the principles of NMR spectroscopy. Polarization of nuclear spins by a magnetic field is perturbed by application of a radiofrequency (rf) pulse. The resultant signal is Fourier transformed, to yield a spectrum reflecting the number and environments of nuclei in the sample.

a receiver coil that are subsequently amplified, digitized, and processed to yield spectra, as illustrated in Fig. 12.2. The most commonly studied NMR-active nucleus is the proton, ^{1}H, but in modern NMR experiments ^{2}H, ^{13}C, and ^{15}N nuclei are also very important. For these heteronuclei it is common to isotopically enrich the sample because of their low natural abundance. This is particularly important for studies of proteins, as will become apparent later in this chapter. Occasionally, other nuclei find specialist applications. For example, in fluorine-containing drugs it is possible to use sensitive ^{19}F-NMR signals to monitor interaction with target proteins, as described later in this chapter.

In modern spectrometers the rf energy is supplied in the form of short pulses (typically, ~10 μs) that simultaneously excite all nuclei of a given isotope type (e.g., all protons or all ^{13}C nuclei). Nuclei of a given isotope that are in different chemical environments by virtue of their atomic locations in the molecule have slightly different resonance frequencies and lead to different oscillating voltages in the receiver coil. The resultant combined signal, termed a ***free induction decay*** (FID), is Fourier transformed to give a spectrum that is basically a plot of peak intensity vs. frequency, with one peak for each chemically distinct nucleus. These features are schematically illustrated in Fig. 12.2. The frequency axis is termed the ***chemical shift*** because it reflects the local chemical environment of each nucleus. The range of chemical environments of nuclei in a molecule is such that chemical shifts range up to only a few hundred parts per million (ppm) of the base resonance frequency for ^{13}C and ^{15}N. For ^{1}H the range is smaller still, covering only about 10 ppm. Despite this small range, chemical shifts provide valuable diagnostic information on the environment of the nucleus giving rise to the signal.

The chemical shift is an extremely important NMR parameter but there are many other parameters that can be discerned from NMR spectra. Indeed, NMR is unique among many forms of spectroscopy in that there are so many parameters associated with a spectrum other than just peak intensity and frequency. These include coupling constants, which provide information on local conformations and also on molecular connectivities; nuclear Overhauser effects (NOEs), which provide information on internuclear distances: and relaxation parameters, which provide information on molecular dynamics. Table **12.1** summarizes the main NMR parameters that may be measured and highlights their applications in the drug discovery process.

The following sections of this chapter provide specific examples of how these various parameters are useful in the drug discovery process. Before doing this, though, it is useful to consider some of the limitations of one-dimensional (1D) NMR spectroscopy, particularly when the detected nucleus is ^{1}H, as is most commonly the case. With one signal coming from each chemically distinct proton and with those signals spread only over 10 ppm, it is

Table 12.1 NMR Parameters and Their Applications in Drug Design/Discovery

Parameter	Information Provided Relevant to Drug Design
Chemical shift	Reflects local chemical environment; provides a fingerprint marker of structure (particularly in HSQC spectra)
Coupling constants	Conformational analysis, establishing molecular connectivity
Nuclear Overhauser effect	Determining interproton distances, three-dimensional structures
Relaxation times	Molecular dynamics
Line-shape	Detecting and quantifying chemical exchange processes
Peak intensities	Reflect relative number of nuclei, molecular symmetry
Amide exchange rates / temperature coefficients	Hydrogen bonding or solvent exposure of amide protons

clear that spectral overlap can potentially be a major problem for anything but the simplest of molecules. The development of higher field NMR spectrometers, which effectively provide greater dispersion in the frequency dimension, has contributed significantly to overcoming this limitation and increasing the application of NMR for studying pharmaceutically relevant molecules. In addition to such instrumental developments, methodological advances have also played a key role in extending the use of NMR. Multidimensional NMR methods have revolutionized biomolecular NMR spectroscopy by removing the limitations of a single frequency dimension, leading to the development of 2D, 3D, and 4D spectra.

A simple way of illustrating multidimensional NMR is through reference to heteronuclear correlation spectroscopy, in which two or more separate frequency dimensions are correlated with one another. For example, a particularly valuable 2D experiment is ^{1}H-^{15}N heteronuclear single quantum correlation (HSQC) spectroscopy, in which the resultant spectrum has two frequency axes, corresponding to ^{1}H and ^{15}N frequency dimensions, and one intensity axis. Analogous ^{1}H-^{13}C HSQC spectra are also widely used. Such spectra are normally represented with the intensity axis in contour form so that they may be drawn in two dimensions as a set of contour peaks. Spectral peaks occur for pairs of $^{15}N/^{1}H$ or $^{13}C/^{1}H$ nuclei that are directly bonded to one another, and with each frequency being characteristic for the local chemical environment they represent a relatively simple, but highly characteristic fingerprint of the sample. Figure 12.3 shows the relationship between 1D and 2D spectra for the immunosuppressive drug cyclosporin, and includes a region of both the $^{1}H/^{15}N$ and $^{1}H/^{13}C$ HSQC spectra. In HSQC spectra-overlap problems are alleviated because, even if two protons have the same chemical shift and would hence be overlapped in a 1D spectrum, chances are that the respective heteronuclear signals will not be overlapped, allowing the signals to be resolved in the 2D spectrum. HSQC spectra are widely used in NMR-based drug screening and we will return to them later.

Multidimensional NMR spectra are not restricted to cases where the separate frequency axes encode signals from different nuclear types. Indeed, much of the early work on the development of 2D NMR was performed on cases where both axes involved ^{1}H chemical shifts. The main value in such spectra comes from the information content in cross peaks between pairs of protons. In COSY-type spectra (COSY = Correlation SpectroscopY) cross peaks occur only between protons that are scalar coupled (i.e., within 2 or 3 bonds) to each other, whereas in NOESY (NOE Spectroscopy) spectra cross peaks occur for protons that are physically close in space (<5 Å apart). A combination of these two types of 2D spectra may be used to assign the NMR signals of small proteins and provides sufficient information on internuclear distances to calculate three-dimensional structures. Figure 12.3 includes a panel showing the COSY spectrum of cyclosporin and highlights the relationships between 1D ^{1}H-NMR spectra and corresponding 2D homonuclear (COSY) and heteronuclear (HSQC) spectra.

Homonuclear 2D spectra are generally applicable for the study of proteins up to only approximately 80 amino acids in size. For

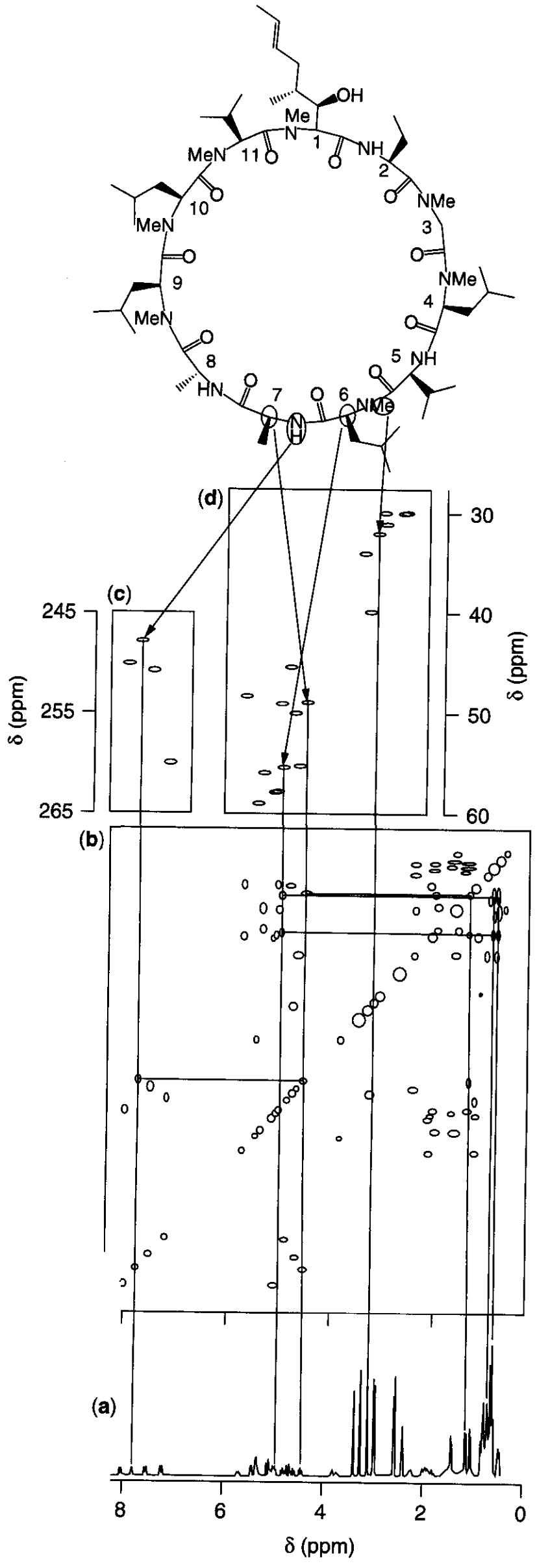


Figure 12.3. A schematic representation of the (a) 1D ^{1}H; (b) 2D DQF-COSY; (c) $^{15}N/^{1}H$-HSQC; and (d) $^{13}C/^{1}H$-HSQC spectra of the immunosuppressive agent cyclosporin. Example resonances/correlations from residues 6 and 7 have been highlighted to illustrate the assignment process.

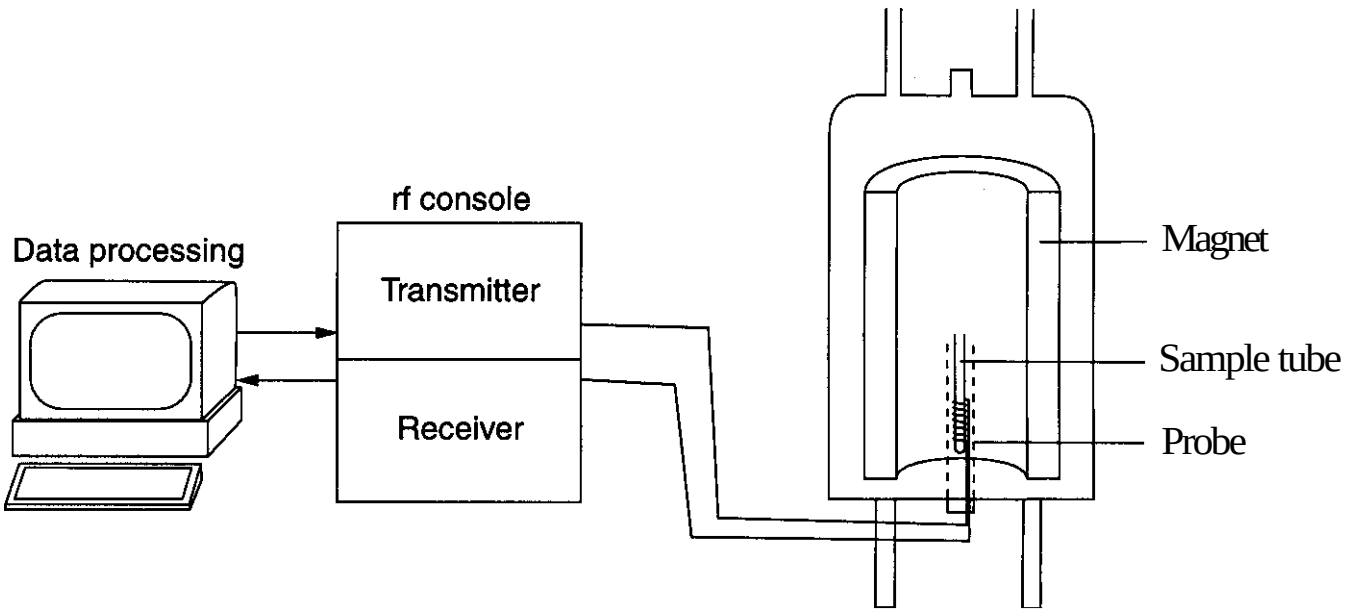


Figure 12.4. Block diagram of a modern NMR spectrometer. These systems use superconducting magnets that are based on a solenoid of a suitable alloy (e.g., niobium/titanium or niobium/tin) immersed in a dewar of liquid helium. The extremely low temperature of the magnet itself (4.2 K) is well insulated from the sample chamber in the center of the magnet bore. The probe in which the sample is housed usually incorporates accurate temperature control over the range typically of 4 to 40°C for biological samples. The rf coil in the probe is connected in turn to a preamplifier, receiver circuitry, analog-to-digital converter (ADC), and a computer for data collection.

larger proteins the increased number of signals leads to overlap problems and, in addition, COSY-type spectra suffer from poor sensitivity when the signal linewidths are of the same order as or larger than ^{1}H, ^{1}H scalar coupling constants. Such limitations are reduced by use of spectra of higher dimensionality (i.e., 3D or 4D spectra) that are based on correlations involving heteronuclear rather than homonuclear coupling constants. Such spectra are important in the structure determination process for larger proteins and are typically recorded for samples that incorporate uniform labeling with ^{15}N, or both ^{13}C and ^{15}N nuclei. Multidimensional spectra that involve irradiation of ^{1}H, ^{13}C, and ^{15}N nuclei are referred to as ***triple resonance spectra.***

The details of how multidimensional spectra are obtained is beyond the scope of this chapter, but it suffices to say that, like most other modern NMR experiments, they involve irradiation of the sample with a set of rf pulses of defined length, frequency, and phase, with specific interpulse delays. The pulse programs for such experiments are commonly provided with the spectrometer as part of a standard library of experiments and may easily be run by novice users after input of an appropriate set of parameters to define the relevant spectral widths and type of experiment required.

The above discussion provides a basic overview of some of the methods important in modern NMR spectroscopy. Before examining specific applications in drug discovery it is useful to describe the instrumental requirements for such studies.

1.4 Instrumentation

NMR spectrometers constitute a powerful and homogeneous magnet, a radiofrequency console for generating appropriate rf pulses, a probe for applying this rf energy to the sample and receiving the resultant signals, and a computer console for controlling the experiments and acquiring the resultant data. These features are summarized in Fig. 12.4. Spectrometers are normally specified in terms of the resonant frequency of protons at the given magnetic field (e.g., 500 MHz corresponds to a magnetic field of 11.7 Tesla). Both sensitivity and dispersion of signals increase with increasing magnetic field.

There have been some major breakthroughs in both NMR instrumentation and methodology over the last decade that have greatly increased the utility of NMR for drug discovery applications. These are summarized in Table 12.2, which also includes some of the earlier milestones in the development of NMR. Most notable among recent innovations

Table 12.2 Milestones in the Development of NMR Spectroscopy

Year	Development	Nature
1970	FT NMR	Instrumental
1975	Superconducting magnets	Instrumental
1980	2D NMR	Methodological
1985	Protein structure determination	Methodological
1990	Isotope labeling/multidimensional NMR	Methodological
1990	Pulsed field gradients	Instrumental/methodological
1995	NMR screening	Methodological
1997	TROSY	Methodological
1998	LC-NMR/LCMS-NMR	Instrumental
2000	Cryoprobes	Instrumental

are the use of pulsed-field gradient methods for improving spectral quality and allowing new types of experiments to be performed, transverse relaxation-optimized spectroscopy (TROSY) methods (20) for increasing the size of macromolecules that can be examined, and cryoprobes for enhancing sensitivity. The development of cryoprobes has resulted in the biggest single gain in sensitivity over recent years, effectively giving 500-MHz spectrometers the sensitivity of 800-MHz spectrometers (although without the gain in resolution!). The enhanced sensitivity is obtained by cooling the receiver coil and associated circuitry to near liquid helium temperatures, thereby reducing thermal noise. There were considerable technical barriers to be overcome in developing such probes because of the large difference in temperature between the receiver coils and the sample, which are only a few millimeters apart. These barriers have now been overcome and cryoprobes are being installed in a large number of laboratories. They are also becoming available for higher field systems (800 MHz), thus providing further sensitivity gains.

Although the basic configurations of instruments tailored for structure-based design or for NMR drug screening are similar, there are some minor differences. For structure-based design applications a relatively high field spectrometer is required (>500 MHz), usually equipped with three or four radiofrequency channels for the simultaneous irradiation of ^{1}H, ^{13}C, ^{15}N, and in some cases ^{2}H nuclei. The greatest sensitivity and dispersion are obtained with the highest possible magnetic field. Instruments of up to 900 MHz are currently available but, at the time of writing, only a few have been installed. Numerous 800-MHz systems dedicated to structure-based design have been installed in pharmaceutical laboratories. The high field instruments provide another advantage in that TROSY experiments (20) can be used to produce a marked improvement in spectral quality for larger proteins. Such developments promise to push higher the size of proteins whose structure can be determined by NMR.

For NMR drug screening programs, the basic requirement of a spectrometer of 500 MHz or greater remains, but in addition, an interface that allows the spectrometer to sample a library of compounds of potential binding ligands needs to be present. This may be done either by use of a discrete sample changer or a flow-type system. Flow systems have the potential advantage of increased throughput but have the potential disadvantage of precipitation of protein samples. In practice this appears not to have been a major problem and both types of systems are in use in the pharmaceutical industry. Sample changer systems currently have the advantage that they may be adapted for use with cryoprobe technology (currently unavailable for flow systems). Cryoprobes allow dramatically enhanced sensitivity gains, which bring particular advantages to the study of macromolecule-ligand interactions used in screening programs (21).

Pulsed-field gradients have become integral to most modern NMR spectrometers and are routinely used both for structure determination and screening experiments. Another recent development has been the interface of NMR spectrometers with other instrumenta-

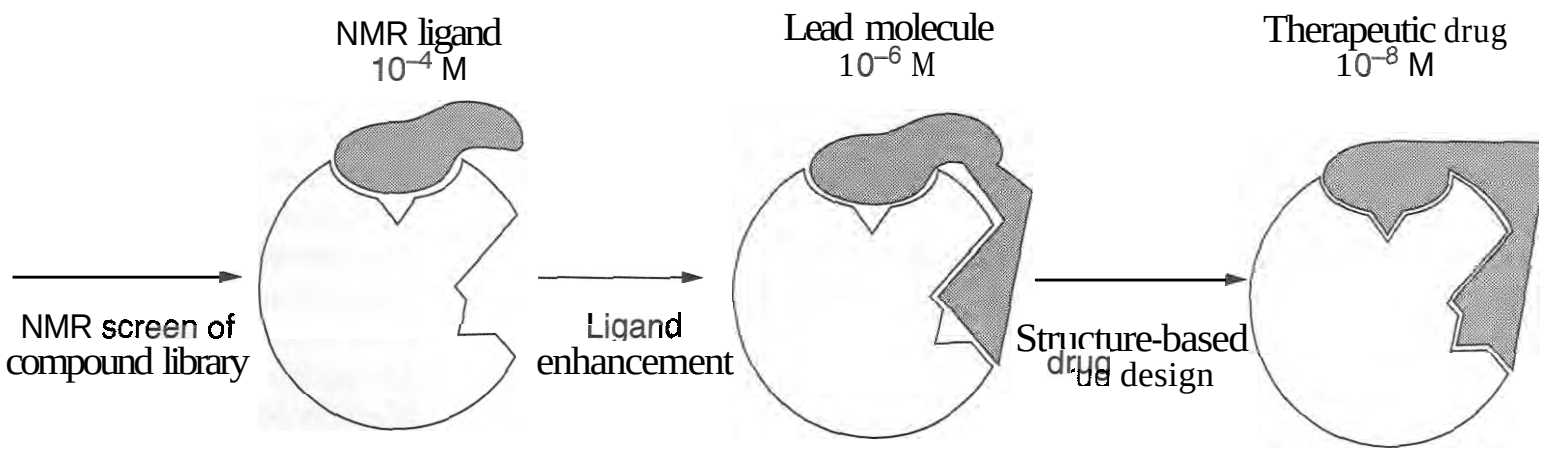


Figure 12.5. **A** summary of the relationship between NMR screening and structure-based design. **(Adapted** from Ref. 15.)

tion such as liquid chromatography (LC) and/or mass spectrometry (MS). The applications of these instrumental developments to drug discovery have been recently reviewed (8, 13, 22).

1.5 Applications of NMR in Drug Design and Discovery

Our focus here is on the use of NMR in the discovery and design phase of drug development. The major role of NMR in the ***design*** process comes about by its exquisite ability to provide structural information, whereas the major role of NMR in ***discovery*** comes through its use as a screening tool to detect the binding of novel ligands to macromolecular targets. As already noted, the latter application is a relatively recent development but has created much interest in the pharmaceutical industry and promises to significantly enhance applications of NMR in this industry. The impact of the methodology is already becoming evident even at this early stage, with several SAR-by-NMR-derived leads currently in clinical development. As already noted, though, the discovery and design phases are often intimately connected, with lead molecules discovered in screening programs routinely being optimized by use of structure-based design approaches (Fig. **12.5).**

In the context of this chapter structure-based design refers to the process of determining the three-dimensional structure of **a** lead molecule or macromolecular target, or determining the structure of the macromolecule-ligand complex, and using this information to design new drugs. The questions that may be asked when embarking on structure-based design projects are:

- What are the solution and bound conformations of the ligand?
- What is its charge/tautomeric state?
- Which functional groups bind to the receptor and what charge state are they in?
- What is the structure of the receptor?
- Which parts interact with the ligand?
- What is the geometry of the ligand-receptor complex?
- What are the kinetics of binding and are there dynamic motions of ligand, receptor, or the complex?

Table **12.3** summarizes these and other questions and indicates the type of NMR approaches that can provide answers. Remaining sections of this chapter are organized around the headings identified in Table **12.3.**

In considering these questions it is convenient to distinguish between ligand-based design, where the structural focus is on the small lead molecule, and receptor-based design, where the aim is to determine the structure of the macromolecular target. The NMR methods used in ligand-based design have been well established for many years, based on the use of NMR by organic and natural product chemists for more than four decades. However, there have been some important recent advances in NMR methods such as the use of pulsed field gradients, and in the combination of NMR

Table 12.3 Information on Ligands, Macromolecules, and Their Complexes Sought in Structure-Based Design and Relevant NMR Technologies Used to Derive This Information

Target	Information	NMR Technology
Ligand	Solution conformation	1D/2D NMR
	Chargeltautomeric state	Chemical shift/titrations
	Solution dynamics	Line-shape/relaxation analysis
	Pharmacophore models	All of the above, and TrNOE, of multiple ligands
	Bound ligand conformation	TrNOE
Macromolecule	3D structure	2D/3D/4D NMR
	Macromolecular dynamics	Relaxation time measurements
	Structure of articulated macromolecules (e.g., multimeric or membrane-bound receptors)	TROSY
Ligand–macromolecular complex	Stoichiometry of complex	Chemical shift/titration
	Kinetics of binding	Line width, titration analysis
	Location of interacting sites	HSQC, isotope editing
	Orientation of bound ligand	NOE docking
	Bound ligand conformation	TrNOE
	Structure of complex	3D/4D NMR
	Dynamics of complex	Relaxation time measurements

with other technologies such as LC and MS that promise to enhance applications in this field **(13).** The use of NMR to determine the three-dimensional structures of macromolecules is a newer field, commencing only in around 1985, and is one that is still rapidly evolving. NMR screening is a still newer approach, developed since around 1996. Ligand-based and receptor-based design are examined in Sections **2** and **3,** respectively, and screening-based approaches are examined in Section 4.

2 LIGAND-BASED DESIGN

Many naturally occurring molecules have potent bioactivity that renders them useful leads in the drug design process. These may be naturally occurring hormones, neurotransmitters, or other endogenous molecules, or they may be bioactive molecules from plants or microorganisms. Furthermore, screening programs on synthetic compound libraries frequently result in the discovery of bioactive molecules that then become starting points in drug design. The general aim of ligand- or analog-based design is to determine the structure and conformation of a known bioactive molecule and then mimic this conformation in a designed lead compound, with the aim of improving the activity or druglike properties. The following sections examine various aspects of ligand-based design and illustrate them with examples.

2.1 Structure Elucidation

If the bioactive molecule is a synthetic product, its structure may be rapidly deduced by a simple comparison of NMR parameters (often combined with MS) of the product relative to those of the known precursor, to see whether the desired chemical transformation has taken place. If the bioactive compound is an unknown molecule discovered in an active fraction in bioassay-guided screening, then the first step is to elucidate its structure. Typical molecules that form the basis of such natural products-based drug discovery studies include "organic" natural products as well as small peptides and proteins. The approaches to structure elucidation for natural products and peptides/proteins are a little different from each other and are described in turn.

2.1.1 Structure Elucidation of Natural Products. In the case of nonpeptidic natural products the main structural focus initially is to elucidate the carbon framework. This nor-

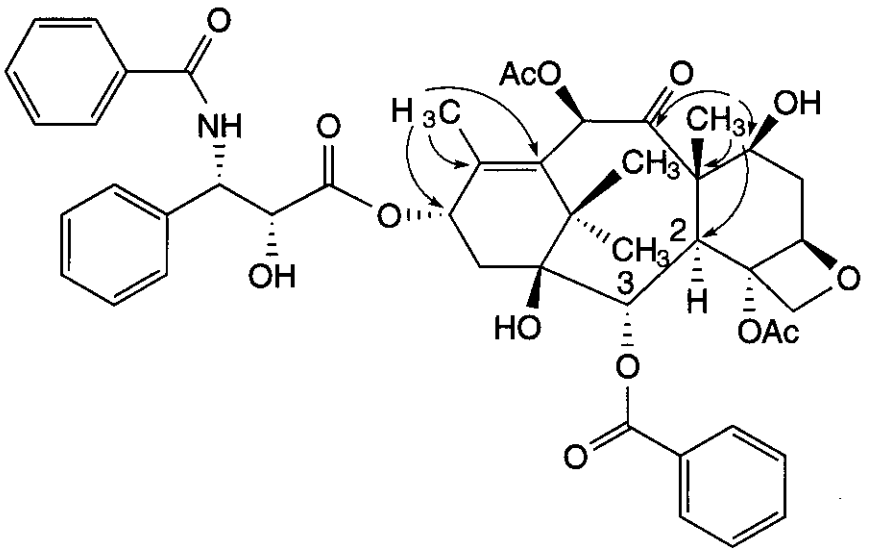


Figure 12.6. Illustration of the HMBC correlations (arrows) used to assign the positions of two of the methyl quaternary methyl groups in taxol.

mally involves a combination of 1D ^{1}H and ^{13}C-NMR, followed by homonuclear (DQF-COSY, TOCSY, ROESY, or NOESY) and heteronuclear (HSQC, HMBC) 2D experiments. Heteronuclear multiple bond correlation (HMBC) spectra are particularly valuable because they assist in tracing the backbone of the molecule. Such spectra display cross peaks between a ^{13}C nucleus and protons connected within two or three bonds and, in doing so, provide valuable information on molecular connectivity. Figure 12.6 shows typical HMBC correlations seen for selected regions of taxol, a plant-derived natural product that is currently a leading treatment for breast and ovarian cancers. Although the structure of taxol itself was originally deduced from a combination of X-ray crystallography on a degradation product and a range of ^{1}H and ^{13}C spectra in the 1970s, before HMBC spectra had been invented, HMBC spectra have been widely used for studies of the many taxol derivatives that have been examined in the last decade.

Elucidation of the carbon framework of natural products often yields substantial information about the three-dimensional structure at the same time. but if there are remaining questions on the stereochemistry of chiral centers or other factors affecting the three-dimensional structure, these can usually be resolved from NOESY svectra and/or an analysis of coupling constants. We will return to the taxol example later in Section 2.2 when describing conformational analysis.

2.1.2 Structure Determination of Bioactive Peptides. In contrast to the process described for organic molecules, the structure elucidation of peptide-based natural products involves two distinct steps: (*1*) the elucidation of the primary structure (amino acid sequence) followed by (2) a determination of secondary/tertiary structure. The primary structure determination is routinely done through Edman sequencing, or more recently by MS-MS methods. NMR plays a key role in the elucidation of the secondary and tertiary structure of peptides, mainly based on 2D homonuclear NMR spectroscopy. A combination of DQF-COSY and TOCSY (TOtal Correlalation Syectroscopy) spectra are used to assign spin systems to amino acid types and then NOESY spectra are used to sequentially assign the resonances to individual protons in the peptide (23). The three-dimensional structure is then determined by deriving a series of internuclear distance restraints from the NOESY spectrum and using them in a simulated annealing algorithm to calculate a family of structures consistent with them.

Because the structure determination of peptides and proteins represents a very important contribution of NMR to the drug development process, it is informative to describe the yrocess in more detail. To do this we will use the recently developed peptide-based drug MVIIA as an example.

2.1.2.1 NMR Structure of Ziconotide: A Novel Treatment for Pain. MVIIA, now known as Ziconotide, is a 25-amino acid peptide originally discovered from the venom of the marine cone snail, *Conus magus*. Like other ω-conotoxins it is a potent blocker of N-type calcium channels, giving it a wide range of potential therapeutic applications. When delivered intrathecally (i.e., through spinal infu-

sion), it is approximately 1000 times more potent than morphine as an analgesic and has great potential for the treatment of intractable cancer pain (24). Figure 12.7 shows the peptide sequence and illustrates selected regions of the TOCSY and NOESY spectra.

As seen in Fig. 12.7, the TOCSY experiment is useful for classifying spin systems to amino acid type, with typically the most useful region being the "skewers" emanating from individual NH shifts (–7–10 ppm). For each NH proton in the peptide a series of cross peaks to the a, β, and other side-chain protons is observed and these patterns define the spin system as belonging to a particular type of amino acid. Note, however, that there is some degeneracy in the resultant patterns. The NH side-chain pattern is truncated if there is a break of more than three bonds between protons within the spin system. This means, for example, that the skewers for aromatic residues extend only as far as the β-protons and they therefore appear similar to other "AMX" residues such as Cys, Ser, Asp, or Asn. Nevertheless, the ability to assign signals to either individual amino acid types or to the AMX group is a useful starting point in the assignment. However, such spectra provide no information about the sequential location of an amino acid if it is not unique in the sequence. These sequential assignments are obtained from the NOESY spectrum, as illustrated in the sequential walk shown in the middle panel of Fig. 12.7. The aim of the sequential assignment process is to locate adjacent amino acid spin systems, principally through a cross peak between the αH proton of one residue (i) and the NH of the following residue ($i + 1$), often denoted as dαN(i, i + 1). Additional support for the assignment is usually also sought in dβN(i, i + 1) and dNN(i, i + 1) correlations. At the early stages of an assignment it is impossible to be certain whether a particular cross peak is a sequential or longer range cross peak; however, as the assignment procedure progresses, ambiguities become resolved. The assignment process is generally highly convergent, in that once a series of correct assignments is made the number of choices for remaining cross peaks diminishes, in principle making their assignment easier.

Because peptides are polymers of amino acids units, the repeated NH, Ha, and side-chain protons tend to fall in characteristic chemical shift ranges that can be useful in looking for patterns to identify amino acid types. Table 12.4 shows typical chemical shifts for each of the 20 common amino acids when located in a "random-coil" environment (23, 25. 26). It is important to stress that these shifts can vary quite considerably in structured proteins (by up to several ppm) and are more useful for pattern recognition purposes than for exact identification of a particular residue. In the case of the Ha protons, the differences between the actual shifts in a structured protein and these random-coil values have an additional important use, in that they provide an indication of the local secondary structure. Intuitively, the further a chemical shift is from a random-coil value. the more likely it is attributed to that residue's being in a structured environment.

After the assignment is complete it is possible to derive substantial information about the secondary structure from an analysis of chemical shifts, coupling constants, and NOEs, even before the three-dimensional structure calculations are commenced. Figure 12.8 shows a typical summary of the relevant NMR information, again using the data for MVIIA as an example (27, 28). Trends in these, data provide a general indication of major elements of secondary structure. For example, a series of strong dαN(i, i + 1), relative to dNN(i, i + 1) NOEs often indicates an extended or β-type structure, whereas strong dNN(i, i + 1) NOEs indicate local helical structure or turns. Large JαN coupling constants (>8.5 Hz) are associated with extended structure and small ones (<5 Hz) with helical structure. Similarly, deviations of chemical shifts from random-coil values, often represented in terms of "chemical shift indices" (29), indicate extended (positive values) or helical structure (negative values).

An additional useful parameter is the exchange rate of amide protons after dissolution of the sample in D_2O. Slowly exchanging amide protons indicate protection from solvent and possible involvement in intramolecular hydrogen bonds associated with elements of secondary structure. All of the NMR and slow exchange data can be consolidated to give

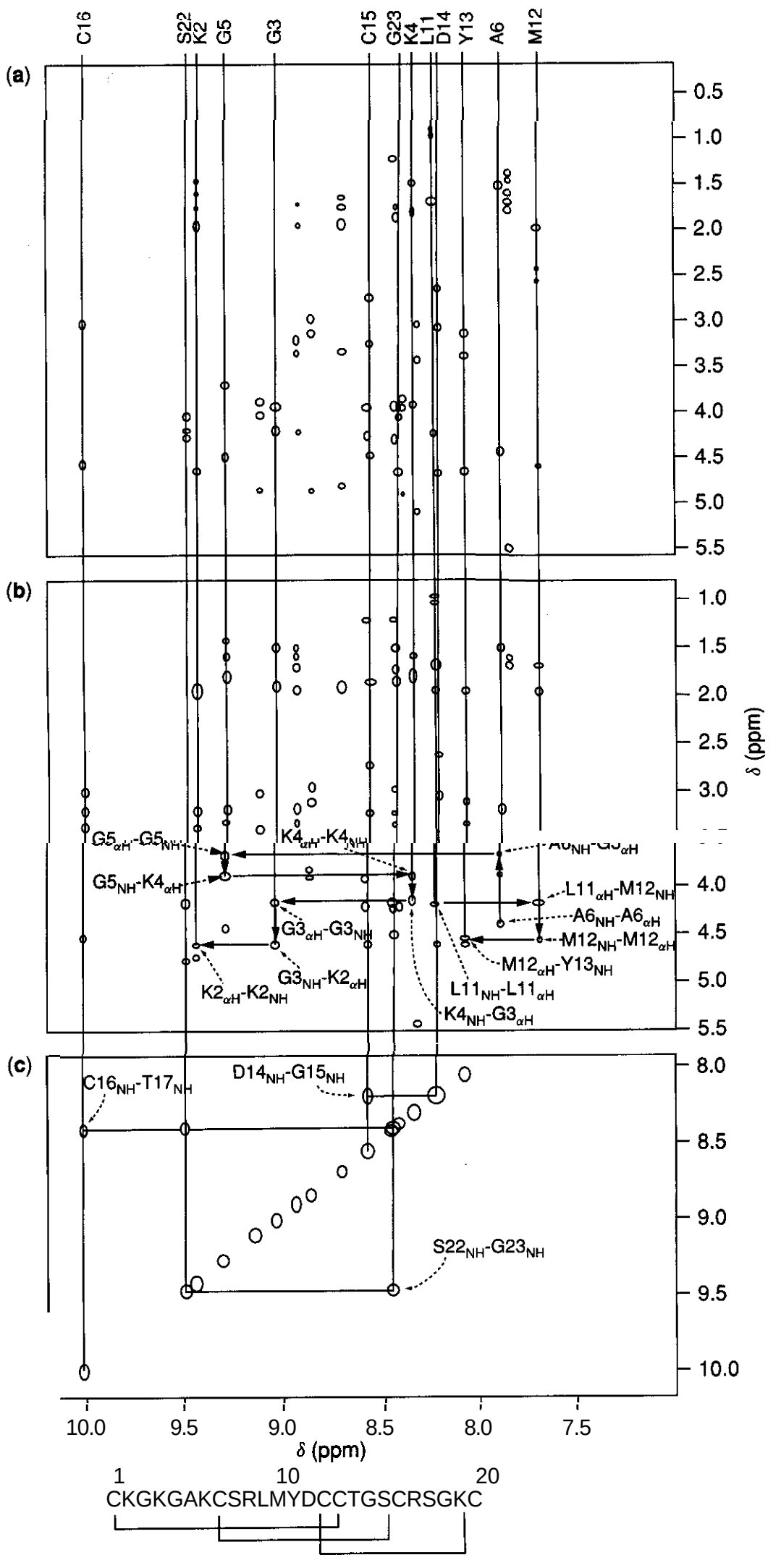


Figure 12.7. Schematic representations of 2D-NMR spectra of the conotoxin MVIIA. (a) The fingerprint region of the TOCSY spectrum with selected spin systems marked. (b) Fingerprint region of the NOESY spectrum showing two (**K2-A6** and **L11-Y13**) sequential walks. (c) NH-NH region of the NOESY spectrum showing correlations between the NH protons of D14 and G15; C16 and T17; and S22 and G23.

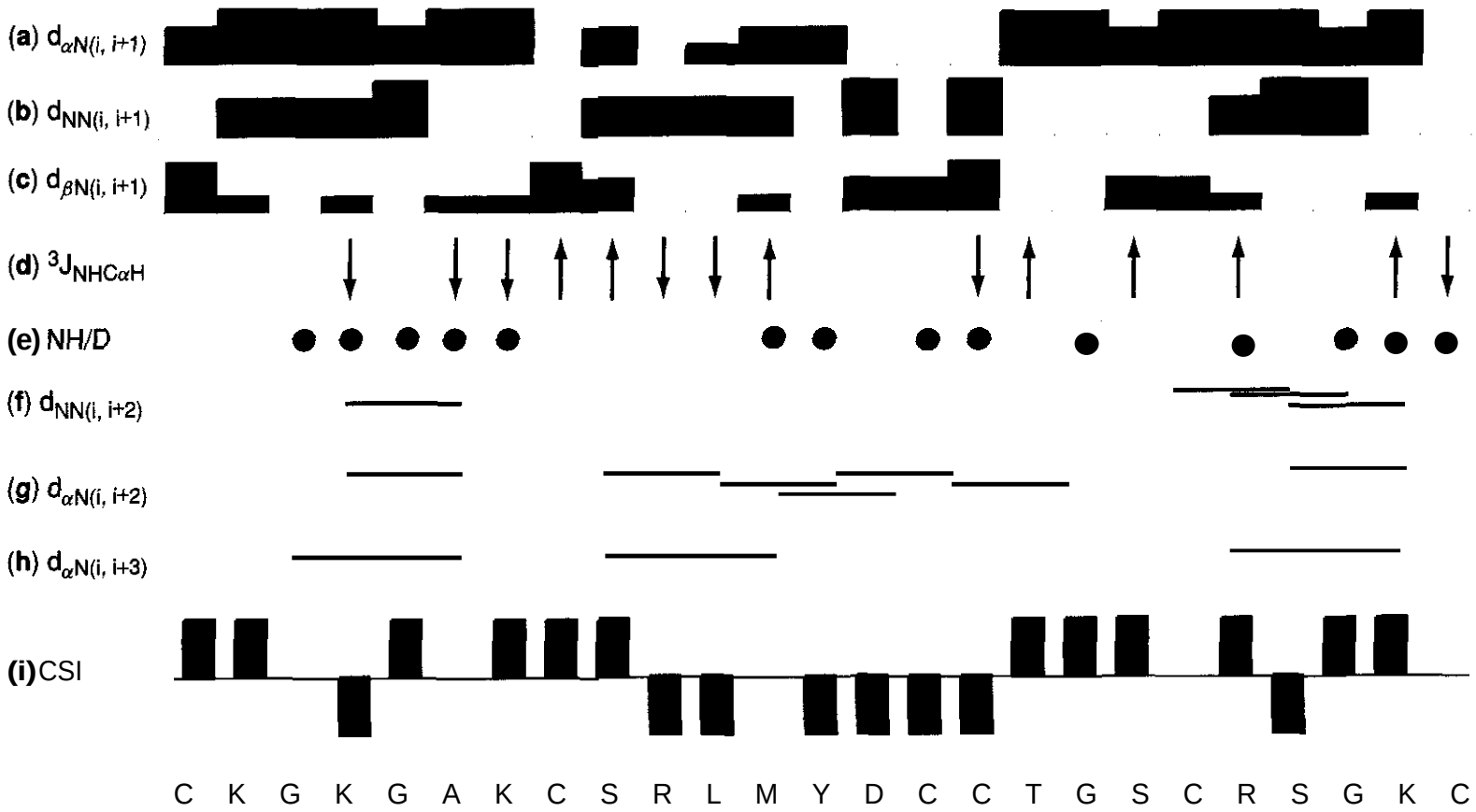
Table 12.4 ^{1}H Chemical Shifts for the 20 Common Amino Acid Residues in Random-Coil Peptides[a]

Residue	NH	αH	βH	Others
Ala	8.24	4.32	1.39	
Arg	8.23	4.34	1.89, 1.79	γCH$_2$ 1.70, 1.70
				δCH$_2$ 3.32, 3.32
				NH 7.17, 6.62
Asn	8.40	4.74	2.83, 2.75	γNH$_2$ 7.59, 6.91
Asp	8.34	4.64	2.84, 2.75	
Cys	8.43	4.71	3.28, 2.96	
Gln	8.32	4.34	2.13, 2.01	γCH$_2$ 2.38, 2.38
				δNH$_2$ 6.87, 7.59
Glu	8.42	4.35	2.09, 1.97	γCH$_2$ 2.31, 2.28
Gly	8.33	3.96		
His	8.42	4.73	3.26, 3.20	2H 8.12
				4H 7.14
Ile	8.00	4.17	1.90	γCH$_2$ 1.48, 1.19
				γCH$_3$ 0.95
				δCH$_3$ 0.89
Leu	8.16	4.34	1.65, 1.65	γH 1.64
				δCH$_3$ 0.94, 0.90
Lys	8.29	4.32	1.85, 1.76	γCH$_2$ 1.45, 1.45
				δCH$_2$ 1.70, 1.70
				εCH$_2$ 3.02, 3.02
				εNH$_3^+$ 7.52
Met	8.28	4.48	2.15, 2.01	γCH$_2$ 2.64, 2.64
				εCH$_3$ 2.13
Phe	8.30	4.62	3.22, 2.99	2,6H 7.30
				3,5H 7.39
				4H7.34
Pro		4.42	2.28, 2.02	γCH$_2$ 2.03, 2.03
				δCH$_2$ 3.68, 3.65
Ser	8.31	4.47	3.88, 3.88	
Thr	8.15	4.35	4.22	γCH3 1.23
Trp	8.25	4.66	3.32, 3.19	2H 7.24
				4H 7.65
				5H 7.17
				6H 7.24
				7H 7.50
				NH 10.22
Tyr	8.12	4.55	3.13, 2.92	2,6H 7.15
				3,5H 6.86
Val	8.03	4.12	2.13	γCH$_3$ 0.97, 0.94

[a]The backbone shifts (αH and NH, ppm) are from Wishart et al. (26). The remaining shifts are from Wiithrich 1986 (23).

an accurate representation of secondary structure, as indicated in the lower panel of Figure 12.8. In the case of MVIIA a triple-stranded β-sheet may be deduced on the basis of the local NOE, coupling, chemical shift, and amide-exchange NMR data.

Once all peaks in the 2D spectra have been assigned, cross peaks in the NOESY spectrum are used to derive a series of interproton distance restraints. These are then used in a simulated annealing algorithm to calculate a family of 3D structures consistent with the input restraints. Fig. 12.9 shows two commonly used methods of representing such NMR-derived structures, either as a stereoview of the superimposed family of structures or as a ribbon diagram, in which elements of secondary structure are highlighted. For the latter rep-



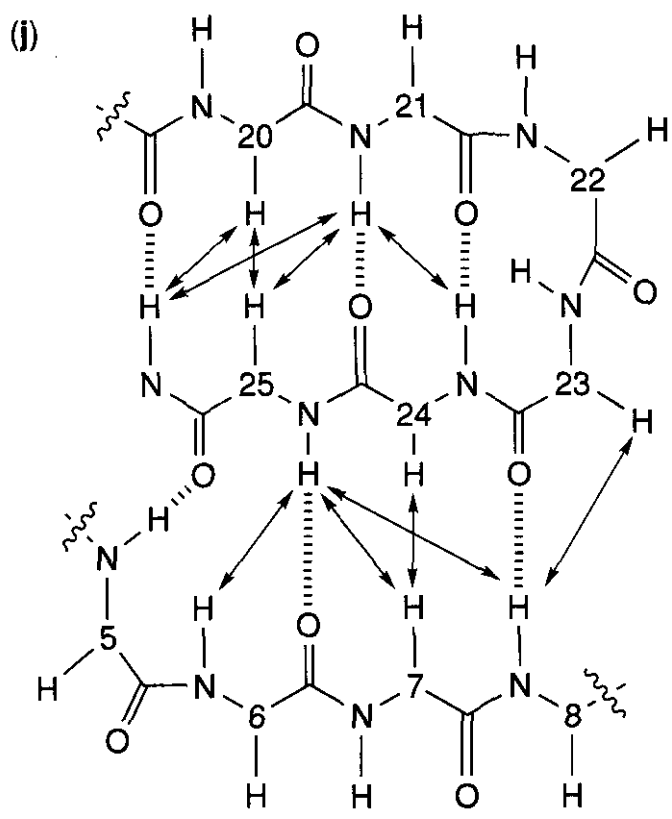


Figure 12.8. A summary of the NMR data observed for MVIIA. (a) Ha-NH sequential NOEs. (b) NH-NH sequential NOEs. (c) Hβ-NH sequential NOEs. (f-h) Other short-range NOEs. The thickness of the bar indicates the strength of the observed NOE (weak, medium, or strong). (d) Three-bond NH-Ha coupling data, where upward-pointing arrows indicate a large coupling (>8 Hz) and downward-pointing arrows indicate a small coupling (<5 Hz). (e) H/D exchange data, where a filled circle represents a slow exchanging NH. (i) Chemical shift index (CSI) data. The CSI uses a scoring system that compares Ha shifts to random-coil chemical shifts. A sequence of consecutive +1 scores is indicative of β-structure, whereas a sequence of consecutive –1 scores suggests helical structure. (j) The β-sheet of MVIIA. Double-headed arrows indicate observed NOEs and broken lines indicate proposed H-bonds.

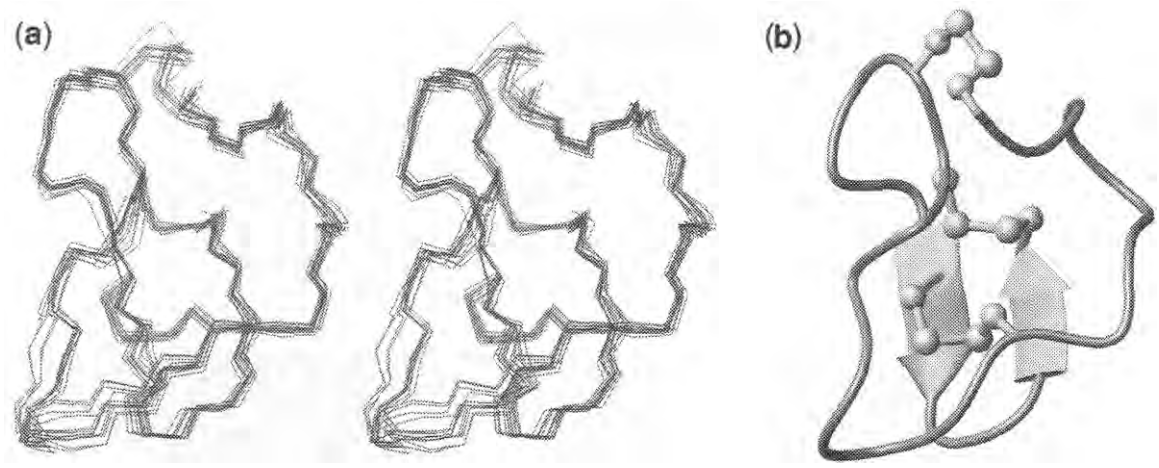


Figure 12.9. (a) A stereoview of the superimposed backbone structures of the 20 lowest energy conformations for MVIIA(27). (b) Ribbon diagram of MVIIA.

resentation the lowest energy or average member of the ensemble is often chosen as representative of the structure. It is important, however, to examine the full ensemble to gain a complete understanding of the structure. Regions of disorder in the ensemble can be indicative of a lack of sufficient distance restraints, perhaps attributable to overlap or assignment errors, or may be related to local flexibility.

In the case of MVIIA the peptide itself is being clinically developed as the active drug for administration through the intrathecal (spinal infusion) route. However, in general, peptides have a range of potential disadvantages as drugs, including poor bioavailability and susceptibility to proteolytic breakdown. Thus, for many cases involving peptide-based leads the structural information of the type described above might be used as a starting point to design smaller constrained peptides or nonpeptidic mimics. This is the case, for example, in the development of endothelin antagonists described below.

2.1.2.2 Endothelin as a Lead in Ligand-Based Design. Endothelin (ET), shown in Fig. 12.10, is a 21-amino acid endothelial-derived constricting factor that has gained prominence as a pharmacological lead molecule. Interest in the peptide arose because of its poterit renal, pulmonary, and neuroendocrine activities. Endothelin and its isoforms have been implicated in a wide variety of disease states including ischemia, cerebral vasospasm, stroke, renal failure, hypertension, and heart failure (30). It exerts its pharmacological effect by acting on specific G-protein–coupled receptors. In mammalian species two receptors, ET_A and ET_B, have been cloned; both are widely distributed in human tissue and are distinguished by different responses to various ET isoforms.

The NMR-derived three-dimensional structure of ET-1 consists of several distinct regions, including a random-coil N-terminus, a β-turnover residues 5–8, followed by a short helical region and a flexible C-terminal tail (as summarized in Ref. 31). The presence of the flexible tail in solution is not surprising, as may be imagined from the primary sequence shown in Fig. 12.10. Although solution structures of ET and its analogs (32–46) determined by NMR have been valuable in defining the gross conformation of these molecules, the flexibility of the tail in solution makes it difficult to extrapolate to the bound state. Indeed, an X-ray structure of ET-1 has quite a different structure for the C-terminal tail than for the random-coil arrangement in solution (46). The bound conformation may be different again.

There is clearly an advantage to having lead molecules with reduced flexibility, given that their solution conformation will intrinsically provide a better reflection of the bound conformation. In addition, the development of a more rigid drug will reduce unfavorable entropic contributions to binding energy. Indeed, a range of small cyclic peptides that are ET_A- or ET_B-selective antagonists have been discovered and provide valuable leads to the development of potential therapeutics. NMR studies have been instrumental in determining their solution conformations. For exam-

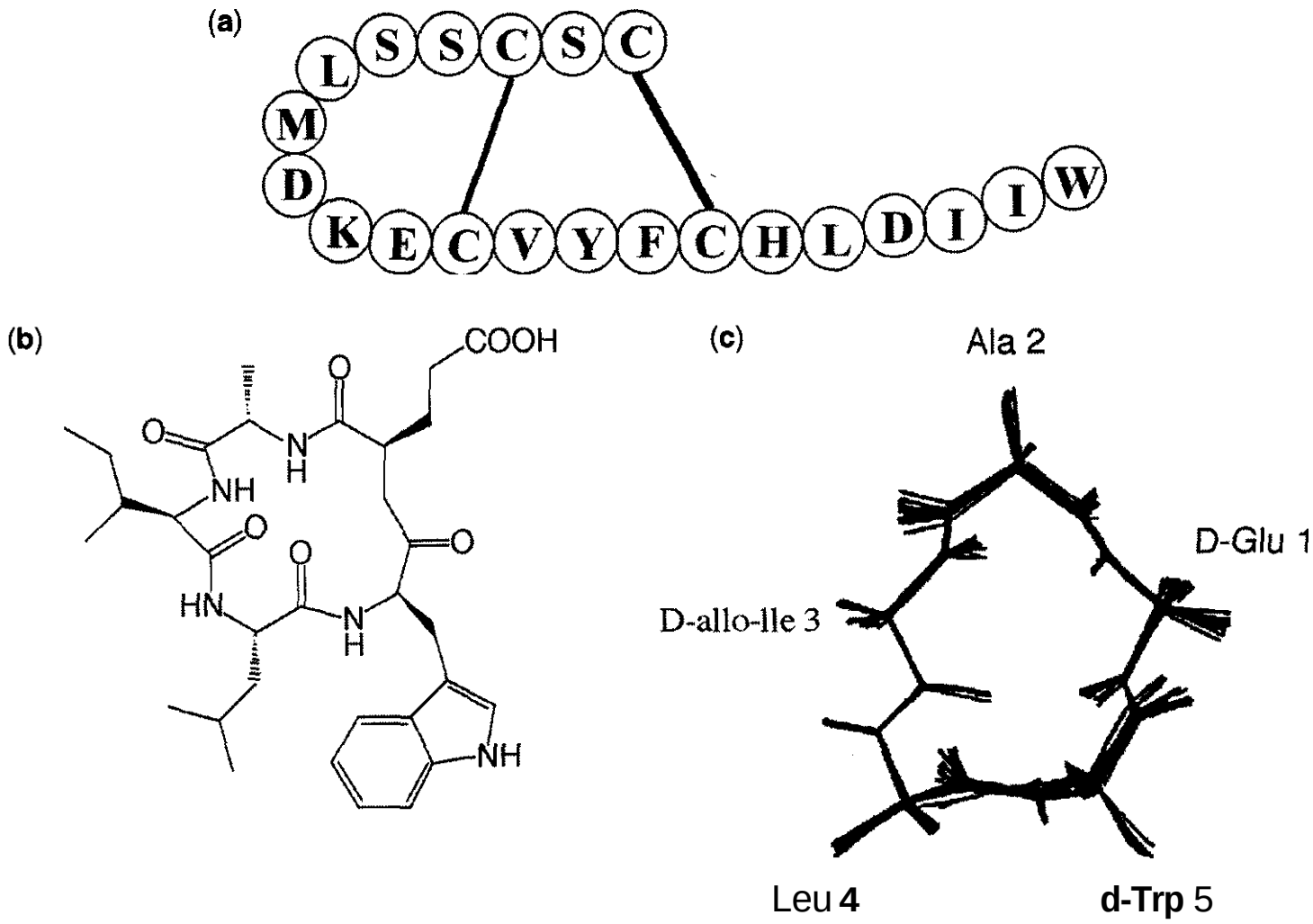


Figure 12.10. (a) Primary sequence and disulfide connectivities of endothelin-1 **(ET-1).** (b) Primary structure of the cyclic endothelin antagonist BE18257B and (c) a family of 36 NMR structures, which demonstrate the well-defined nature of the cyclic peptide backbone.

ple, the rather well defined solution conformation (47) of the ET_A-selective antagonist BE18257B (shown in Fig. 12.10) contrasts with the flexibility of the tail region of ET that this peptide is thought to mimic. The discovery and development of these molecules illustrate the principle that cyclic peptides are often more suitable than linear peptides as lead ligands in drug design. In addition to their better-defined and less-flexible conformations than those of their linear counterparts, they generally have improved bioavailability and resistance to protease attack.

We shall return to endothelin as a lead in drug design, in relation to a nonpeptidic antagonist. The underlying theme illustrated by the endothelin example is that ligand-based design often proceeds from initial studies of flexible endogenous molecules (particularly peptides) to constrained mimics (e.g., cyclic peptides) and often culminates in the development of nonpeptidic drug leads. NMR assists by defining the structures of the lead and subsequent molecules.

2.1.3 Instrumental Advances and their Impact on Structure Elucidation. Over the last few years there have been several exciting instrumental developments that promise to dramatically expand the role NMR will play in the drug discovery process. These relate to the combination of NMR with other technologies such as LC and/or MS and the use of NMR to directly monitor reactions carried out on solid-phase resins (8, 13, 22). The latter promises to indirectly enhance drug discovery programs by improving the monitoring and hence efficiency of solid-phase combinatorial synthesis. Effectively, resin-based syntheses can be monitored at successive stages without the need to cleave intermediate products from the resin.

As already mentioned, the additional sensitivity brought about by cryoprobe technology promises to enhance a wide range of NMR applications, but will be particularly important in natural products-based drug discovery. In many cases only limited amounts of pure compounds are isolated from natural products extracts and sensitivity has been a major limit-

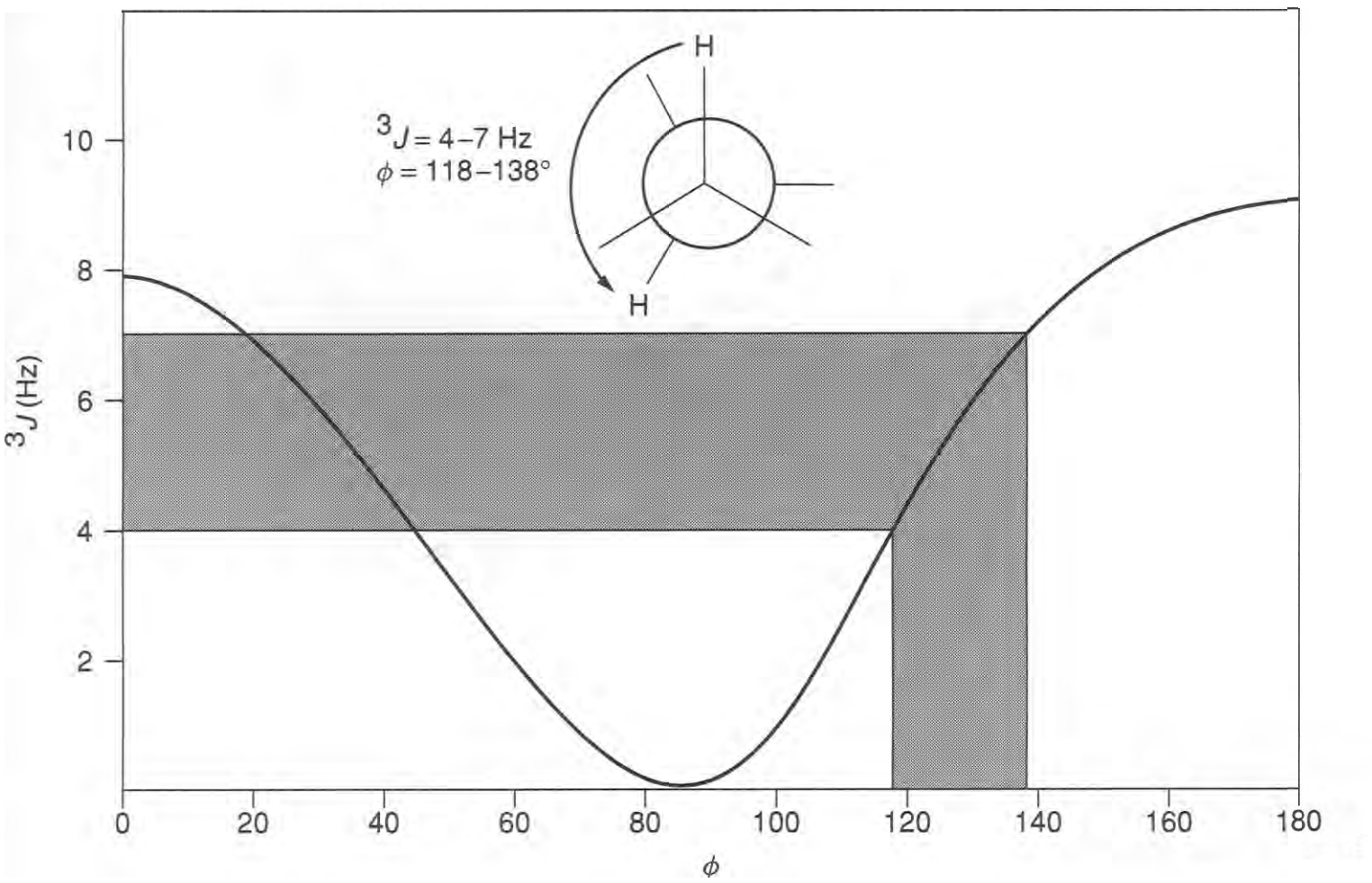


Figure 12.11. Illustration of the Karplus relationship between three-bond scalar coupling constants and the dihedral angle of the intervening bond. The relationship is indicated for the ϕ torsion angle of the H2 and **H3** protons within the rigid core of taxol and related derivatives. See Fig. 12.6 for the structure of taxol.

ing factor on structure elucidation. LC/MS/NMR systems will greatly improve the efficiency of such analyses by minimizing the need for separate sample-handling steps for the different analytical technologies.

2.2 Conformational Analysis

Usually only 1D or 2D NMR methods are required to determine the solution conformation of bioactive ligands. Useful tools include analysis of chemical shifts, coupling constants, and NOEs. An assumption inherent in the application of such studies to drug design is that the solution conformation will be maintained on binding to the receptor. This is justified in the case of relatively rigid ligands. However, for potentially flexible ligands the possibility of changes in conformation on binding must be considered, as noted above for the case of endothelin.

Coupling constants and NOEs are the main NMR parameters used in determining the solution conformations of drug leads. NOEs provide information about through-space proximity. Three-bond vicinal coupling constants are particularly valuable because their dependency on the intervening dihedral angle' through the Karplus relationship allows local geometry to be determined. This is illustrated in Fig. 12.11 for taxol. Although there are several vicinal coupling constants in this molecule (Fig. 12.6), only one $^3J_{H2H3}$ occurs in a region of the molecule that is expected to be conformationally rigid and thus suitable for conformational determination by use of coupling constants. In taxol and a range of analogs this coupling is in the range 4–7 Hz, consistent with partially eclipsed dihedral angles of approximately 120–140" for this ring-constrained structure. This is in good agreement with the X-ray structure of a taxol analog, where the angle is 120". Note that, in general, such a Karplus analysis does not give a unique solution unless several coupling constants sampling the same dihedral angle are present and is reliant on the assumption that the molecule exists only in a single conformation in solution. Although it is generally believed that

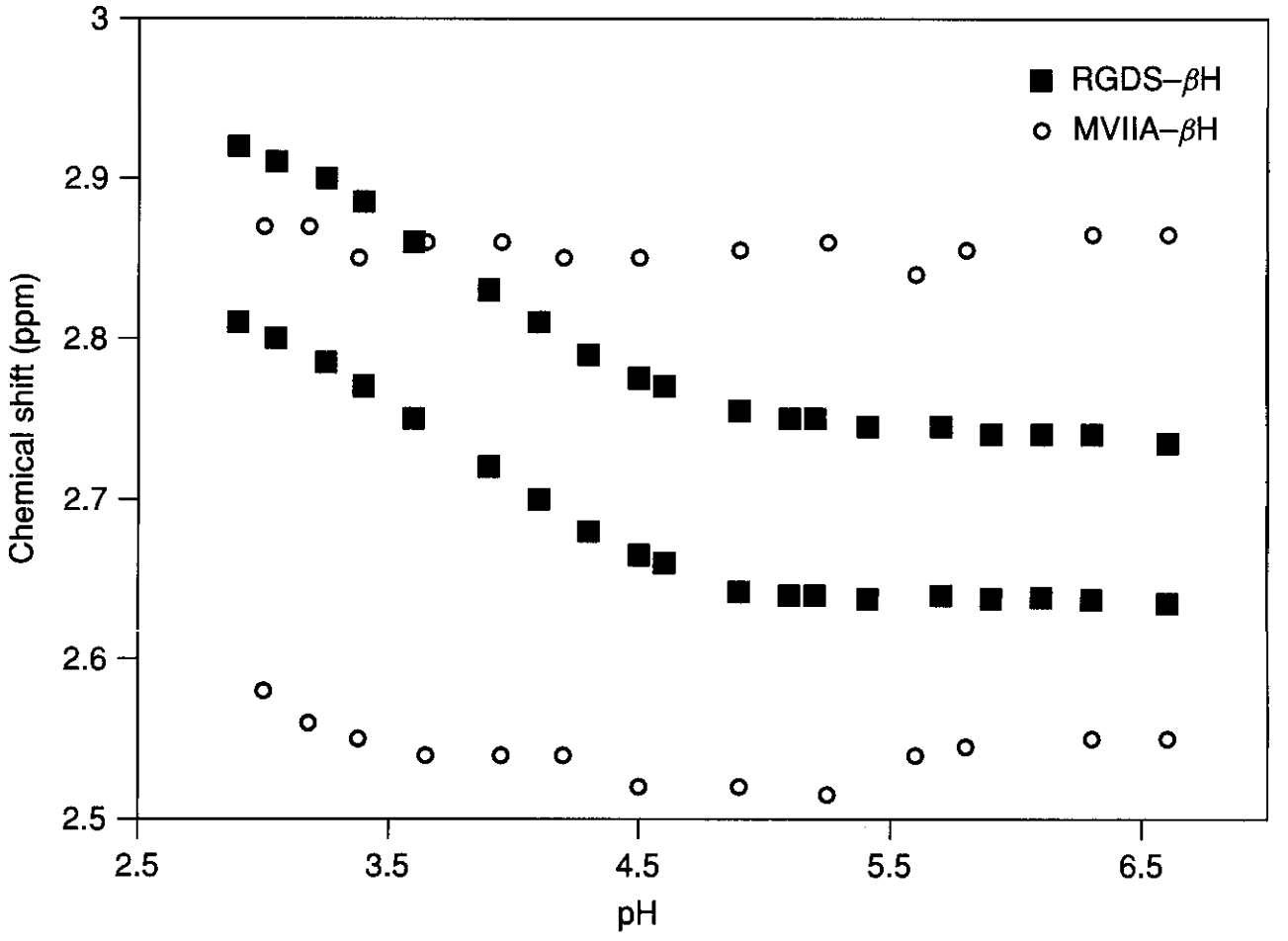


Figure 12.12. Chemical shift changes of the β-protons of Asp14 in MVIIA illustrating the lack of titration of the adjacent carboxyl group, indicating its invovlment in salt bridge. By contrast, the shift of a control random-coil peptide varies with an apprent pK_a value of 3.7, as expected for an uncomplexed carboxyl moiety in peptides.

this is the case for the core of taxol, recent relaxation data (48) described in section 2.4 suggest that this conclusion may need to be reexamined.

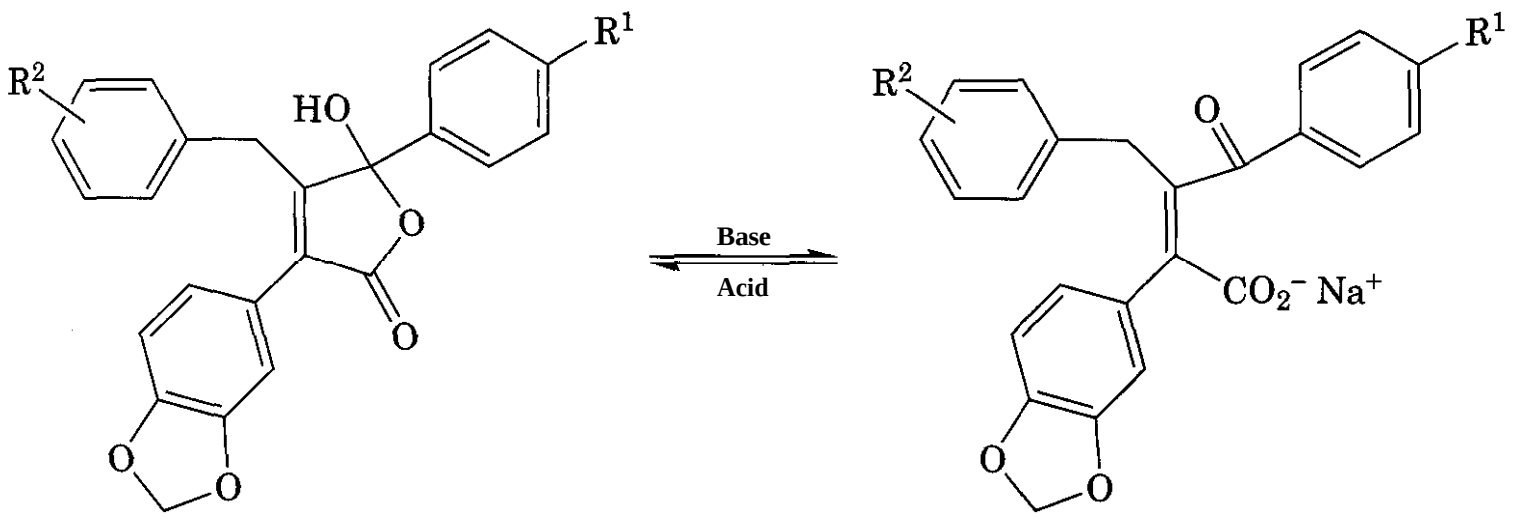
In addition to studies of the taxol core, there have been a large number of studies of the conformations of the side chains of taxol and it appears that these are certainly flexible and that the molecule may adopt both extended and folded conformations of the side chains. In a case like this the observed vicinal-coupling constants are a weighted average of those from the participating conformers.

2.3 Charge State

An advantage of NMR over other structural techniques such as X-ray crystallography is that it has the potential to provide information not only on structure but also on the electronic properties of molecules. Many drug leads contain ionizable groups and a determination of their charge state in solution and/or at the bound site is important in the design of analogs. Simple plots of chemical shifts as a function of pH for nuclei near these ionizable groups provide a convenient way of determining the pK_a value and hence charge state. This is illustrated for ziconotide in Fig. 12.12, where it was suspected that one of the ionizable groups in the molecule, Asp^{14}, may be involved in a stabilizing salt-bridge interaction (28). This was confirmed by noting that the pK_a value for this residue is lowered considerably relative to the usual value for Asp. The β-proton chemical shifts were essentially independent of pH over the range 3–7 (indicating a $pK_a < 3$), whereas those of a control, random-coil peptide, titrated as expected over this range.

2.4 Tautomeric Equilibria

Tautomerization is a relatively common feature of drug molecules that is amenable to analysis through the use of chemical shifts or coupling constants as probes. This was recently demonstrated in a study of some non-peptide endothelin analogs (49). Starting from the modestly active compound (1) (Table

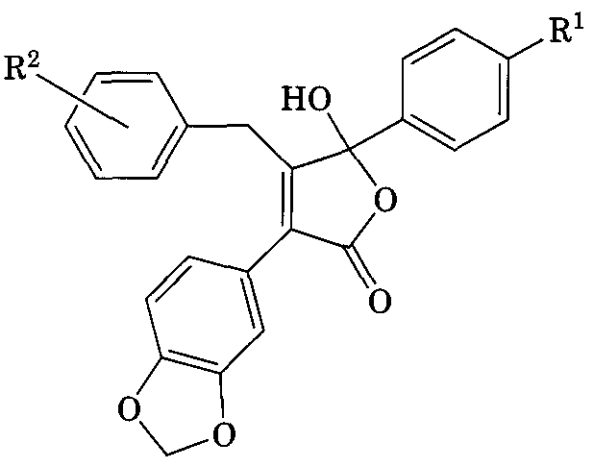


12.5), derived by screening a compound library for ET_A antagonists, the nanomolar inhibitor (2) was developed. Further optimization through examination of electronic and structural requirements led to the subnanomolar inhibitor (3), which was subsequently put forward for evaluation in a number of preclinical disease models for stroke.

These molecules display keto-enol tautomerization, as illustrated in the following structures. The open form keto-acid salts and the closed form butenolides exist in a pH-dependent equilibrium in solution, and at physiological pH both forms exist. In principle, the biological activity could reside in either or both forms.

The extent of tautomerization was established by evaluation of NMR spectra as a function of pH, from 2.65–9.05 (49). At acidic pH, compound (2) exists essentially in the closed butenolide form. Because the pH is slowly raised by addition of NaOD, the spectrum begins to exhibit properties associated with the open form keto-acid, and at basic pH the compound is essentially all in the open form. The coupling pattern shown by the benzylic protons is a particularly characteristic marker of the tautomeric process. At acidic pH the benzylic protons exhibit an AB quartet pattern consistent with the ring-closed structure. As the pH is raised this pattern coalesces to a singlet, broad at neutral pH and sharp at basic pH, as would be expected with the open form keto-acid structure. After the pH was basic, addition of DCl to acidify the solution caused

Table 12.5 Substitution Pattern and Receptor-Binding Affinity of Nonpeptidic Endothelin Antagonists[a]



Compound	R^1	R^2	IC_{50}(n*M*) ET_A	IC_{50}(n*M*) ET_B
(1) PD012527	Cl	H	430	27000
(2) PD155080	OCH_3	H	>0.4	4550
(3) PD156707	OCH_3	3,4,5-OCH_3	0.3	780
(4)	OCH_3	3,5-OCH_3,4-$O(CH_2)_3SO_3Na$	0.38	1600

[a]From Refs. 49 and 50

the spectrum to return to its original appearance, consistent with a reversible tautomerization process.

Identical biological results were obtained with the salt and closed butenolide form in all pharmacological assays, reflecting equilibration at physiological pH. This made it difficult to identify the biologically active form from these experiments alone, although methylation of the OH group in compounds 1–3 resulted in a loss of activity. Because these analogs cannot tautomerize to form open keto-acids, it seems likely that the open form is responsible for activity.

In addition to its impact on the biologically active form, the tautomeric process has profound implications for formulation of drug candidates, as illustrated in some recent further development work on compound (3)(50). Although it is easy to synthesize and isolate water-soluble salts of the keto-acids, once they are placed in aqueous solution the tautomeric equilibrium determines how much of each form is present. Indeed, if the closed butenolide tautomer is sufficiently water insoluble, it can precipitate out of solution and the equilibrium can drive the complete precipitation of the compound. Although (3) has good oral activity, its intravenous use is limited by the insolubility of the closed-form butenolide tautomer without the use of a specific and complex buffered formulation. Thus in recent work a series of water-soluble butenolides was developed (50) to overcome this limitation for parenteral uses. This culminated in the development of (4) (Table 12.5), currently in preclinical evaluation.

This description of the development of (4) provides a good illustration of the fact that the availability of an active molecule is not the end of the drug development pathway, and that formulation considerations can be critical. In this case NMR played a significant role in understanding tautomeric processes that had a direct bearing on solubility and hence formulation.

2.5 Ligand Dynamics: Line-Shape and Relaxation Data

It is increasingly being recognized that the solution molecular dynamics of drugs may have an important role in modulating biological activity (48, 51–54). For example, dynamics may influence entropic contributions to the free energy of binding. In general the more flexible a ligand is, the more unfavorable will be the loss in entropy on binding, assuming a relatively rigid bound state of the ligand. However, in some cases flexibility of a ligand may be a positive factor. This applies, for example, if a degree of flexibility is required to allow a ligand access to a buried active site, or if activation of a receptor requires a conformational change mediated by ligand binding (9). Therefore, a knowledge of the flexibility of lead molecules is an important supplement to the structural and electronic information available from NMR.

The two major NMR methods for obtaining information on ligand flexibility are line-shape analysis and relaxation measurements (usually ^{13}C or ^{15}N T_1, T_2, or heteronuclear NOE measurements). In general terms, the former is sensitive to motions on the milli- to microsecond timescale and the latter to nanosecond timescales. To some extent, structure calculations on peptide-based lead molecules can also give an indication of regions of flexibility from an examination of local regions of disorder among a family of calculated structures. Caution must be exercised because other factors can contribute to disorder, although in many cases there is a connection between disorder in a structural ensemble and molecular flexibility (55). A recent example concerns the solution structures of three isomers of the α-conotoxin GI (56). Attempts to increase structural diversity through the engineering of nonnative disulfide bonds showed that nonnative isomers were not only different in conformation but were also considerably more flexible than the native isomer and had reduced activity.

In an example that illustrates the application of NMR relaxation measurements for studying ligand flexibility, Kessler and colleagues (57) investigated the role of disulfide bonds in the a-amylase inhibitor tendamistat. This small protein contains two disulfide bonds (Cll-C27 and C45-C73) and opening of the latter is known to reduce the melting temperature of the protein (i.e., reduce its stability), but in this case does not affect its α-amylase inhibitor function. The latter observation

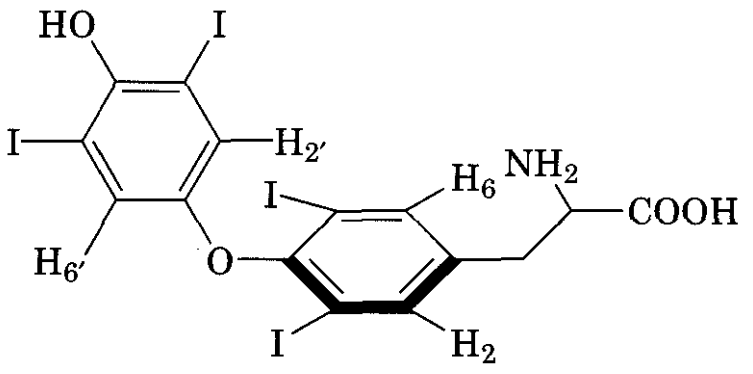
Table 12.6 ^{13}C-NMR Chemical Shift and Relaxation Data for Thyroxine

Position	Chemical Shift (ppm)	Experimental[a]		Theoretical[b]			
				Isotropic Motion		Two-State Internal Motion	
		T_1 (s)	NOE	T_1 (s)	NOE	T_1 (s)	NOE
C2′,6′	127.3	0.63	2.53	0.63	2.96	0.63	2.53
C2,6	142.6	0.63	2.63	0.63	2.96	0.63	2.58
C-α	57.1	0.51	2.37	0.51	2.94	0.51	2.57
C-β	36.4	0.64	2.29	0.64	2.96	0.64	2.51

[a]Measured relaxation data at 75 MHz and 305 K (52).
[b]Theoretical best-fit values based on the indicated models for molecular motion.

suggests that this disulfide bond may not affect either the structure or the dynamics of the molecule. To investigate this possibility heteronuclear NOE measurements were used to determine the effect of the selective removal of the single disulfide bond on the dynamics of the molecule. To assess structural changes, chemical shift differences, intrastrand NOE effects, and protected amide protons were measured for the disulfide-deficient variant C45A/C73A and wild-type tendamistat. Removal of the C45-C73 bond by the C45A/C73A mutation was found to have no influence on the β-barrel structure, apart from very local changes at the mutation sites. ^{13}C and ^{15}N relaxation data showed that the only region of significant internal mobility in either the wild type or variant protein was at the N-terminus. There was little difference in internal flexibility between the two proteins, consistent with their similar a-amylase inhibitory activity. In this case it appears that the role of the C45/C73 disulfide bond is more related to thermodynamic stability and the secretion efficiency of the protein rather than to limiting its flexibility.

The thyroid hormones, exemplified by thyroxine (5), provide an example of the use of both line-shape analysis and NMR relaxation time measurements, to give an insight into the internal flexibility, and perhaps the mode of action, of pharmaceutically important molecules (52, 53, 58). The thyroid hormones act by binding to a nuclear receptor and appear to control receptor function by inducing a conformational change that directs the alignment of functionally critical secondary-structure elements of the receptor (59). Synthetic thyroxine is widely used for the treatment of thyroid disorders and indeed has been the second most widely prescribed drug in the United States over recent years.



(5)

Table **12.6** shows experimental ^{13}C T_1 and heteronuclear NOE data for thyroxine together with best-fit theoretical values for these parameters based on two different models of the motion of thyroxine in solution. A simple isotropic model, in which the drug is regarded as a rigid body tumbling in solution, is clearly unable to simultaneously fit the T_1 and NOE data, whereas the two-state jump model, which incorporates a degree of internal mobility based on rapid flipping between conformational states, is better able to account for the experimental data. Although it is difficult to precisely define the nature of the internal motion, it is clear that rapid (nanosecond) internal motion is present, and this appears to involve small-amplitude torsions about the aromatic ring axes. The correlation time for overall tumbling of thyroxine was deduced to be approximately 0.35 ns, with the internal motion approximately 30-fold faster (52).

Studies of the ^{1}H-NMR spectrum of thyroxine as a function of temperature had earlier

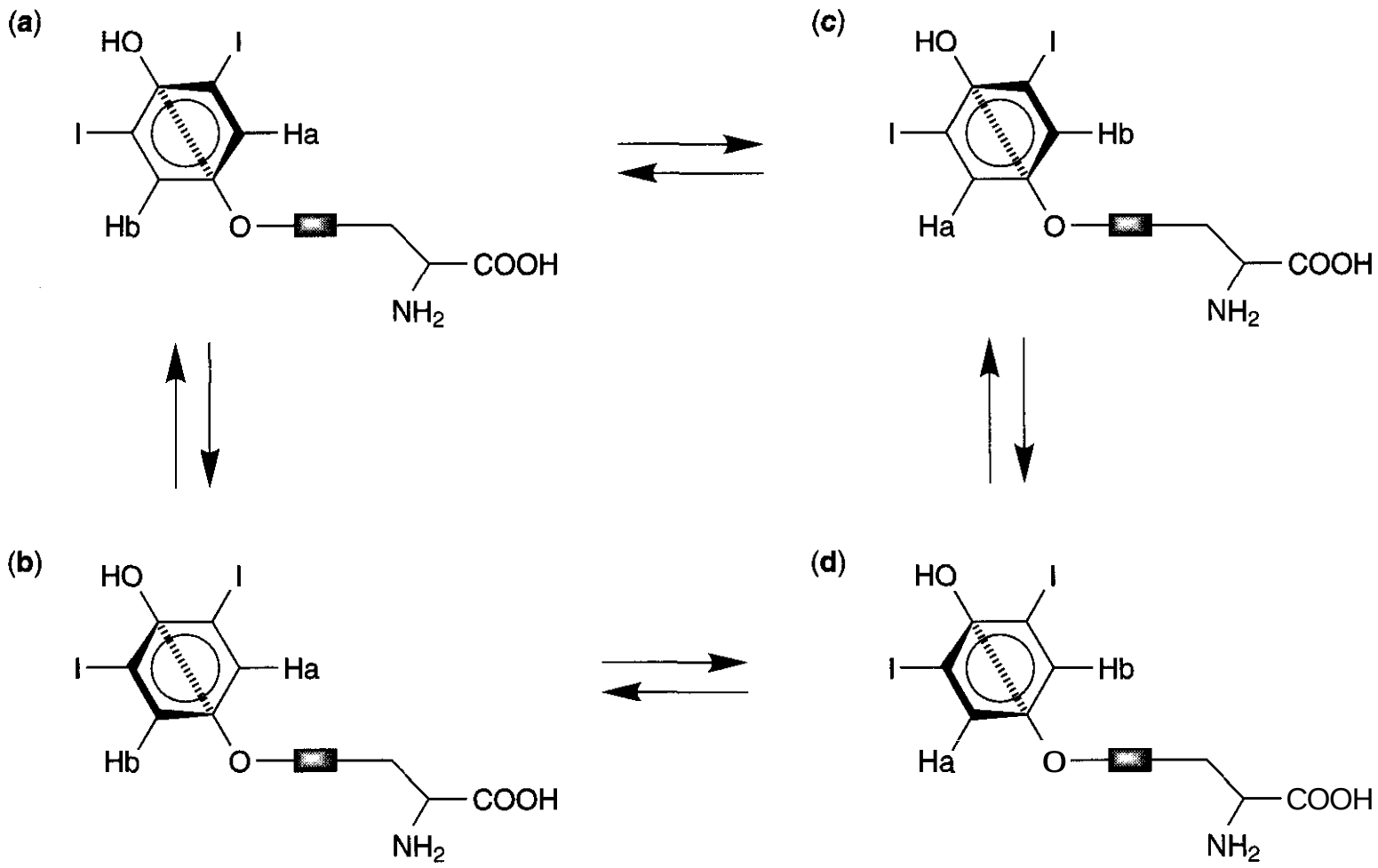


Figure 12.13. Schematic illustrations of motions of the outer ring of thyroxine. The dotted line through the outer ring shows the jump axis about which the ring rotates. (a) Ha is shown in the proximal position and is closer to the viewer than Hb because the torsion angle ϕ' is greater than 0°. This conformation corresponds to one of the two states of the two-state jump model and agrees with the "twist" of the outer ring observed in the crystal structure. (b) Rotation about the dotted line through the center of the outer ring moves Ha away from the viewer and brings Hb toward the viewer. This corresponds to the second state of the outer ring in the two-state jump model. (c) Hb is now in the proximal position and closer to the viewer than Ha. (d) Hb is in the proximal position and is now further from the viewer than Ha. Transition from a to b and from c to d involves small amplitude jumps on the nanosecond timescale and is detected by NMR relaxation measurements. Although not illustrated in the figure, the inner ring also exhibits this type of motion. Transitions a to c and b to d result in 180" flips of the outer ring and exchange of the environments of Ha and Hb. This ring flip occurs on a microsecond timescale and is detected by variable temperature line-shape studies. (Reprinted with permission from Ref. **52.** Copyright 1996 American Chemical Society.)

demonstrated the presence of additional larger amplitude, but slower ring flips (60). At low temperature two signals were seen for the H2′ and H6′ protons. These signals broadened with increasing temperature, then coalesced and sharpened as the temperature was further increased. This was attributed to exchange of the environments of the two protons brought about by 180" rotation of the "outer" ring of thyroxine. Substitution of the observed coalescence temperature (T_c) and the chemical shift difference of the two signals at low temperature ($\delta\nu$) allowed the free energy of activation for this slow ring flip process to be established from equation **12.1** (53, 60).

$$\Delta G^{\neq} = 19.14T_c[9.97 + \log(T_c/\delta\nu)] \quad (12.1)$$

The derived barriers for several thyroid hormones are in the range **36–38 kJ/mol,** which corresponds to large-amplitude ring flips on the milli- to microsecond timescale. From a combination of the relaxation data and the dynamic line-shape analysis data it was possible to propose a unified model that accounts for both the fast and slow internal motions, as summarized in Fig. **12.13.**

In this model, both aromatic rings of the thyroid hormones jump rapidly between two energetically equivalent conformations on a

nanosecond timescale (a ⇔ b and c ⇔ d in Fig. 12.13). The half-angle of the jump varies, depending on the solvent, corresponding to an average displacement of about 90" between the two extreme jump positions. These separate states are not detectable on the chemical-shift timescale but lead to an average proximal environment for Ha and an average distal environment for Hb (attributed to rapid interchange between a and b in Fig. 12.13), which are seen in the low temperature spectra. However, these fast motions are detected by relaxation studies. Although the rate of this motion is rapid, its amplitude is not sufficient to average the environment of proximal and distal protons. Occasionally (about once every 1000 jumps) the outer ring jumps further than the nominal 90" range, exchanging the environments of the proximal and distal protons (a ⇔ c and b ⇔ d in Fig. 12.13). Although the actual rate of an individual ring flip is rapid, the effective rate of the process is on the microsecond timescale, because on average a large number of small amplitude jumps occur for every large amplitude ring flip. It is the exchanging of proximal and distal protons on the microsecond timescale that is detected by the variable-temperature line-shape studies.

The fact that thyroxine is apparently able to so freely move over a moderately large region of conformational space has implications for receptor binding. The crystal structure of the thyroid receptor ligand-binding domain complexed with the thyroid agonist 3,5-dimethyl-3'-isopropylthyronine (59) shows that the thyroid hormones bind at the center of the hydrophobic core of the ligand-binding domain and may play a structural role in the conformational changes that activate the receptor. The structures of the retinoid-X receptor ligand-binding domain (61) and the retinoic acid-retinoic acid receptor ligand-binding domain complex (62) indicate that significant conformational changes accompany ligand binding in those cases. The conformational flexibility exhibited by the thyroid hormones may also be required for binding. It has been suggested that the rapid "wiggling" of the aromatic rings could enable the hormone to work its way to the center of the ligand-binding domain as the protein reorders itself about the ligand and may in fact trigger receptor conformational changes (52).

As briefly mentioned earlier, taxol provides another example where relaxation time measurements provide an insight into dynamics processes. Although it is generally thought that the taxane core is rigid, ^{13}C relaxation data suggest that a degree of flexibility (on the nanosecond timescale) may be present and may vary for different taxol analogs (48). In particular it appears that the removal of certain side chains may introduce additional flexibility into the core region that would not easily be predicted based on a simple inspection of the structure.

Another example of the application of line-shape analysis to ligand dynamics is described in Section 3.2 for the drug trimetrexate when bound to dihydrofolate reductase (DHFR). From that example and earlier studies on DHFR (63–65), it is clear that the techniques described above can equally be applied to ligands when bound to their receptor. In some cases significant but highly specific mobility appears to be present at the bound site.

2.6 Pharmocophore Modeling: Conformations of a Set of Ligands

Determination of the conformations of a range of ligands that all act at the same receptor site can provide significantly more information than just a single ligand structure. With a sufficiently broad range of ligands, it is often possible to generate a pharmocophore model of the receptor site, deduced based on conserved structural features and the conformations of the ligands. This has been done recently, for example, for the o-conotoxins, the broad class of conotoxins to which Ziconotide (or MVIIA, mentioned above) belongs. From structural studies of a range of o-conotoxins and from literature data on various mutants with altered binding affinities, it was determined that only a localized region of the surface of these molecules is involved in receptor binding (66). This allowed a pharmocophore model of putative receptor-binding pockets to be developed.

The advantage of such a pharmocophore model is that smaller, nonpeptide molecules that might have improved stability and bioavailability over their peptidic counterparts

can, in principle, be designed. The NMR approach used in such pharmocophore modeling often involves a combination of many of the techniques already described. By determining information about structure and electronic properties for a range of different ligands, all acting at the same receptor site, it is often possible to infer information about the binding site, even if direct structural studies of this site are not possible.

2.7 Limitations of Analog-Based Design

Although a determination of the structure of bioactive molecules is of key importance, there are distinct limitations on the use of solution structures for drug design. In particular, unless the molecule is rigid there is no certainty that the solution conformation is the same as the bioactive bound conformation. For this reason there has been a shift over recent years to approaches in which information about the bound state is obtained. The other approach has been to probe the bound conformation by making a range of constrained analogs of a flexible lead molecule, as illustrated earlier for endothelin.

The most direct way of determining the conformation of a drug lead is to determine the full three-dimensional structure of its receptor complex. This has now been achieved in a significant number of cases but represents a substantial undertaking, as described in later sections of this chapter. A simpler approach that has also been applied is to use transferred NOE methods, as described below. This approach fits at the interface of ligand-based design and receptor-based design. It fits with the former because no knowledge of the receptor structure is required, but it also fits with the latter because it requires the macromolecule of interest to be included in the mixture to be analyzed. It is appropriate therefore to introduce the topic here but also to discuss it further in Section 3.

2.8 Conformation of Bound Ligands: Transferred NOEs

In ligand-based drug design it is not necessary to know the structure of the receptor, or even the location of the binding site, although the conformation of the ligand bound to the receptor is crucial. It is clearly better if this can be measured directly rather then be inferred from the conformation of the free ligand. In certain circumstances this information on the bound conformation can be obtained from the transferred NOE (TrNOE) technique (67, 68). This method takes advantage of the fact that NOEs build up more rapidly in a ligand-macromolecule complex than they do in free ligand, and given appropriate exchange conditions for a mixture of ligand and macromolecule (typically satisfied for $K_D \geq 10^{-7}$ M^{-1}), then signals from a free ligand may be used to determine the bound conformation.

The theory of the technique was reviewed previously (69) and recent developments that minimize potential artifacts from spin diffusion have been described (5). Because it is not necessary to monitor signals from the macromolecule in this technique, it is usually present in substoichiometric amounts, thus requiring only minimal amounts of what is sometimes the more expensive component of ligand-macromolecule complexes. In addition, the molecular weight restrictions inherent in full 3D-structure determinations of complexes are ameliorated and the conformations of ligands bound to very large macromolecules may be determined. For example, the technique was recently used to determine the structure of an antibiotic bound to the ribosome (70). A range of other applications including enzyme-substrate, protein-carbohydrate, and protein-peptide interactions have recently been summarized (5).

In addition to its application as a tool for determining bound conformations of ligands, the TrNOE method has also been used recently as a screening aid for the identification of ligands from mixtures th atbind to a protein of interest. This application is addressed in more detail later in this chapter.

3 RECEPTOR-BASED DESIGN

Receptor-based design refers to the process of determining the three-dimensional structure of a macromolecular target and using this information to design ligands to interact with it. In general there have been few cases where the structure of a macromolecule or receptor

alone has been successfully used to design, de novo, a ligand to interact with that receptor. However, such an approach is likely to become more common with improved computer-based approaches to molecular design in the future (71).Currently, the most common approach is to study a ligand-macromolecule complex and to initiate the design process based on the interaction of the lead ligand with the macromolecule.

Although the structure of the macromolecule alone is of less interest than that of the complex, in many cases a determination of the structure of the complex follows from earlier studies on the unbound macromolecule. It is thus useful to describe the approaches to structure determination of macromolecular targets. This is followed by a discussion of the dynamic aspects of protein structures in Section 3.2, before addressing the main topic of macromolecule-ligand interactions in Section 3.3.

3.1 Macromolecular Structure Determination

The two major techniques for determining three-dimensional structures of proteins or nucleic acids are X-ray crystallography and NMR spectroscopy. The crystallographic approach to structure determination is described elsewhere of this volume and here the focus is on NMR. NMR has been used to determine the structures of proteins only for about the last 15 years, with the first NMR structure determination being made in 1985. NMR has a number of advantages over X-ray crystallography, including the fact that the requirement that the protein needs to be crystallized is avoided, and that the dynamic information available from NMR studies complements the structural information. A major disadvantage of NMR spectroscopy, though, is that it is currently limited to the determination of structures of <35 kDa. With the development of new NMR techniques, such as TROSY (20), this seems certain to increase significantly over coming years, although the fact remains that among all structures currently deposited in the protein database the average size of NMR structures is about 8 kDa (72), substantially smaller than the average size of protein structures determined by X-ray crystallography. Despite this limitation, NMR has made major inroads into the macromolecular structure determination process, and currently approximately one-fifth of all new structures deposited in the protein database have been determined by NMR spectroscopy.

3.1.1 Overview of Approach. The basis for structure determination by NMR is that, by determining a large number of distance restraints between pairs of protons, it is possible to reconstruct a three-dimensional image of the molecule. These distance restraints are derived primarily from nuclear Overhauser effect (NOE) measurements, which detect distances up to about 5 Å. Over recent years such distance restraints have been supplemented by a range of other restraints, including dihedral angle restraints derived from coupling constant measurements and orientation restraints derived from residual dipolar couplings. These restraints are input into a simulated annealing algorithm, which is used to calculate a family of structures consistent with the restraints.

NMR is unique in that it can provide detailed and specific information on molecular dynamics in addition to structural information. The use of relaxation time measurements allows the relative mobility of individual atomic positions within a macromolecule to be determined. The dynamic information obtained includes not only the rates or frequencies of internal motions but also their amplitudes. Such amplitudes are often expressed by order parameters. Not surprisingly, it is observed in many cases that the termini of proteins are more flexible than internal regions. More interestingly, NMR has provided a number of examples where internal loops in proteins have been shown to have dynamics that may be associated with their function. A good example of this is HIV protease, where NMR studies have identified reduced-order parameters in the flap region of the molecule that may reflect flexibility to allow entry of substrates or inhibitors into the active site.

In summary, a major strength of NMR is that a global picture not only of the structure but also of the dynamics of the macromolecular target is obtained. Further, NMR provides information on ionization states of titratable

groups and other electronic features within macromolecules that may have an impact on ligand binding and function.

3.1.2 Sample Requirements and Assignment Protocols. Structure determination by NMR typically requires 500 μL of a 1–2 mM solution of the protein of interest. It is important that the macromolecule does not aggregate because this causes spectral broadening and may preclude assignment. The sample should preferably be stable in solution over the extended period of time required to collect the range of NMR experiments (73–75) needed for assignment and structure determination. Individual experiments may last from a few hours to several days, with several weeks of data acquisition required for studies of larger proteins.

The particular set of NMR experiments required for NMR structure determination depends on the size of the protein. It suffices to say that for smaller proteins ($\leq$7 kDa) it is usually possible to determine the structures, mainly using 2D NMR, without the need for isotopic labeling, by use of procedures described in detail above for Ziconotide. For proteins in the range 7–14 kDa, ^{15}N-labeling and a combination of 2D/3D NMR experiments is usually sufficient, whereas for larger proteins ^{13}C/^{15}N labeling and 3D or 4D NMR is more or less mandatory. For proteins at the top end of the currently accessible range (25–35 kDa), there are additional advantages associated with partial deuteration of the protein.

3.1.3 Recent Developments. A number of recently developed methods offer the potential for improving the quality of NMR structures and for increasing the size of proteins that can be examined. In particular, the use of residual dipolar couplings and of anisotropic contributions to relaxation provide new kinds of restraints that promise to lead to more accurate NMR structures (74, 76). As already mentioned the TROSY method (20) exploits relaxation phenomena to produce spectra with narrow lines and promises to significantly expand the size of protein targets that can be examined by NMR, from the current limit of about 35 kDa to perhaps >100 kDa.

Another development that is likely to have a significant impact is the increasing number of structural genomics programs being developed. The demands arising from such programs will no doubt stimulate new methods for the large-scale production of labeled proteins (77, 78), and for speeding up the rate of structure determination by both NMR and crystallography.

3.1.4 Dynamics. Proteins exhibit a range of internal motions, from the millisecond to nanosecond timescale, and a full understanding of how small drugs might interact with such a "moving target" requires more than just the time-averaged macromolecular structure. Thus, over recent years much effort has been directed toward defining motions within proteins.

The most commonly applied approach has been to use ^{13}C or ^{15}N relaxation parameters such as T_1, T_2, and the heteronuclear NOE to derive correlation times for overall motion, together with rates and amplitudes of internal motions (79). Although the precise interpretation of the NMR relaxation data in terms of motional parameters remains dependent on the appropriateness of the motional model chosen, the results from many studies on the dynamics of proteins are sufficiently clear to confirm that nanosecond timescale motions in proteins are common. The functional significance of motions on the nanosecond timescale remains unclear and so far there have been few cases where significant differences in motions on this timescale between ligand-free and ligand-bound forms of proteins have been measured. It will be interesting to assess the functional significance of such motions as more data become available. However, slower motions have been correlated with function in a number of proteins, with a good example being HIV protease, described in more detail in Section 4.2.

Relaxation measurements require a considerable investment of spectrometer time and in some cases it may be possible to derive basic information about molecular dynamics from the structural ensemble alone. Although regions of disorder can reflect factors other than dynamics, a recent analysis (55) suggests that ill-defined regions in structural ensembles often do reflect slow, large-amplitude motions. Even if relaxation measurements are

done, it is often not necessary to undertake extensive analysis to derive correlation times, given that trends are often apparent from the raw experimental data. For example, in the case of tendamistat, described above, it is clear directly from the heteronuclear NOE data that significant internal mobility is present at the N-terminus.

3.1.5 Nucleic Acid Structures. Most of the discussion on macromolecular targets so far has focused on proteins. DNA represents another valuable target in drug design. Most studies in which DNA is the target are done using short model oligonucleotides to mimic the binding region of DNA. The regular repeating nature of DNA structures makes this a more successful approach than similar attempts to dissect out binding regions of receptor proteins, where often the whole protein must be present to maintain a viable binding site. Similar comments apply for RNA, where improvements in synthetic methods have led to an increasing number of structure determinations over recent years. The principles involved in structure determination of nucleic acid targets are similar to those of proteins, but in practice nucleic acid structures are somewhat more difficult to solve.

3.1.6 Challenges for the Future: Membrane-Bound Proteins. The majority of targets for currently known drugs are membrane-bound receptors, yet this represents the class of proteins for which least structural information is known. Membrane proteins are notoriously difficult to characterize at a structural level because they are difficult to crystallize, thus inhibiting X-ray crystallographic studies, and are both too large and too difficult to reconstitute in suitable media for NMR studies. Nevertheless, solid-state NMR methods are beginning to show promise that eventually such targets may be structurally characterized (80). Rotational resonance solid-state NMR measurements, for example, allow precise distances to be measured in membrane-bound proteins (81).

3.2 Macromolecule-Ligand Interactions

Macromolecule-ligand interactions are integral to a wide range of biological processes, including hormone, neurotransmitter or drug binding, antigen recognition, and enzyme-substrate interactions. Fundamental to each of these interactions is the recognition by a ligand of a unique binding site on the macromolecule. Through an understanding of the specific interactions involved it may be possible to design or discover analogous ligands with altered binding properties that might inhibit the biochemical function of the macromolecule in a highly specific manner. The study of macromolecule-ligand interactions thus forms the cornerstone of most structure-based drug design applications. The macromolecule of interest may be a protein or a nucleic acid, although the majority of drug design applications have focused on protein-ligand interactions. For this reason we will refer mainly to protein-ligand interactions in the following discussion, but will include some examples of drug-DNA interactions.

3.2.1 Overview. There are several important aspects of macromolecule-ligand interactions that have a bearing on structure-based design. The simplest question that might be asked is "what is the strength of the binding interaction?," whereas the most detailed task would be to precisely define the atomic coordinates of the complete protein-ligand complex. In between these extremes there are many other questions important to the drug design process; these include questions about the binding stoichiometry and kinetics, the conformation of the bound ligand, and about the nature of functional group interactions between the protein and bound ligand. These and other important questions were introduced briefly in Table 12.3 and are examined in more detail later in this section. Before doing this it is first necessary to consider NMR timescales because the ability of NMR methods to answer questions about macromolecule-ligand complexes depends critically on the kinetics of the binding interaction. Section 3.2.2 thus describes how various NMR parameters depend on binding kinetics and in particular how fast- and slow-exchange conditions affect the interpretation of NMR data.

Having identified the exchange regime, the task then becomes to decide which NMR parameters can be used to answer the questions

Table 12.7 NMR Parameters and Their Changes on Binding

Parameter	Difference	Typical Magnitude[a]
Chemical shift	$\nu_L - \nu_{ML}$	*0–1000* s^{-1}
	$\nu_M - \nu_{ML}$	*0–500* s^{-1}
Coupling constant	$J_L - J_{ML}$	*0–12* s^{-1}
	$J_M - J_{ML}$	0–12 s^{-1}
Relaxation rate[b]	$1/T_L - 1/T_{ML}$	*0–50* s^{-1} (for T_1, larger for T_2)
	$1/T_M - 1/T_{ML}$	0–10 s^{-1}

[a]Ranges are approximate only and larger effects may be seen in some cases.
[b]$1/T$ refers to either $1/T_1$ or $1/T_2$.

posed above about the complex. Many of the NMR parameters that were described earlier for deriving information about ligands are also applicable to studies of complexes. These include chemical shifts, NOEs, and relaxation parameters. However, the presence of two interacting partners means that there are some differences in the way such parameters are measured and this has led to the development of several techniques that are particularly important for the study of macromolecule-ligand interactions, including chemical-shift mapping, isotope editing, and various NMR titrations. Section 3.2.3 describes these techniques. Finally, illustrative examples of the application of these techniques to specific drug design problems are given in Section 3.2.4.

3.2.2 Influence of Kinetics and NMR Timescales. Macromolecule-ligand interactions are characterized by an equilibrium reaction that potentially has a wide range of affinities and rates:

$$M + L \leftrightarrow ML$$

The rate constant for the forward reaction is referred to as the on rate (k_{on}), whereas dissociation of the complex is characterized by the reverse rate constant, k_{off}. The equilibrium constant for this interaction, represented in terms of the dissociation constant of the complex K_D, reflects a balance of the on and off rates, as shown in Equation 12.2:

$$K_D = [M][L]/[ML] = k_{off}/k_{on} \quad (12.2)$$

For many protein-ligand interactions k_{on} is of the order of $10^8\ M^{-1}\ s^{-1}$, and is typically quite similar for different ligands. The observation that K_D values may vary over a wide range, typically from millimolar to nanomolar (i.e., $K_D = 10^{-3}$–10^{-9} M) for cases of interest, is a reflection of a variation in k_{off} for different ligands. Consideration of the k_{on} value above and the range of K_D values noted suggests a range in k_{off} from 10^{-1} to $10^5\ s^{-1}$. The lifetime of the bound complex ($\tau_{ML} = 1/k_{off}$) may thus vary from much less than a millisecond to tens of seconds (10^{-5} to 10 s based on the above off rates). The exchange rate for the second-order binding process is given by (82):

$$k = 1/\tau = 1/\tau_{ML} + 1/\tau_L = k_{off}(1 + p_{ML}/p_L) \quad (12.3)$$

where p_{ML} and p_L are the mole fractions of bound and free ligand, respectively.

The appearance of an NMR spectrum of a protein-ligand complex is dependent on the rate of chemical interchange between free and bound states. In particular, the effects of exchange on an individual NMR parameter (e.g., chemical shift, coupling constant, or relaxation rate) depend on the relative magnitude of the exchange rate and the difference in the NMR parameter between the two states. The cases where the rate of interchange is greater than, about equal to, or less than, the parameter difference are referred to as fast, intermediate, and slow exchange, respectively, as indicated in Table 12.7.

Table 12.7 shows that the changes in chemical shifts on ligand binding (for signals either from the ligand or from the macromolecule) are in general greater than those for coupling constants or relaxation rates. Given that 100 s^{-1} might represent a typical exchange rate between free and bound states, it is clear that

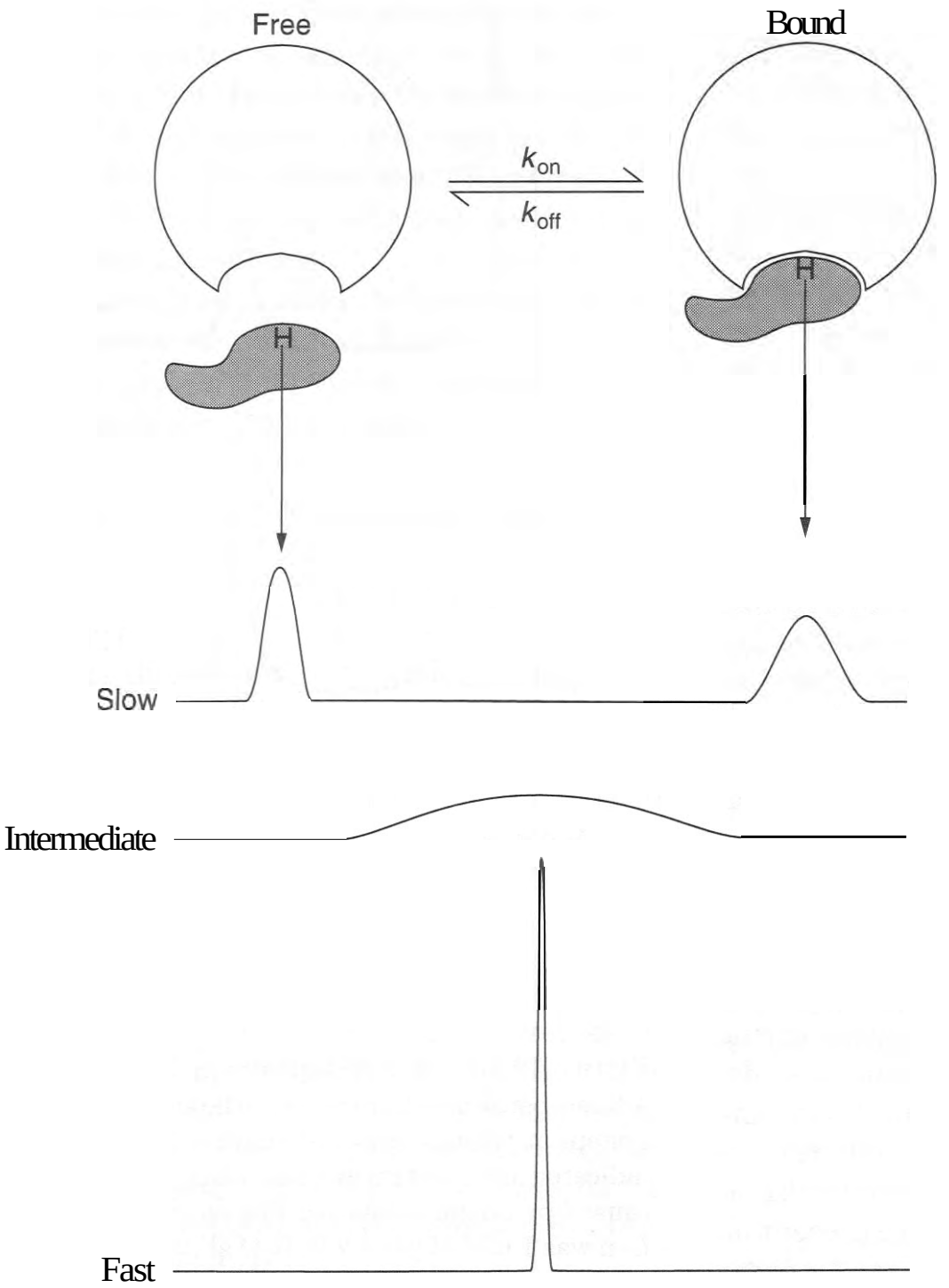


Figure 12.14. Schematic illustration of the effects of slow, intermediate, and fast exchange on the appearance of peaks in NMR spectra of macromolecule-ligand complexes. In the slow exchange case separate peaks are seen for free and bound forms. Note the broader peak for the bound ligand because it now adopts the correlation time of the macromolecule. In the fast exchange case only an averaged peak is observed.

individual NMR signals may be found in either slow, fast, or intermediate exchange on the chemical-shift timescale, but it is more likely that couplings or relaxation parameters will be in fast exchange. Thus, in most cases where the term "NMR timescale" is used in the literature, it refers to the chemical-shift timescale. The table also emphasizes that there are two types of signals that can be monitored, those from the ligand and those from the macromolecule. In general, the typical magnitude of changes to chemical shifts or couplings of either type of signal on binding are similar, although the changes to ligand signals may be larger than those from the macromolecule. However, changes to relaxation parameters for signals from ligands are much more likely to be greater than those for protein signals. This reflects the sensitivity of relaxation parameters to molecular mobility: a ligand undergoes a greater relative change in mobility on binding than does a protein, given that the relative increase in molecular weight in the complex is much greater for the ligand than for the protein.

The exchange regime (slow, intermediate, or fast) determines how a spectrum of a protein-ligand mixture changes during a titration, or as a function of temperature. Figure 12.14 schematically illustrates the various exchange regimes for macromolecule-ligand binding interactions. Slow exchange, corresponding to tight binding, is potentially the most useful regime, given that much detailed information on the nature of a complex can be deduced in this case. Nevertheless, fast ex-

change also allows valuable kinetic and thermodynamic parameters to be derived. The analysis is more complex for intermediate exchange and few quantitative studies are attempted for this situation.

3.2.2.1 Slow Exchange. This situation applies when the rate of exchange is much slower than the difference in chemical shifts between the two states (i.e., $k \ll \nu_B - \nu_F$), where we now change to a nomenclature using subscripts to refer to the bound (B) and free (F) states. It should be understood that the signals may derive from either ligand or macromolecule, so although B is always related to the ML complex, F might refer to either free ligand (L) or free macromolecule (M). In this situation separate peaks are potentially observable for both free and bound states at their respective chemical shifts. Whether such signals are actually observed depends on the mole ratio of the individual species in a titration, and on the signals not being obscured by overlap or broadening.

Addition of a ligand to a solution of a protein results in the appearance of new signals attributed to bound protein resonances, with a concurrent decrease in the intensity of the free protein resonances, reflecting the decreased proportion of free protein during the titration. Once a stoichiometric mole ratio is achieved (usually 1:1, but sometimes 2:1 or higher if multiple binding sites are present on the protein), peaks from free ligand appear with increasing intensity as the excess of free ligand increases.

From such a titration it is possible to determine the stoichiometry of the complex, together with the chemical shifts of the bound states of the ligand and protein. In 1D NMR spectra, overlap of peaks makes it difficult to monitor more than a few resonances from either species and such studies are most readily done when there is a well-resolved signal on one of the interacting species. Selective isotope labels have been used in the past for such studies but it is now more common to use uniform ^{15}N- or ^{13}C-labeling of the protein and detect the chemical shifts in 2D HSQC spectra. It is often more difficult to label the ligand but in some cases the presence of rare nuclei such as ^{19}F can be used to advantage. A good example is the binding of the inhibitor 4-fluorobenzenesulfonamide to the enzyme carbonic anhydrase (83). Figure 12.15 shows ^{19}F spectra of the enzyme inhibitor complex at various mole ratios. The broadened peak for the bound ligand has a chemical shift of approximately 6 ppm and is in slow exchange with the peak from free ligand at 0 ppm. The stoichiometry of the complex in this case is 2:1, so that no signal from free ligand is visible until more than 2 moles of inhibitor are present. Addition of increasing amounts of ligand results in an increase in the free ligand signal, but no change in the bound ligand signal.

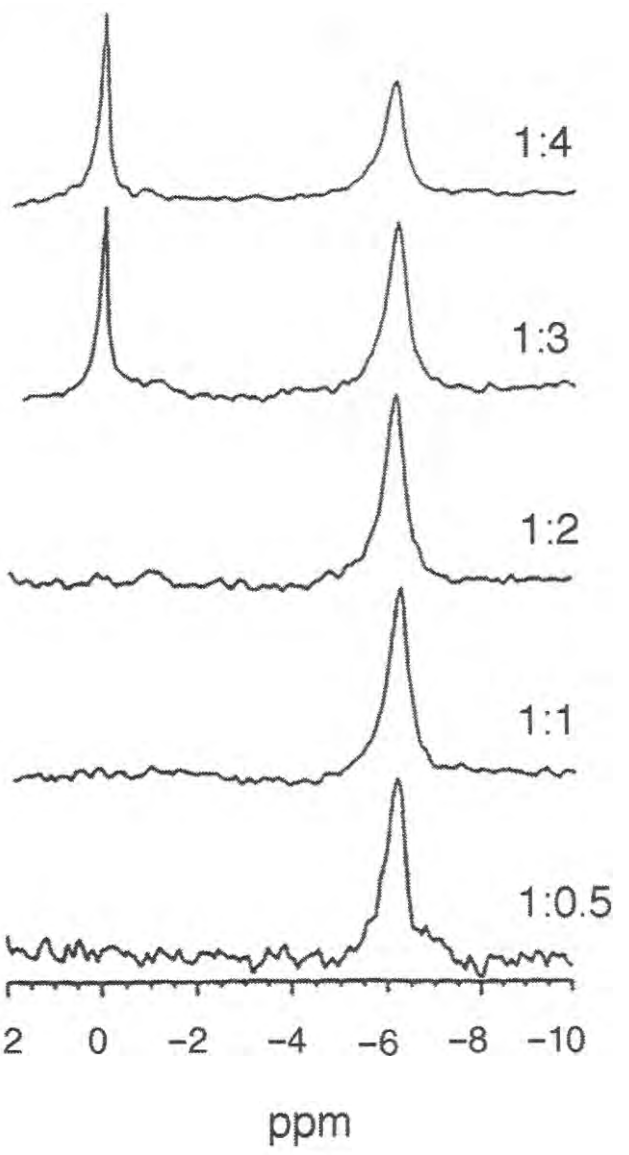


Figure 12.15. ^{19}F NMR spectra at 282 MHz of the 4-fluorobenzenesulfonamide-carbonic-anhydrase-1 system at various ratios of enzyme to inhibitor, as indicated on the traces. The peak at −6 ppm is caused by bound inhibitor. The enzyme concentration was 1 m*M* at pH 7.2 in D_2O at 25°C. (Reprinted with permission from Ref. 83; © 1998, American Chemical Society.)

Determination of the binding constant from slow exchange spectra is not usually attempted. Generally for slow-exchange condi-

tions to exist in the first place, the binding is submicromolar in affinity and non-NMR methods are more suitable for determining affinities in these cases. NMR studies are done at millimolar concentrations, making it difficult to determine K_D with any accuracy for tight binding systems.

In principle, kinetic information on the complex can be obtained from slow-exchange spectra, as seen from the expressions for T, for free and bound ligand signals:

$$1/T_{2F,obs} = 1/T_{2F} + k_{off}p_B/p_F \qquad (12.4)$$

$$1/T_{2B,obs} = 1/T_{2B} + k_{off} \qquad (12.5)$$

Because the linewidth of a peak is related to T_2 by

$$LW = 1/\pi T_2 \qquad (12.6)$$

then measurements of linewidth during a titration can be used to derive k_{off}. Equations 12.4 and 12.5 show that, although the signal from bound ligand is independent of concentration, that of the free ligand decreases in linewidth as more ligand is added. A plot of linewidth vs. ligand/macromolecule mole ratio allows k_{off} to be determined, as illustrated for example in a study of the ^{31}P linewidths for the 2'-phosphate of $NADP^+$ binding to dihydrofolate reductase (84).

Although the determination of off rates is of significance in assessing the stability of the complex, the major interest in more recent studies of complexes in the slow-exchange limit has centered on determining the complete geometry of the complex through the use of intra- and intermolecular NOEs. A recent, but already classic, example of this approach is illustrated by the binding of immunosuppressant peptides such as cyclosporin and FK506 to their receptors. These types of examples are discussed in more detail in Section 3.2.4.2.

3.2.2.2 Fast Exchange. When exchange between free and bound states is very fast, observed NMR parameters are a simple weighted average of those from the two contributing states, illustrated by Equation 12.7 for chemical shifts and Equation 12.8 for linewidths.

$$\nu_{obs} = p_F\nu_F + p_B\nu_B \qquad (12.7)$$

$$1/T_{2,obs} = p_F/T_{2,F} + p_B/T_{2,B} \qquad (12.8)$$

These equations show that, in the fast-exchange limit, addition of a ligand to a protein solution will cause a progressive change in chemical shift. Signals from the protein initially reflect the free state, but as ligand is added the population of bound protein increases and the observed signals move toward those of the bound state. Similarly, when ligand signals are first detected they reflect predominantly the bound state, but with increasing amounts of ligand they move toward the chemical shift of the free state. By regression analysis to Equation 12.7, taking into account the dependency of the mole fractions on K_D by the standard quadratic binding equation, it is possible to obtain estimates of both K_D and the bound shift. The procedure works best for rather weakly binding ligands (e.g., millimolar dissociation constants) (85).

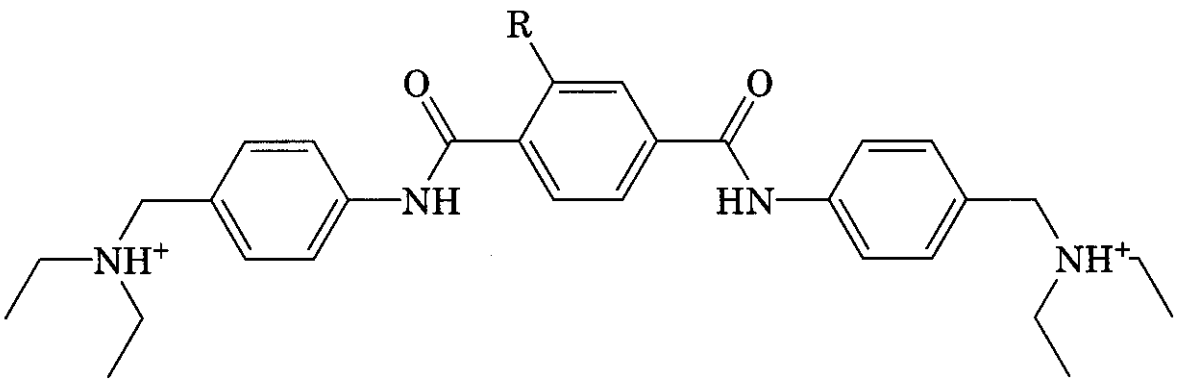
When exchange is somewhat slower, but still within the fast-exchange limit, there is an exchange contribution to linewidth, as shown in Equation 12.9:

$$1/T_{2,obs} = p_F/T_{2,F} + p_B/T_{2,B} + \{[p_Bp_F{}^2 4\pi^2(\nu_B - \nu_F)^2]/k_{off}\} \qquad (12.9)$$

In this case a maximum in the broadening of ligand or protein peaks occurs during the titration at a mole ratio of approximately 0.3, as illustrated below.

The spectral changes that occur in the fast-exchange regime can conveniently be illustrated by studies on the binding of a series of terephthalamide ligands to an oligonucleotide model of DNA. The ligands, referred to as $L(NO_2)$, $L(NH_2)$, and L(Gly), were synthesized as precursors for potential anticancer agents (86). To establish whether they bind in the minor groove of AT-rich DNA, a series of NMR titration experiments was undertaken.

Figure 12.16 shows an expansion of the aliphatic region of a series of ^1H-NMR spectra of 0.5 m*M* of the oligonucleotide d(GGTAATTACC)$_2$, to which increasing amounts of $L(NH_2)$ were added (86), of which the spectra cover mole ratios of ligand to DNA duplex



(6) L(NH_2) R = NH_2
(7) L(NO_2) R = NO_2
(8) L(Gly) R = $NHCOCH_2NH_2$

ranging from 0:1 to 2.6:1. Although the spectra are complicated by overlap in some regions, it is clear that addition of the ligand causes significant changes to the DNA peaks. A typical example is seen for the T6 methyl peak, for which addition of ligand causes both an upfield shift and broadening of the peak at certain stages of the titration. The chemical shift moves monotonically with ligand concentration up to a mole ratio of 1:1 and then reaches a plateau, remaining constant as larger amounts of ligand are added. Broadening of the peak reaches a maximum at a ligand: DNA mole ratio of approximately 0.3. Both observations are consistent with there being moderately fast exchange on the chemical-shift timescale between the free and ligand-bound forms of the DNA in solution. In this case, the observed spectral peaks reflect neither the free nor the bound form of DNA, but are averaged signals.

Ligand peaks are also in fast exchange, as seen with the L(NH,) methyl peak, which first appears at a ligand:DNA ratio of 1.36:1 as a shoulder on the overlapped T3 and T7 methyl peaks at approximately 1.27 ppm. This peak is not initially visible in spectra at low ligand: DNA mole ratios because of the small population of bound species and the overlapping DNA peaks. It moves upfield with increasing ligand concentration and, again, represents an averaged peak intermediate in chemical shift between free and bound forms, reflecting fast-exchange kinetics. Eventually, the chemical shift of this signal approaches that of the free ligand at 1.1 ppm, measured in a separate experiment with a solution of ligand alone.

In the fast-exchange cases such as this it is possible to obtain an estimate of the dissociation constant for the complex (K_D) and the bound chemical shift (ν_B) of DNA resonances by fitting the observed chemical shift changes as a function of ligand concentration to equation 12.7 (85). The parameters that best fit the experimental data for the T6 methyl peak were $K_D = 1.2 \times 10^{-6}$ M and ($\nu_B - \nu_F$) = 46 Hz. Limitations on the accuracy of K_D values derived in this way were described previously (85).

To further define the thermodynamic constants associated with binding, the linewidth data were also quantitatively examined by use of Equations 12.6 and 12.9. In the case of moderately fast exchange, a maximum linewidth is predicted at a ligand:DNA mole ratio of 0.33 (82, 85), and this was indeed observed in the current case. Derived binding parameters were $K_D \leq 1.0 \times 10^{-6}$ M, $k_{off} \approx 250\ s^{-1}$, ($\nu_B - \nu_F$) = 49 Hz, and LW_B = 12 Hz, consistent with the values derived from the analysis of chemical shifts. Subsequent studies with the related ligand L(NO,) showed similar binding to L(NH_2). However, a third ligand, L(Gly), was found to bind somewhat more tightly, with some signals in the intermediate exchange regime.

3.2.2.3 Intermediate Exchange. In this regime the rate of exchange between bound and free states is comparable to the differences in NMR parameters associated with the exchange. In general the spectral peaks often become very broad and analysis is difficult. This is the case, for example, for L(Gly). In the methyl region of the spectra shown in Fig.

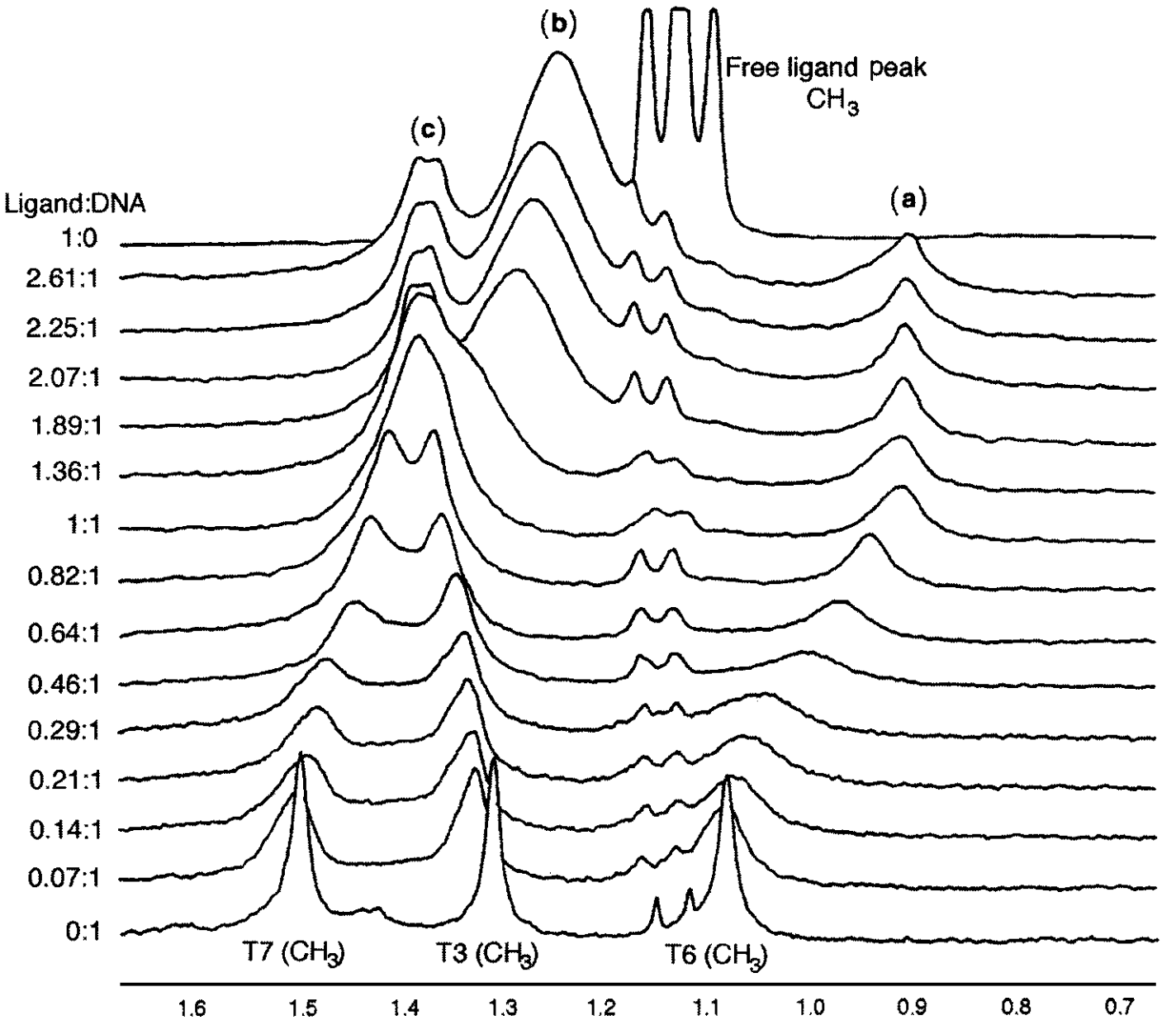


Figure 12.16. Expanded regions from 300-MHz ^{1}H-NMR spectra for complexes between L(NH_2) and d(GGTAATTACC)$_2$ recorded at 10°C. The two small peaks at 1.12 and 1.14 ppm arise from an impurity. Increasing ligand concentration causes an upfield shift of the T6 methyl resonance (a), and causes the T7 and T3 resonances to become overlapped at later stages of the titration (c). Peak (b) is an averaged resonance from the ligand methyl groups intermediate in shift between the bound and free forms of the ligand. (Reprinted with permission from Ref. 86.)

12.17, the T7 CH_3 signal moves upfield and the T3 CH_3 signal moves slightly downfield with increasing ligand concentration, as seen previously for L(NH_2) and L(NO_2). However, in contrast to the case for the other ligands, the characteristic broadening of peaks at intermediate ratios is non-Lorentzian, suggesting kinetics in the intermediate exchange regime. The T6 CH_3 peak does not shift in the characteristic fast-exchange manner but, instead, a new broad resonance appears close to the expected position of the bound T6 CH_3 chemical shift on the first addition of ligand, and increases in intensity with increasing ligand concentration. This observation is consistent with the ligand being in slow to intermediate exchange between the free and bound forms, with $k_{off} \approx (\nu_B - \nu_F)$. Based on the magnitude of $\nu_B - \nu_F$ for this resonance, k_{off} for L(Gly) is estimated to be 50 s^{-1}, which is significantly slower than that for L(NO_2) and L(NH_2).

At a ligand:DNA ratio of approximately 1:1, the ratio of the integrals of the T6 methyl peak and the overlapped T3 and T7 methyl peaks is about 1:6. The expected value is 1:2, which indicates that the bound ligand methyl peak (4 $\times$ CH_3) is overlapped with the T7 and T3 methyl peaks, as observed with L(NH_2) and L(NO_2). When the ligand:DNA ratio is increased beyond a 1:1 ratio, a new peak appears at about 1.15 ppm and increases in intensity as

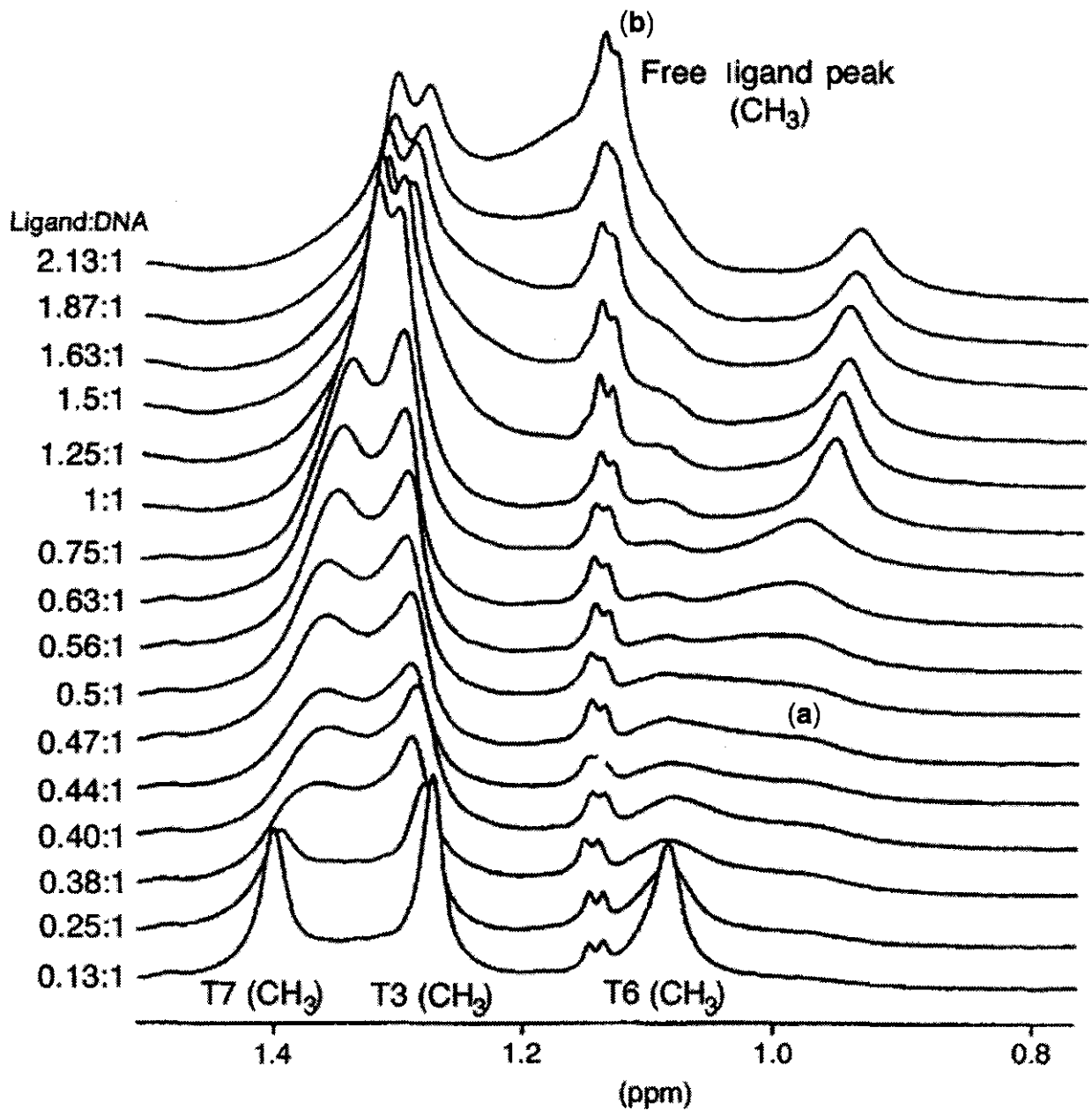


Figure 12.17. Expansions from the 600-MHz ^{1}H-NMR spectra for complexes formed between L(Gly) and d(GGTAATTACC)$_2$ showing the methyl resonances. The two small peaks at 1.12 and 1.14 ppm are attributed to an impurity. The complex nature of the T6 methyl resonance at ligand:DNA ratios less than 1:1 (a), and the manner in which signal intensity increases at about 1.15 ppm at DNA:ligand ratios greater than 1:1 (b), are indicative of intermediate exchange. (Reprinted with permission from Ref. 86.)

the ligand concentration is increased. This new peak corresponds to the methyl peak of the free ligand and its appearance in this manner is consistent with slow exchange on the chemical-shift timescale. To confirm this, spectra of a 2:1 mixture of L(Gly) and d(GGTAATTACC), were acquired at different temperatures (86), as illustrated in Fig. 12.18.

At low temperatures, signals at 1.15 and 1.30 ppm (overlapped with the T7 CH, and T3 CH, peaks) attributable to the methyl groups from the free ligand and bound ligand, respectively, are distinguishable. As the temperature is increased, a broad peak appears between these two signals (at −1.22 ppm). At the lower temperatures $k_{off} \leq (\nu_B - \nu_F)$, so that methyl resonances of the ligand have complex characteristics reflecting slow-intermediate exchange. At higher temperatures, $k_{,,} \geq (\nu_B - \nu_F)$, so the signal appears as a fast-exchanged average between the free and bound resonances. From a qualitative analysis of the spectra, k_{off} for L(Gly) was estimated to be 50–60 s^{-1} at 283 K.

The fact that some peaks (e.g., the oligonucleotide T7 and T3 methyl signals) exhibit fast exchange, whereas others in the same spectrum of the same complex exhibit slow-intermediate characteristics, is a reflection of the different $(\nu_B - \nu_F)$ values for different peaks. This emphasizes the point made earlier that the "exchange regime" is a relative expression

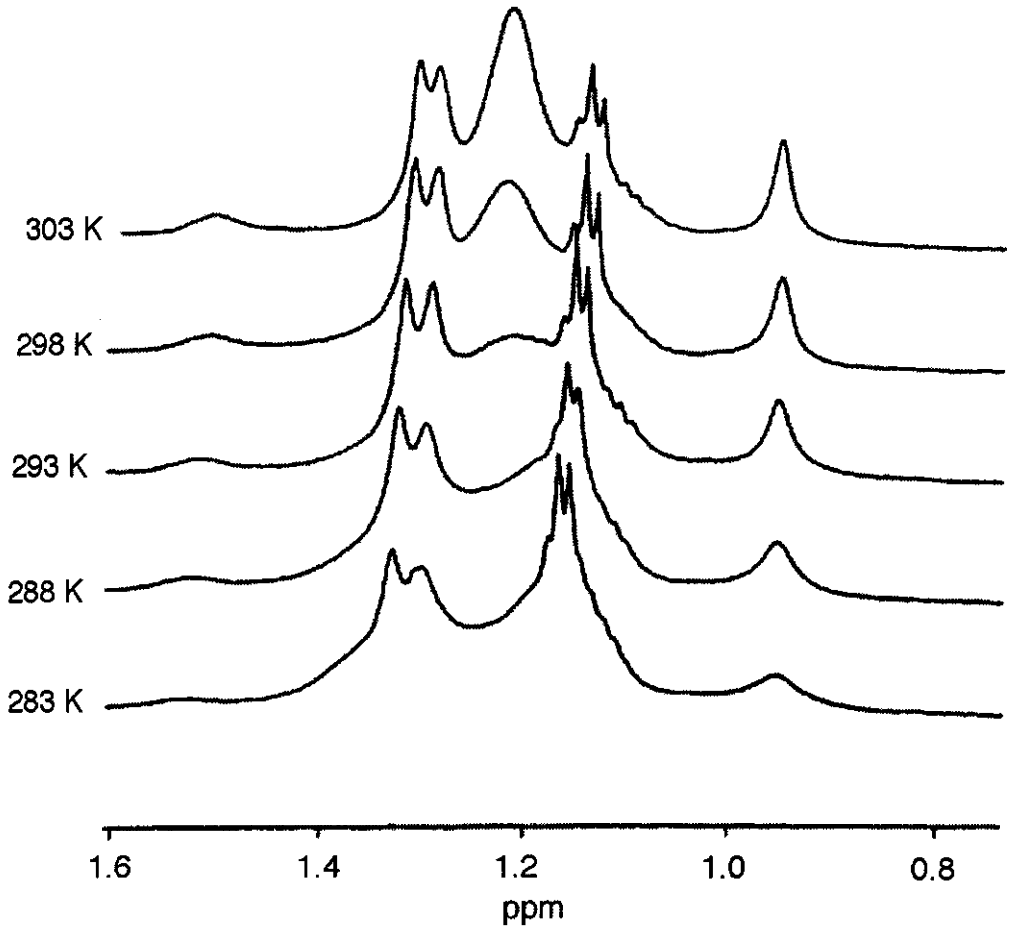


Figure 12.18. Expansions of 'H-NMR spectra of a 2:1 mixture of L(Gly) and d(GGTAATTACC)$_2$ acquired at different temperatures. (Reprinted with permission from Ref. 86.)

and depends not only on the rate of exchange, but also the size of the chemical shift differences involved. In summary, the observations suggest that k_{off} for binding of L(Gly) to the oligonucleotide duplex is much slower than that for the other two derivatives. This provides an illustration of the value of NMR as a quick method for comparing the binding of different ligands and for confirming ligand-binding hypotheses.

The change in binding kinetics may be rationalized by considering the different structure of the L(Gly) ligand relative to the other ligands. It was anticipated that, upon binding to the minor groove, the terephthalamides would adopt a conformation in which the substituent on the central ring would form part of the convex edge of the ligands and therefore be directed toward the "mouth" of the groove. Given this binding arrangement, the ligand L(Gly) would have a positively charged alkylamine group positioned to interact with the negatively charged phosphate groups of the DNA backbone. The L(Gly) derivative also has a bulkier substituent than that of the other ligands and this is also consistent with some differences in its binding.

3.2.3 NMR Techniques. NMR is a particularly versatile tool for the analysis of protein-ligand interactions. As well as being able to observe different nuclei, measurements may be made of a range of different NMR parameters, including chemical shifts, linewidths, coupling constants, and relaxation parameters. In addition, there are several specific NMR techniques that have been applied for the measurement of these parameters. The, techniques that are particularly valuable for the study of macromolecule-ligand interactions are described in the following sections.

3.2.3.1 Chemical-Shift Mapping. Chemical shifts are exquisitely sensitive markers of the local charge state and environment. Although it is not possible to construct an accurate model of a binding site from a knowledge of the chemical shifts of a bound ligand, a qualitative interpretation of *changes* in chemical shifts of the macromolecule on binding provides significant insight into the location of the binding site. Traditionally, such studies were done using 1D NMR but are now increasingly done by 2D HSQC spectra. By simultaneously obtaining information on chemical shifts for a large number of sites in a macromolecule and seeing which ones change when a ligand binds, and which ones do not, it is possible to deduce the location of the binding site. This procedure is referred to as chemical-shift mapping. A prerequisite of the approach

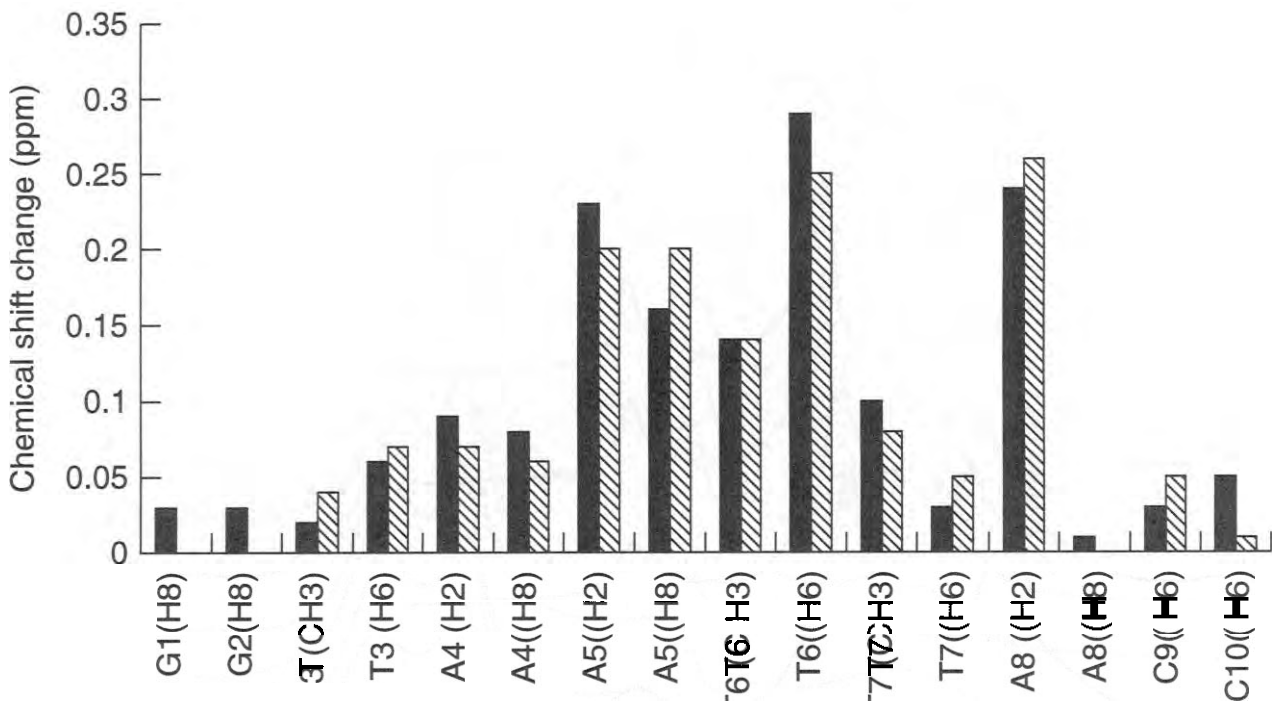


Figure 12.19. Chemical-shift perturbations of DNA protons upon ligand binding. The lighter and darker columns represent shifts attributed to L(NO,) and L(NH_2) derivatives, respectively.

is that the chemical shifts have been assigned. Chemical-shift mapping by use of HSQC spectra is widely used in NMR screening approaches and we will defer a more detailed discussion on it until Section 4.

The relative simplicity of how chemical shift information localizes binding sites may be illustrated by continuing with the example introduced above of terephthalamide binding to DNA. Figure 12.19 shows that, upon binding of the terephthalamides L(NH_2) and L(NO_2) to d($GGTAATTACC)_2$, the DNA protons on the four base pairs between A5 and A8 are perturbed to a much larger degree than protons in the rest of the sequence. It is thus likely that these four residues form the binding site.

A more detailed analysis allows the binding site to be further localized to the minor, rather than to the major, groove in the region of these bases. A4, A5, and A8 are the only residues containing easily detectable minor groove protons (H2). These resonances, which originate from the floor of the minor groove, are shifted downfield with ligand binding, whereas most other resonances are shifted upfield. This observation is consistent with the ligands binding in the minor groove and has been reported for other minor groove binders such as Hoechst 33258 (87) and SN-6999 (88, 89), where adenine H2 protons on the floor of the groove experience deshielding ring current effects. However, significant chemical shift changes were also observed for some major groove protons. This illustrates the general point that sometimes allosteric effects can cause changes at sites not directly involved in binding. In the case of DNA, binding perturbations in the major groove have also been observed for other established minor groove binders such as distarnycin (90), netropsin (91), and Hoechst 33258 (92). Based on NOE and crystallographic data, it was concluded that the effects were caused by distortions of the B-DNA duplex, including changes in the "base roll" of residues within the binding site, upon complexation. Electronic effects arising from the close proximity of charged groups on the ligand to neighboring nucleotides were also found to perturb major groove resonances.

In the case of the terephthalamides a comparison of the minor and major groove perturbations for a particular residue shows that the minor groove protons are affected to a much greater extent. This is particularly evident for A8, where the H2 proton shifts by approximately 0.25 ppm and the H8 proton is not affected (Fig. 12.19). It is difficult to conceive of a binding mode in the major groove that would account for such a large effect on the minor groove A8 H2 resonance without a simultaneous effect on the major groove protons of T7 and A8. The observed 1:1 stoichiometry of the

complex excludes the possibility that the ligand binds to the major and minor groove at the same time. It is more likely, therefore, that binding in the minor groove causes distortion of the DNA structure so that perturbations are observed for the major groove protons of A5 and T6, but not neighboring nucleotides.

Other examples of the use of chemical-shift mapping to locate binding sites have been made for ligands binding to a range of drug targets, including immunophilins, matrix metalloproteases, and DHFR. Some of these examples are described in more detail in section 3.2.5.

3.2.3.2 NMR Titrations. There are a number of advantages in undertaking a titration of ligand against macromolecule or vice versa rather than just examining the final complex. These include introducing the possibility of distinguishing signals from the individual components on the basis of intensities at intermediate stages of the titration in the slow-exchange case and obtaining kinetic and thermodynamic parameters associated with the interaction in the fast-exchange case. Such titrations may be done using either 1D or 2D spectra and are very useful for establishing the exchange regime of the complex, as described in Section 2.1. A variety of parameters may be monitored in the titration, although the two most common are chemical shifts and linewidths. Examples of such titrations are given in Figs. 12.16 and 12.17.

3.2.3.3 Isotope Editing and Filtering. Isotope editing provides a powerful way of distinguishing between the components in a complex without the need for a titration. It is one of the most useful tools for the study of macromolecule-ligand complexes, and indeed the background NMR technology that underpins isotope editing was developed specifically for the study of complexes. The principle of the approach is illustrated in Fig. 12.20 and is based on the use of isotopes to select for signals from either the ligand or macromolecule, or signals exclusively linking both of them.

The conceptually simplest approach is to uniformly deuterate the macromolecule, thereby removing its signals from ^{1}H-detected NMR spectra, and allowing signals from only the ligand to be observed. This substantially simplifies the spectrum and allows, for example, the bound conformation of a ligand to be determined from NOESY data recorded in D_2O. By rerunning the spectrum in H_2O, additional NOEs to exchangeable amide protons on the protein may be detected, thereby providing information on contacts between ligand and protein. Alternatively, ^{15}N or ^{13}C signals may be introduced selectively into either the ligand or protein and editing techniques used to select only signals attached to these labels and their proximate protons. This was used in the first example of an isotope-edited study, in this case to examine the binding of a ^{15}N-labeled peptide-based inhibitor to pepsin (93).

Potentially, the most useful approach involves uniform labeling of one of the components with either ^{15}N or ^{13}C and leaving the other component unlabeled. It is then possible to edit the spectrum by selecting for interactions (either through bond or through space) that connect protons that are both one-bond coupled to ^{15}N or ^{13}C. Alternatively, the spectrum may be filtered to specifically remove such signals, thereby selecting only signals involving protons coupled to ^{14}N or ^{12}C (i.e., on the unlabeled component). It is generally easier to uniformly label the protein rather than the ligand, and editing methods are highly efficient, thus making it easy to visualize just the protein. However, because ligand signals , are often of interest, filtering experiments play a valuable role in visualizing them. Unfortunately, filtering experiments are more susceptible to artifacts than are editing experiments, although there have been recent advances in reducing artifacts (94).

Another possibility is to use half-edited/half-filtered 2D experiments to detect NOEs that specifically involve interactions between protons attached to ^{15}N or ^{13}C and those that are not. This approach is used, for example, to detect intermolecular NOEs between a labeled protein and an unlabeled ligand. Examples of isotope editing/filtering are given in section 3.2.4.

3.2.3.4 NOE Docking. In many cases the study of a complex may follow a previous structure determination of the isolated macromolecule and in that case it may be possible to determine much information about a complex by obtaining a relatively small number of NOEs linking the ligand and macromolecule.

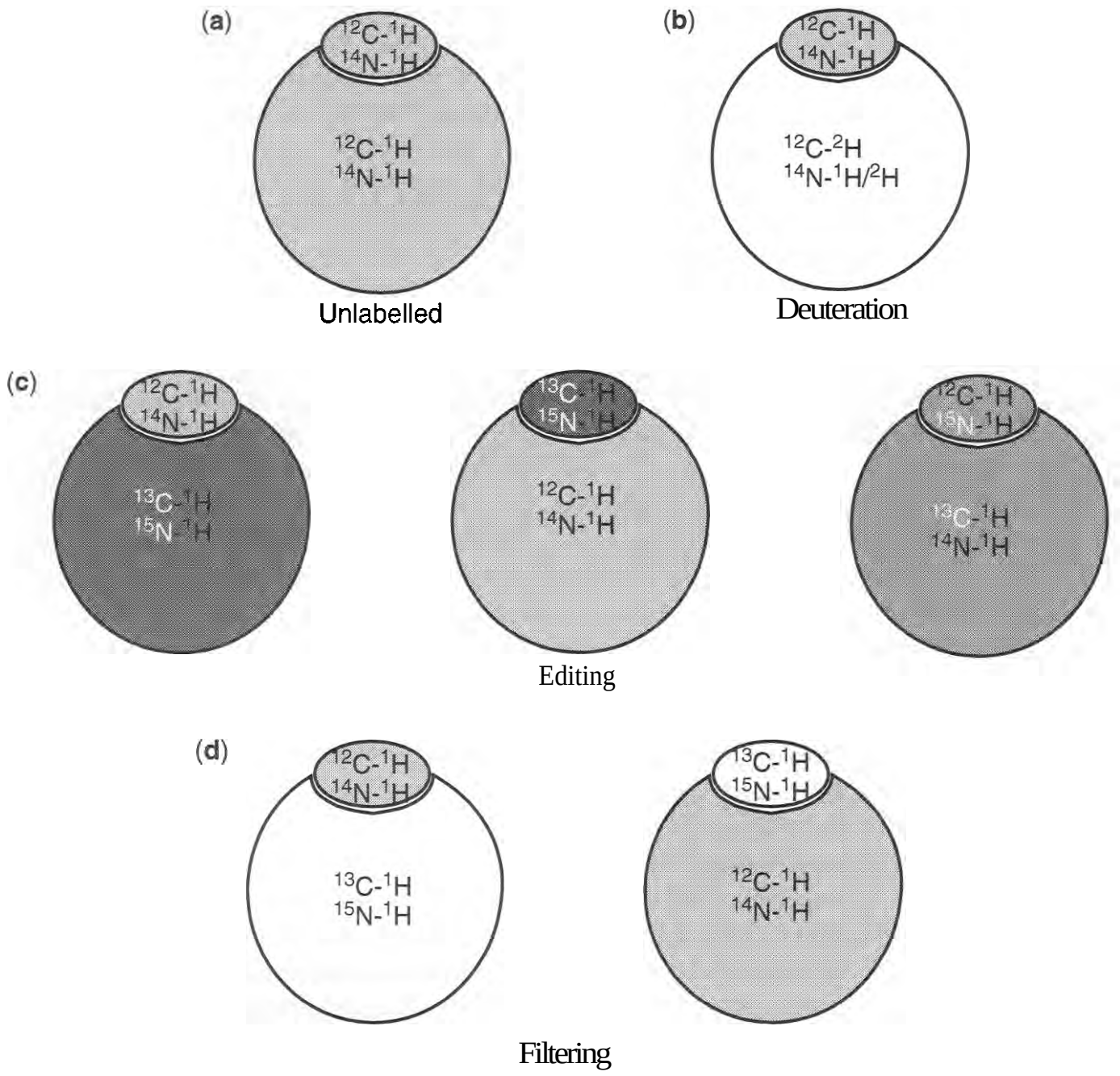
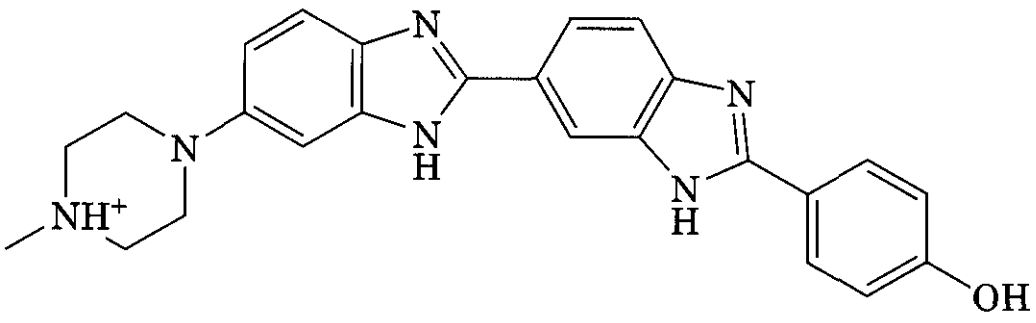


Figure 12.20. Isotope editing and filtering can be used to select signals from either the ligand or the protein. (a) Normal protein and ligand with no filtering or editing. (b) Selection of the ligand signals by 2H labeling of the protein. (c) Selection of protein or ligand signals by ^{13}C and/or ^{15}N labeling/ editing. (d) Removal of protein or ligand signals by ^{13}C or ^{15}N filtering.

Gradwell and Feeney **(95)** recently analyzed factors important in such NOE docking experiments. In their analysis, a high resolution X-ray structure of a protein-ligand complex was used to simulate loose distance restraints of varying degrees of quality that might typically be estimated from experimental NOE intensities. These simulated data were used to examine the effect of the number, distribution, and representation of the experimental constraints on the precision and accuracy of the calculated structures. **A** standard simulated annealing protocol was used, as well as a more novel method based on rigid-body dynamics. The results showed some parallels with those from similar studies on complete protein NMR structure determinations, but it was found that more constraints per torsion angle are required to define docked structures of similar quality. This is because the conformation and orientation of the ligand are defined only by NOEs and not by covalent attachment, as is the case for amino acid side chains in a protein structure. The effectiveness of different NOE-constraint averaging methods was explored and the benefits of using "R^{-6} averaging" rather than "center averaging" with small sets of NOE constraints were demonstrated. With these considerations in mind it appears that NOE docking can be a very cost-efficient procedure for defining the environment, orientation, and conformation of ligands.



(9)

3.2.4 **Selected Examples.** Applications of the various NMR techniaues described are now illustrated with selected examples. The examples have been chosen to give a broad perspective on the types of NMR experiments that can be done and the types of information they provide. Specifically, the first example covers the case of drug-nucleic acid binding and focuses on more traditional NMR experiments, involving relatively standard homonuclear methods. The second example covers binding of moderately large ligands to immunophilins and highlights modern isotope editing techniques. The third example, covering ligand binding to a matrix metalloproteinase, also highlights the importance of these techniques and shows how relatively simple spectra involving ^{19}F-containing ligands can be very informative. The fourth example describes ligand binding to DHFR, one of the most extensively studied systems by NMR, and illustrates the derivation of a broad range of kinetic and geometric information on intermolecular complexes. The final example, on HIV protease, describes how NMR complements X-ray studies and provides information on dynamic motions within complexes.

3.2.4.1 DNA-Binding Drugs. The NMR approaches that have been used to examine the interactions of minor groove binding drugs with DNA can be illustrated with studies on the bisbenzimidazole-based compound, Hoechst 33258, (9).It has been used widely as a fluorescent cytological DNA stain and is also active as an anthelmintic agent. It has activity against intraperitoneally implanted L1210 and P388 leukemias in mice (96).

Footprinting studies (96) have shown that sequences of four AT base pairs are a prerequisite for strong binding to DNA, consistent with similar observations for other structurally related molecules such as distarnycin and netropsin (87, 90, 91, 97–99). The first structural studies of Hoechst 33258 complexed to short sequences of synthetic oligonucleotides were done using X-ray crystallographic methods (100–102). NMR and further X-ray studies followed (92, 103–107). Three of the X-ray studies (100, 101, 103) used the *Eco*RI sequence d(CGCGAATTCGCG)$_2$ and another (102) used the sequence d(CGCGATATCGCG),. Both sequences fulfil the requirement of at least four consecutive AT base pairs, and the resulting complexes showed similar modes of binding. In all of the X-ray studies, the Hoechst ligand was found to bind to the minor groove.

The NMR studies of complexes between Hoechst 33258 and oligonucleotide sequences provided complementary information to the crystal structure data (92, 103–106). Because the binding is reversible, the NMR data offer the opportunity to derive information about the kinetics of the interaction. As with the crystallographic studies, the oligonucleotide sequences were designed to contain runs of AT base pairs. Some NMR studies were performed with dodecanucleotide sequences used in crystallographic studies, including d(CGCGAATTCGCG)„ which allowed a direct comparison with the crystallographic data. Experiments were also performed with sequences specifically designed to investigate different aspects of the interaction. The sequence d(CTTTTGCAAAAG)$_2$ was designed to offer two binding sites, and it was shown that two Hoechst molecules interacted with the DNA duplex in symmetry-related orientations at the 5'-TTTT-3' and 5'-AAAA-3' sites (92).

3.2.4.1.1 Stoichiometry and Kinetics. The starting point in studies of ligand-DNA complexes is usually a titration experiment to es-

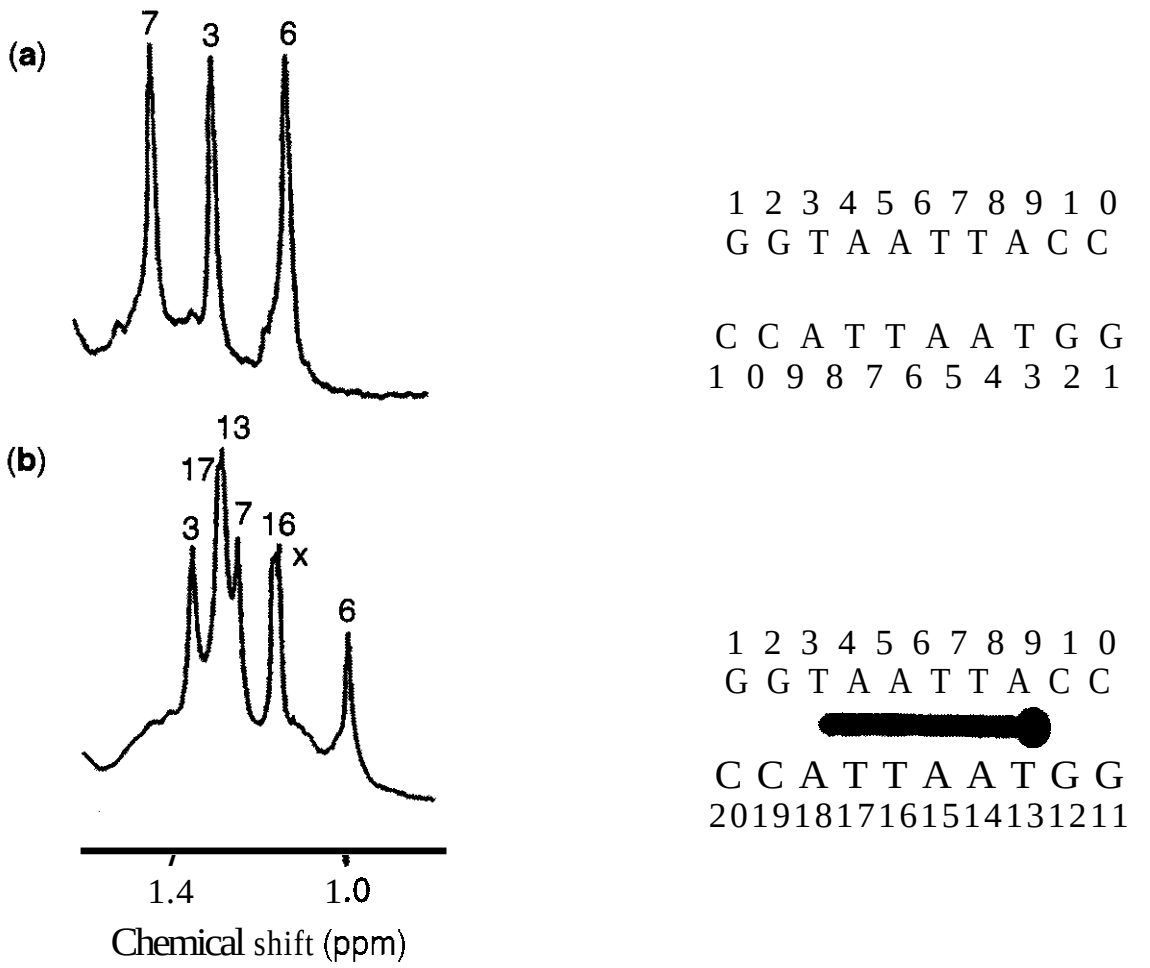


Figure 12.21. 1D ^{1}H-NMR spectra (recorded at 20°C) illustrating the thymine methyl region for the symmetrical ligand-free duplex (a) and for the 1:1 Hoechst:d(GGTAATTACC)$_2$ complex (b), which is no longer symmetrical because of the ligand binding. x corresponds to a small impurity peak. The DNA strands are numbered to the right of the spectra and the approximate location of the ligand is indicated by a black bar. (Reprinted with permission from Ref. 105. Copyright 1993, Blackwell Publishing Science.)

tablish the nature and stoichiometry of the complex. Complexes between the ligand and DNA duplex are obtained by adding small aliquots of ligand solution to a sample of the DNA duplex with one-dimensional ^{1}H NMR spectra acquired after each addition. The effects observed on the NMR spectrum after each addition reveal whether an interaction is taking place and allow the interaction to be characterized as fast or slow exchange on the NMR timescale. The stoichiometry of the interaction can also be determined from the titration.

In general, the addition of Hoechst 33258 to the oligonucleotide duplexes causes a decrease in the intensity of free DNA resonances and a concomitant increase in the intensity of new resonances, which appear in previously unoccupied spectral regions. This is consistent with the free and bound forms of the DNA duplex being in slow exchange with each other. For example, when Hoechst 33258 is added to d(GGTAATTACC)$_2$, the free DNA signals completely disappear at a DNA:drug ratio of 1:1, and the number of new resonances is twice the number of previously observed free DNA resonances (Fig. 12.21). This is a common feature of complexes with 1:1 stoichiometry and reflects a loss of the dyad symmetry of the duplex attributed to ligand binding.

Upon addition of Hoechst 33258 to d(CTTTTCGAAAAG)$_2$, the free DNA signals completely disappeared at a ratio of 2:1 drug: DNA and there was no doubling of the number of DNA resonances in the spectrum (92). From this, it could be concluded that two molecules were bound per duplex in a manner that retained the dyad symmetry of the DNA duplex. The binding was also determined to be cooperative, in that no intermediate 1:1 complex was detected (92). The formation of a 1:1 complex would have resulted in a very complicated spectrum at intermediate ligand:DNA ratios, given that resonances arising from the free

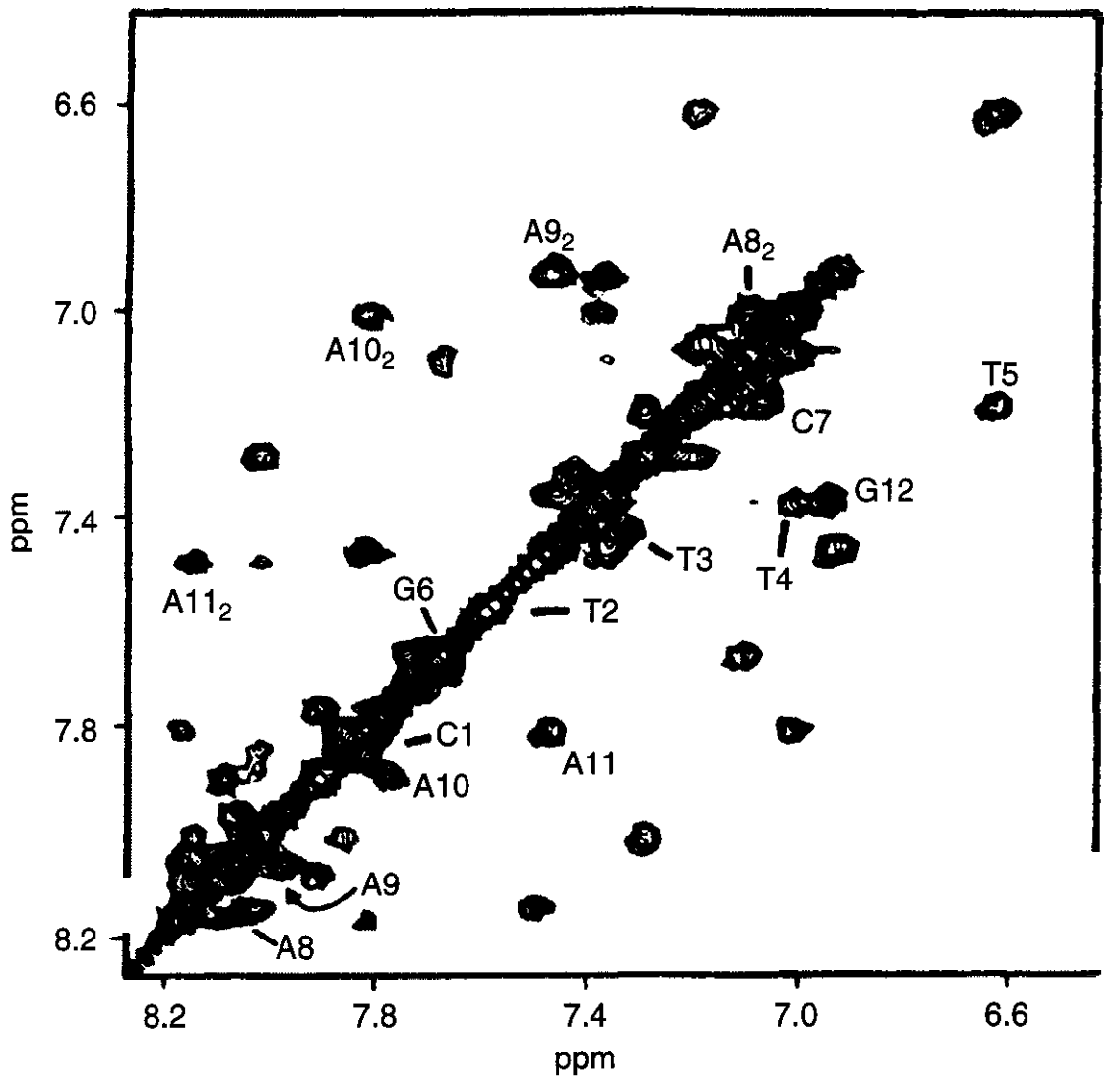


Figure 12.22. Aromatic region of NOESY spectrum of a 1:1 mixture of Hoechst vs. d(CTTTTCGAAAAG)$_2$ recorded with a 200-ms mixing time. Chemical exchange cross peaks between protons of the free DNA and the 2:1 Hoechst:DNA complex are labeled with their identifying base pair. Below the diagonal the H6 and H8 cross peaks are shown, whereas those of the adenine H2 resonances are highlighted in the upper portion of the figure and labeled with a subscript 2. (Reprinted with permission from Ref. 92. Copyright 1990, Oxford University Press.)

DNA, the 1:1 complex and the 2:1 complex, would have produced four times as many observable peaks relative to the free DNA species. At intermediate ligand concentration, however, only two sets of peaks arising from DNA molecules were detected. In the 2:1 complex only four thymine methyl resonances were detected (1.0–1.5 ppm), as expected for a symmetrical DNA duplex. These are all overlapped in the free DNA spectrum. In the 1:1 mixture, only signals from free DNA and from the 2:1 complex were detected.

The reversible nature of the Hoechst:DNA interaction is illustrated by the observation of chemical-exchange cross peaks in NOESY spectra of mixtures of free and complexed oligonucleotides (92, 104). This may be seen in the NOESY spectrum of a mixture of free and complexed d(CTTTTCGAAAAG)$_2$, shown in Fig. 12.22, in which many chemical exchange cross peaks are observed between resonances arising from the free and bound oligonucleotide. In a NOESY spectrum acquired at lower temperature, the intensity of these chemical-exchange cross peaks is significantly reduced, indicating that the exchange is slowed at lower temperatures. The exchange rate was estimated to be $<10\ s^{-1}$ at 10°C (92).

The ability to observe such dynamic exchange phenomena is one of the strengths of NMR relative to X-ray crystallography and several examples of these phenomena are described later in the chapter.

3.2.4.1.2 Binding Site. A combination of chemical shift and NOE information can be used to locate and characterize binding sites. Chemical-shift differences between resonances arising from free and bound forms of DNA are indicative of the nature of the interaction. In all studies of the Hoechst complexes described above (92, 104–107) significant changes to the chemical shifts of thymine H1′ protons and adenine H2 protons were observed, in contrast to the generally small perturbations observed for the base H8/H6 and CH, resonances located in the major groove. perturbations of this nature are consistent with binding to the minor groove. In some instances, significant perturbations were observed to major groove protons located well

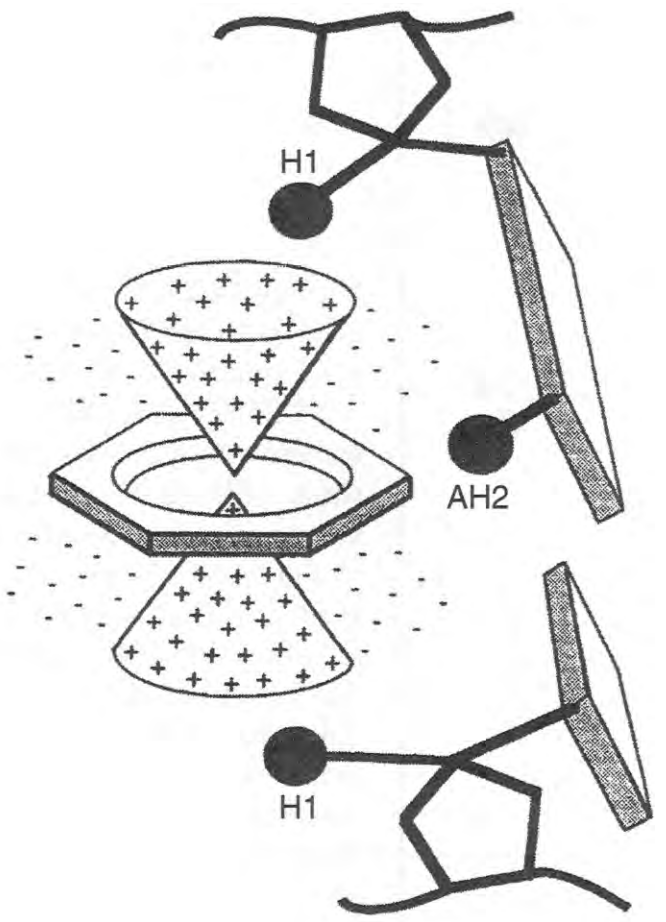


Figure 12.23. Schematic representation of ligand-induced ring-current effects on nucleotide protons that form the walls (deoxyribose H1′) and floor (adenine H2) of the minor groove. (+), shielding effects; (−), deshielding effects. (Reprinted with permission from Ref. 105. Copyright 1993. Blackwell Publishing Science.)

within the binding site, reflecting changes in the conformation of the DNA duplex (e.g., base roll, propeller twisting) (92, 106).

Further evidence of minor groove binding is provided by the fact that resonances arising from protons on the floor of the groove, such as the adenine H2 and imino resonances, are shifted downfield, whereas resonances from protons on the minor groove walls, such as the H1′ protons, are shifted upfield. This is a consequence of the ligand's being inserted edge-on into the minor groove. The deoxyribose protons that form the walls of the minor groove are positioned above the π-plane of the aromatic rings and consequently receive upfield perturbations to their chemical shifts. Protons positioned on the floor of the groove, however, generally lie in the plane of the aromatic rings and experience downfield perturbations to their chemical shifts, as illustrated in Fig. 12.23.

The magnitude of chemical shift changes is a strong indicator of the location of the binding site. In the case of the 2:1 complex with d(CTTTTCGAAAAG)$_2$ (92), the largest chemical shift changes occur over the 5'-TTTT-3' and 5'-AAAA-3' regions of the duplex. In the case of 1:1 complexes, where the DNA duplex contains an AT base-pair segment located at the center of the sequence, greater chemical shift perturbations are observed for resonances in that region (104–106), consistent with the binding site's being located there.

Assignment of the bound ligand and DNA resonances enables the identification of intermolecular NOEs, which are required for a precise determination of the binding site. The interaction of Hoechst 33258 with the oligonucleotides produced a large number of intermolecular NOEs (~25–30), placing considerable constraints on the structure of the complex and enabling the orientation of the ligand within the binding site to be determined. The NOE contacts observed for different complexes have a few features in common. The contacts generally involve DNA protons associated with the minor groove, such as ribose H1′ and adenine H2, clearly locating Hoechst in the minor groove. Protons of all four spin systems of the ligand show NOEs to protons of the DNA, demonstrating that the interaction occurs along the entire length of the drug. Typically, protons along one edge of the ligand (e.g., NH and H4′/H4″) exhibit close contacts to protons on the floor of the minor groove, showing that the bound drug is crescent shaped and isohelical with the DNA (92, 104–106).

Models of the interaction of Hoechst 33258 with the oligonucleotides studied were generated based on the intermolecular NOEs. The models of the 1:1 complexes indicated that the ligand interacted with the four AT base pairs located at the center of the sequence. Interestingly, there was no evidence for interactions with GC base pairs on the periphery of the binding sites. In the 2:1 complex reported by Searle and Embrey (92), the array of contacts observed located the ligand in the minor groove at the center of the 5'-TTTT-3' and 5'-AAAA-3' sites, as illustrated in Fig. 12.24.

As well as defining the location of the binding site, intermolecular NOEs can be used to determine the orientation of the ligand at that site. In the case of the 2:1 complex, the *N*-

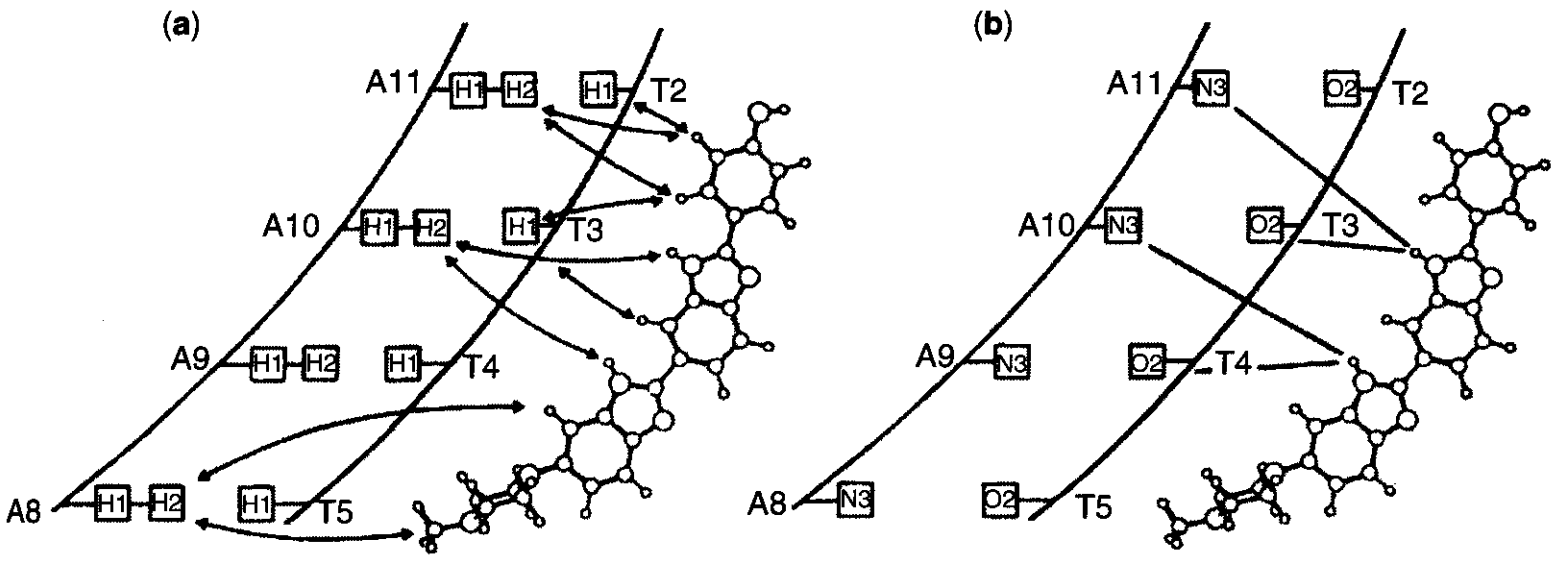


Figure 12.24. Schematic representation of Hoechst 33258 bound to the minor groove of the 5′-TTTT sequence. (a) Highlights of some of the NOEs that determine the position and orientation of the Hoechst molecule within the minor groove. (b) Intermolecular hydrogen-bonding scheme. Molecular-modeling studies with an idealized B-DNA helical structure indicate that the benzimidazole H3′ is capable of forming bifurcated interactions with A11N3 and T302, whereas the benzimidazole H3″ hydrogen bonds in a similar manner, but with A10N3 and T402. In the proposed model of the complex, these distances fall within 3.5 Å and are thus within acceptable hydrogen-bonding limits. (Reprinted with permission from Ref. 92. Copyright 1990 Oxford University Press.)

methylpiperizine moieties were found to point toward the center of the duplex, as indicated by NOEs between the protons from the piperizine ring and the 5'-terminus of the adenine tract (Fig. 12.24). Corresponding NOEs were also observed between the drug phenolic protons and the 5'-terminus of the thymine tract, as well as the 3'-terminus of the adenine tract of the complementary strand. This model did not indicate any interaction with the central GC base pairs (92).

The orientation of the ligand was similarly determined in the 1:1 complexes based on intermolecular NOEs between protons located at the extremities of the Hoechst molecule and protons of the binding site. For example, in the interaction with $d(GTGGAATTCCAC)_2$, Fede et al. (106) reported NOEs between protons from the piperizine moiety and the H2 and H1′ protons of the dinucleotide fragment d(A5T5)·d(A6T6).

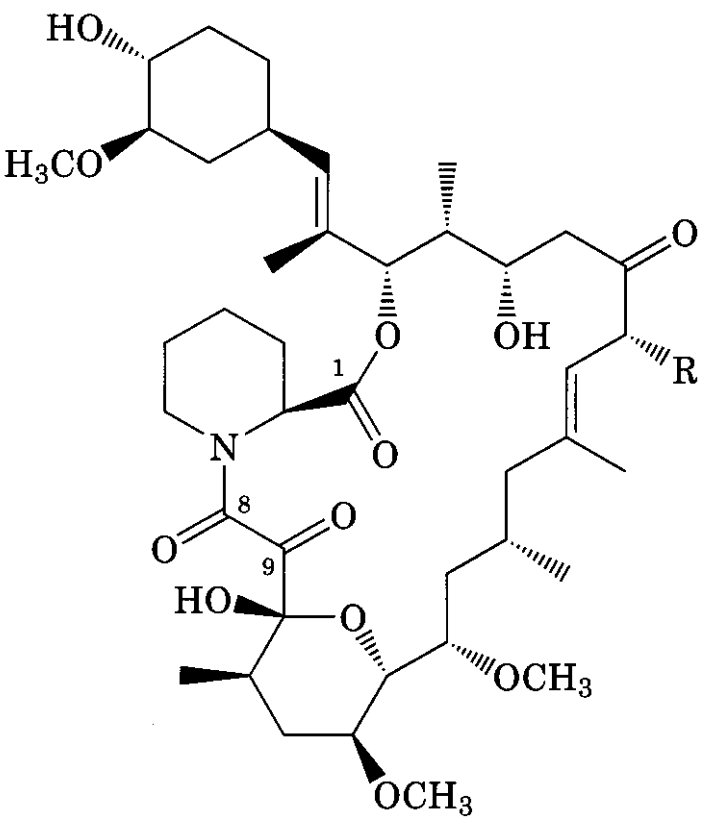
3.2.4.1.3 Dynamic Processes. The binding of the Hoechst molecule to the self-complementary oligonucleotide duplexes in a 1:1 ratio lifts the dyad symmetry of the duplexes so that two sets of DNA resonances are observed. This indicates that the drug is in slow exchange between the free and the bound forms. Close examination of the 2D NOE data, however, reveals the presence of chemical-exchange cross peaks between symmetry-related protons on opposite sides of the dyad axis of the DNA duplex. The mechanism by which this occurs has been described as dissociation of the Hoechst molecule from the duplex, followed by a 180" reorientation and rebinding (105, 106). The self-complementary nature of the sequences ensures that the same complex is formed for either ligand orientation but with the net effect of interchanging the two strands with respect to the orientation of the Hoechst molecule. The rate at which this process occurs was estimated using cross-peak intensities in the NOESY spectrum (106). When interacting with $d(GGTAATTACC)_2$ and $d(GTGGAATTCCAC)_2$, the lifetime of the complex in each state ($1/k_{ex}$) was reported to be approximately 0.8 and 0.45 s, respectively (105, 106). These values indicate a small but significant difference in the affinity of Hoechst for TAATTA and GAATTC sites.

Intramolecular dynamic processes that are fast on the NMR timescale are also observable in the ^{1}H-NMR spectrum of the bound Hoechst molecule. Resonance averaging is observed for the H2/H6 and H3/H5 protons of the phenol group, which is consistent with the environments on either side of the ring being averaged by rapid ring-flipping motions about the C4-C2′ axis. This occurs despite the appar-

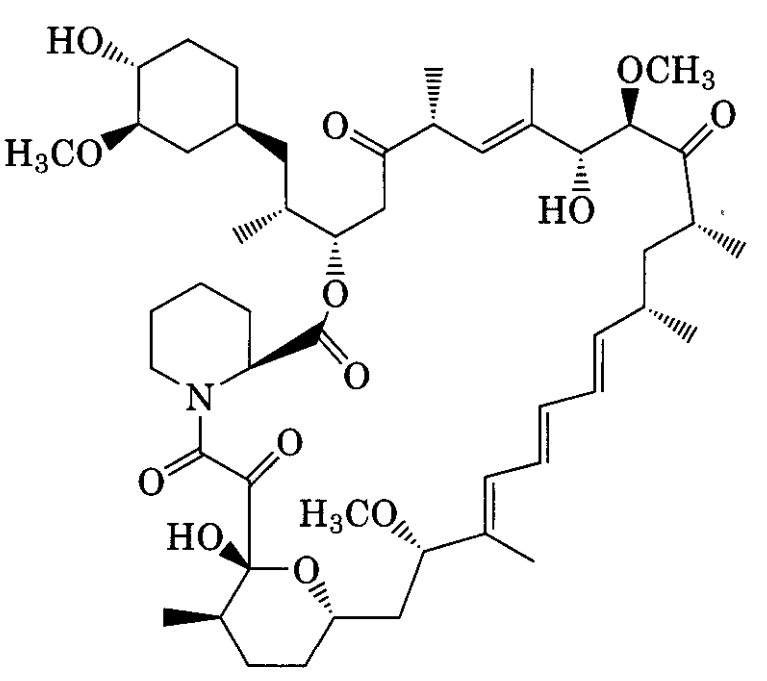
ent tight fit between the phenyl ring and the walls of the minor groove, which, in a static model of the complex, must present a large barrier for rotation. It was estimated (105) that the rate for this process is as high as 1000 s^{-1}. This is much higher than the rate of interconversion between free and bound forms of the duplex; thus, dissociation of the drug from the complex cannot be the rate-limiting factor for phenol ring flipping. Dynamic fluctuations of the DNA conformation are more likely to provide the rate-limiting step.

3.2.4.1.4 Summary of Solution Studies. The data obtained from these NMR studies are consistent with the bound ligand fitting tightly within the minor groove of AT tetramers, with the aromatic rings of the ligand being roughly coplanar. The AT tract provides the key recognition features required for binding, including the narrowness of the minor groove. The importance of van der Waals interactions is evident, given the large number of NOE contacts between the ligand and the walls and floor of the groove. Hydrogen bonding also plays a significant role in stabilizing the interaction, as do electrostatic interactions between the positively charged piperizine ring and the minor groove. Electrostatic interactions are also likely to play a significant role in orienting the ligand within the binding site, as shown in the 2:1 complex, where the piperazine rings point toward the center of the duplex where the positive charge is best stabilized (92). The information derived from these studies, as well as from NMR studies of the interactions of other minor groove binders with DNA, is useful for the design of ligands with altered specificity or increased binding affinity, with the overall goal being the development of novel drugs.

3.2.4.2 Immunophilins: Studies of FK506 Analog Binding to FKBP. Some of the most detailed investigations of the interaction between ligands and their target proteins have been made for the immunophilin class of proteins. The major FK506 binding protein (FKBP) has a molecular mass of about 11.8 kDa, whereas cyclophilin (Cp) has a mass of about 17 kDa. These proteins are unrelated in amino acid sequence but both have peptidyl-prolyl *cis-trans* activities that are inhibited by immunosuppressants that block signal transduction pathways leading to T-lymphocyte activation. FK506 (10) and rapamycin (11) inhibit the *cis-trans* isomerase activity of FKBP, whereas cyclosporin A (structure shown in Fig. 12.3) inhibits that of Cp. NMR has contributed significantly to the understanding of binding interactions to both proteins.



(10) FK506 R = CH_2CHCH_2
(12) Ascomycin R = CH_2CH_3



(11) Rapamycin

Initial studies on FK506 focused on the structure of the free ligand to aid in the design of further analogs (108–110). However, it was established from studies of the cyclosporin A-

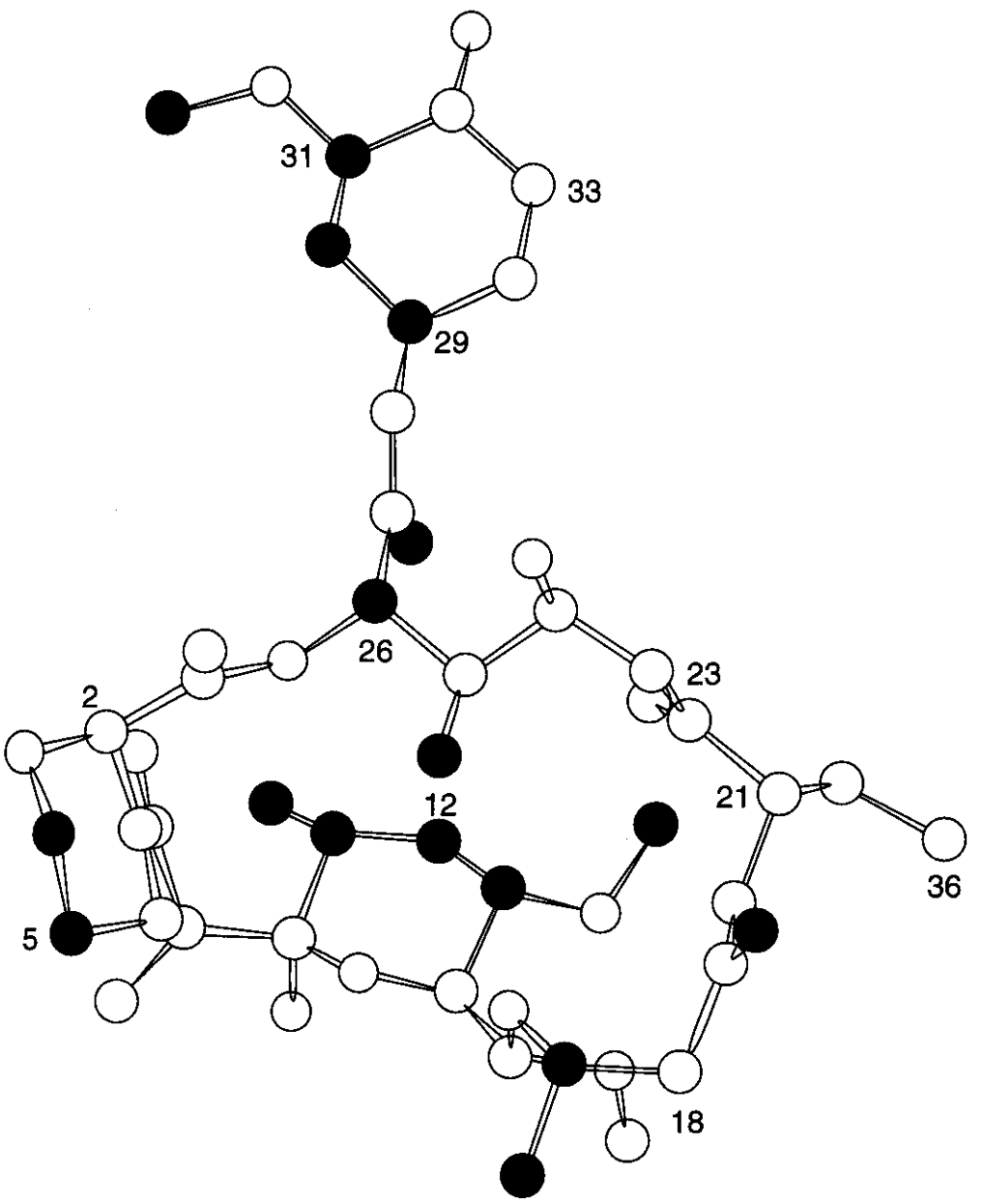


Figure 12.25. Three-dimensional structure of ascomycin bound to FKBP. Protons on the ligand that showed NOEs to the protein are denoted by a black shading of the carbons to which they are attached. Although no NOEs were observed from the protons at position 3 to the protein, the upfield shift of their resonances, −1.09 and 0.25 ppm, suggests that they are in close proximity to an aromatic region of FKBP. (Reprinted with permission from Ref. 116; © 1991, American Chemical Society.)

cyclophilin complex that the conformation of a molecule bound to its target site may be very different from that in the free state (111–113). In addition, analog design is assisted by knowing the location of the binding region of the ligand. Studies were therefore undertaken to determine the bound state of the ligand as well as to identify those portions of the drug interacting with the binding protein.

The first investigations involved the analysis of ^{13}C carbonyl chemical shifts of C8 and C9 and the ^{1}H chemical shifts of the piperidine ring of FK506 bound to FKBP (114,115). The upfield shifts of the piperidine ring protons, as well as NOEs observed between these protons and aromatic protons of FKBP, suggested that the bound site on FKBP resided in an aromatic-rich domain, and allowed a putative binding site on FKBP to be proposed. It was also evident that the pipecolinyl functionality of FK506 and analogs was involved in the binding face of the ligand.

In another study (116), a uniformly ^{13}C-labeled ascomycin, (12), was prepared, allowing the bound conformation of ascomycin to be determined in the presence of FKBP. The enhanced ^{13}C signals were used to edit the ^{1}H NOESY spectra used for the structural analysis. Not only were the assignments of side-chain methyls made possible by the ^{13}C enrichment, but ligand resonances could be distinguished readily from those of the protein. The conformation of the ligand was determined from NOEs observed in a 3D HMQC-NOESY spectrum. The resulting ascomycin structure (Fig. 12.25) differed considerably from that of the uncomplexed FK506 obtained by X-ray crystallography, but was similar to that of rapamycin. In particular, the bound ascomycin displayed a *trans* orientation of the 7,8-amide bond, whereas this bond is *cis* in free FK506 and *trans* in rapamycin. The backbone structure of the macrocyclic ring differed from that of uncomplexed FK506, but

showed a similarity in the piperidine ring region to that of rapamycin. This study also showed that both the piperidine ring and the pyranoyl moiety of ascomycin are involved in the binding interface in the complex with FKBP. Ligand protons that show NOEs to the protein are in bold in Fig. 12.25. X-ray studies since have confirmed these results for both the FK506-FKBP and rapamycin-FKBP complexes, showing the *trans* orientation of the ligand amide bond in the bound conformation, and verifying the involvement of the piperidine and pyranoyl regions of the ligand in the binding interface (117).

The binding site of the FKBP complex has also been investigated through use of NMR spectroscopy. Michnick et al. (118)and Moore et al. (119) solved the structure of uncomplexed FKBP by use of ^{1}H-NMR methods. Although spectral overlap did not allow every structural constraint present to be identified unambiguously, convergent structures defining the global fold of the 107-residue FKBP protein were obtained. Previous biochemical data allowed the extensive aromatic cluster within the core of the structure to be identified as the ligand-binding pocket. The loop regions of the protein between residues 37–43 and 83–90, situated at the open end of the binding pocket, were also of interest. The loops were the least well defined regions of FKBP and were thought to be flexible, and perhaps involved in the binding interaction. Examination of ^{1}H and ^{15}N chemical-shift changes on addition of ligand supported this notion and suggested that significant structural changes in these loop regions occurred upon ligand binding (118).

In a later study, a high resolution structure of the complete ascomycin-FKBP complex was calculated by heteronuclear 3D and 4D NMR by Meadows et al. (120). Uniformly labeled [^{15}N]FKBP and [^{13}C,^{15}N]FKBP were prepared and incubated with unlabeled ascomycin to form the complexes. Three-dimensional NOESY-HSQC spectra, resolved according to ^{15}N shifts, were used to obtain the NH-NH NOEs within FKBP. CH-NH NOEs were derivedfrom a4D [^{13}C,^{1}H,^{15}N,^{1}H]-NOESY spectrum of the doubly labeled material in H_2O and CH-CH NOEs from the same experiment repeated in D_2O. Hydrogen bond constraints were obtained by the identification of slowly exchanging amide protons from a series of HSQC spectra acquired over several days. Torsional angle constraints were obtained from coupling constants measured in a 2D HMQC-J spectrum of [U-^{15}N]FKBP/ascomycin. In all, 1958 distance constraints were applied to the structure calculation, with the extra resolution afforded by isotopic labeling, as compared with the 590 and 1047 restraints used in earlier homonuclear studies (118, 119). Restraints defining the structure of bound ascomycin were obtained from the previously reported data of Petros et al. (116)and, along with the intermolecular NOE-derived distance constraints also reported in their study, the complete ascomycin/FKBP solution structure was calculated.

The extra detail afforded by the multidimensional NMR approach allowed the ligand-protein contact area to be located unambiguously and even specific intermolecular hydrogen bonds identified. The structure of the complexed FKBP was essentially similar to that of the uncomplexed structure, except that the "ill-defined" loop regions between residues 36–45 and 78–92 were found to adopt well-defined conformations in the complexed proteins, as preempted by previous studies. Although this difference may partially be a result of the differences in resolution achieved in the complexed and uncomplexed FKBP NMR studies, generally it was thought that binding involved some rearrangement of the 36–45 and 78–92 loops. This provides a good example of the dynamic nature of protein binding as revealed by NMR spectroscopy.

The dynamic aspects of the ligand-FKBP complex formation were pursued by Cheng et al. (121) through analysis of ^{15}N-NMR relaxation data. In particular, the increased backbone mobility for several residues within the 36–45 and 78–95 loops compared with that of the rest of the protein was noted. From analysis of the ^{15}N relaxation rates of FKBP complexed with FK506, it was found that flexibility was restricted along the entire polypeptide chain (122). This confirmed the proposition that the binding in-

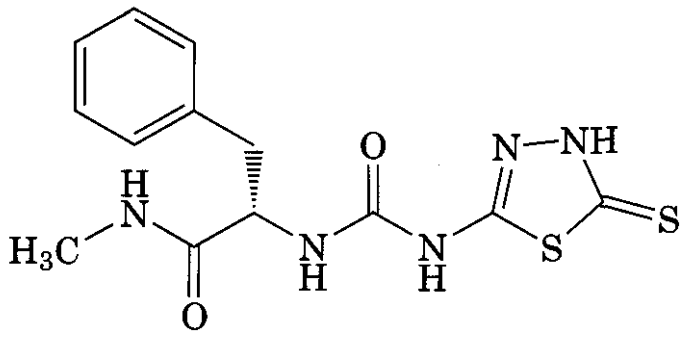
teraction of FKBP with ligand involves stabilization and structuring of the protein loops adjacent to the binding site.

In summary, it was possible not only to define the free and bound conformations of the ligand but also to identify the two binding interfaces involved in the interaction and demonstrate a reduction in protein mobility in a defined region of the protein upon binding. This level of analysis was possible because of the tight binding of the FKBP-ligand complex, its small size, and the availability of labeled species. The information proved to be complementary to X-ray crystallographic studies and will help to clarify the role of FKBP complex formation in immunoregulation.

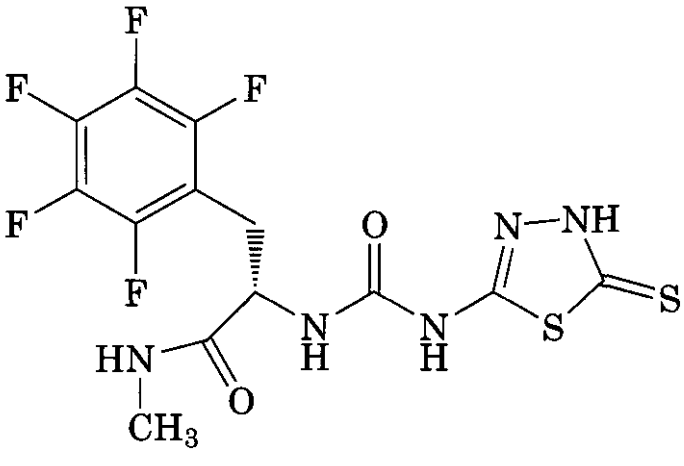
3.2.4.3 ***Matrix Metalloproteinases.*** Matrix metalloproteinases (MMPs), including stromelysin, collagenase, and gelatinase, are involved in tissue remodeling associated with embryonic development, growth, and wound healing. Unregulated or overexpressed MMPs have been implicated in several pathological conditions, including arthritis and cancer, and inhibitors of stromelysin and other MMPs have attracted much interest because of their potential for the treatment of these diseases.

Several NMR structural studies of stromelysin (123–127) and collagenase (128, 129) complexes have been reported. The secondary structure and global fold have been found to be quite similar for the catalytic domains of both enzymes and their various complexes with ligands. The active site in each enzyme is a cleft spanning the width of the enzyme, with a catalytic zinc atom coordinated by three histidine residues located in the center. Different dynamic properties of active-site residues in stromelysin/ligand complexes (3) and of collagenase with and without bound inhibitor (128, 129) have been reported. It has been proposed that structural and dynamic differences can be exploited in structure-based drug design to achieve broad inhibitor activity against several MMPs or to obtain more selective inhibition (3).

Of recent interest have been structural data on a novel class of MMP-binding inhibitors, represented by PNU-107859 (**13**) and PNU-142372 (**14**), which contain a thiadiazole moiety that coordinates the catalytic zinc atom through its exocyclic sulfur atom (130).



(13)



(14)

Isotope editing/filtering studies played an important role in defining interactions between the ligands and stromelysin. For example, for the stromelysin/PNU-107859 complex a 3D ^{12}C-filtered, ^{13}C-edited NOESY spectrum recorded on the [^{12}C,^{14}N]PNU-107859/[^{13}C,^{15}N]-stromelysin complex was used to assign protein/ligand NOEs. Of the 11 observed NOEs between the ligand and protein aliphatic protons, nine involved the aromatic ring of (13) and one involved the terminal methyl group. NOEs were observed between (13) and protons of Tyr^{155}, His^{166}, Try^{168}, and Ala^{169}. All four of these residues are located in the S_1-S_3 binding sites on one side of the active site. Comparison of 2D ^{1}H-^{15}N HSQC spectra showed that differences between the ^{1}H and ^{15}N chemical shifts for the stromelysin/**13** and stromelysin/**14** complexes are concentrated in the active site, indicating that no gross conformational differences in protein structure exist. The aromatic rings of (**13**) and (14) bind in the same region of the protein.

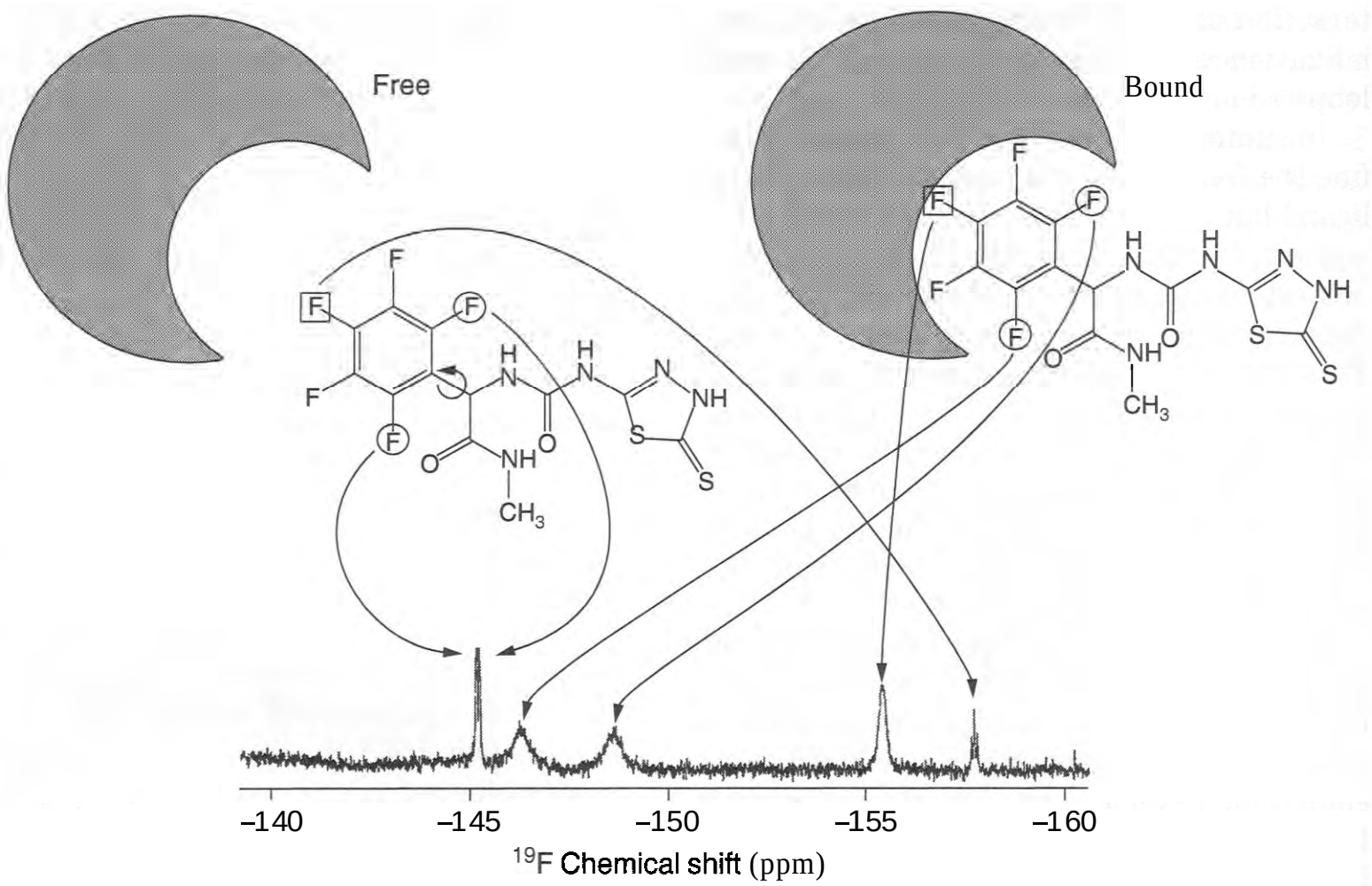


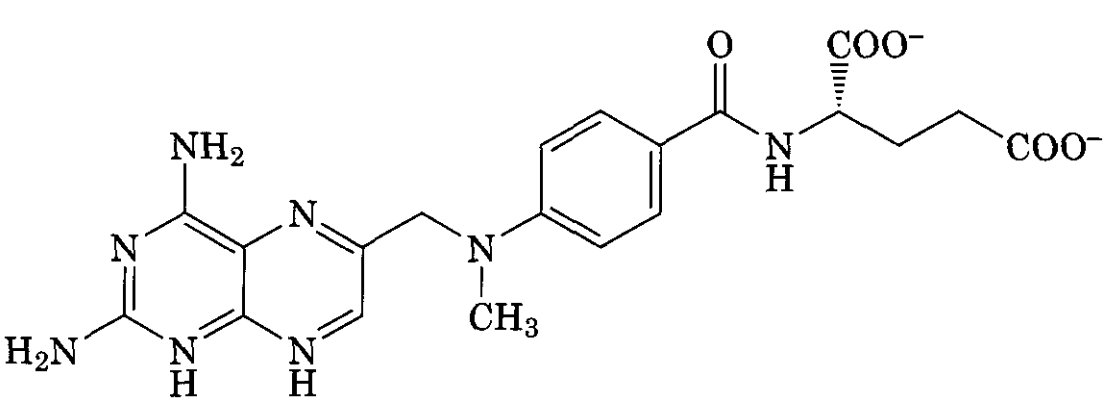
Figure 12.26. Region of the 1D ^{19}F spectrum of the stromelysin/PNU-142372 complex. Signals from free (sharp) and bound (broad) PNU-142372 are observed. (Adapted from Ref. **3** and reprinted with permission from Elsevier Science.)

A region of the 1D ^{19}F spectrum of the stromelysin/**14** complex is shown in Figure 12.26 (3). Two separate resonances were observed for the two *ortho* fluorine atoms of the bound ligand in contrast to the single resonance observed for both *ortho* protons of stromelysin-bound (**13**), indicating that the ring flip rate (rotation about the C^P-C^Y bond) is reduced for stromelysin-bound (14) compared to stromelysin-bound (13). A ring flip rate of approximately 100 s^{-1} was estimated from the difference in linewidths for the bound *ortho* and *para* fluorine atom resonances of (**14**), more than two orders of magnitude slower than the ring flip rate for (13). The ^{19}F spectrum in Figure 12.26 illustrates several general principles that are useful in NMR studies of ligand macromolecule complexes. First, note that the use of a rare probe nucleus such as ^{19}F produces spectra of elegant simplicity. Because there is no naturally occurring ^{19}F in the macromolecule, it generates no interfering signals. Second, the offset in chemical shift between bound and free signals reflects the different environment of the bound and free states. Third, signals from the bound ligand are broader than those from the free ligand because of the higher molecular weight of the complex but are still clearly visible for a complex of this size.

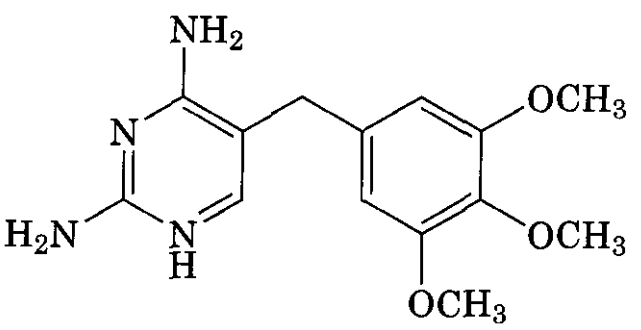
NMR studies have also been reported for ligands bound to collagenase. Interest so far has focused on hydroxamate-containing ligands, where it has been shown that binding causes a decrease in mobility of some but not all active-site residues (128, 129). Interestingly, some active-site residues adjacent to residues that interact directly with inhibitor were found to have high mobility both in the presence and the absence of inhibitor (129). This contrasts with what is observed for stromelysin complexed to hydroxamate ligands and a more complete understanding of the dynamics of the respective interactions may provide critical information for drug design (3).

Hydroxamate-containing ligands have also featured in other NMR studies, this time using transferred NOEs to determine their bioactive conformations (131). TrNOE data were used to determine the conformation of the inhibitors when bound to stromelysin. The NOE-derived structures of the bound inhibitors were used as templates to screen a database of 260,000 compounds. Eighteen of the 23 compounds identified for which stromelysin binding data were available had affinities less than 200 n*M*, demonstrating the value of deriving a conformationally restricted template for structure-based drug design (131). This study also demonstrates the close synergy that exists between structure-based design and screening approaches, either *in silico* or experimental.

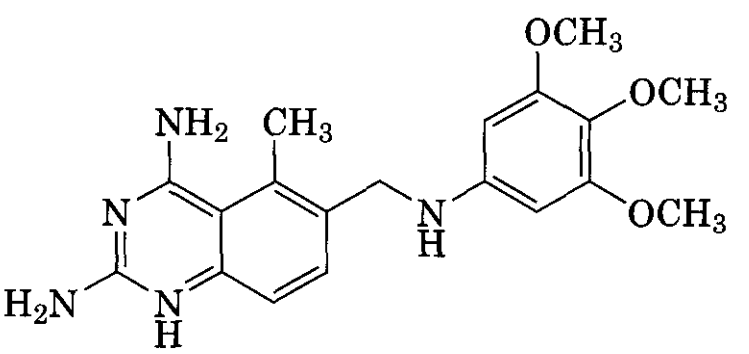
3.2.4.4 ***Dihydrofolate Reductase.*** Dihydrofolate reductase (DHFR) is an important intracellular enzyme that is the target of several clinically used drugs, including methotrexate (**15**), an anticancer compound, and trimethoprim (**16**), an antibacterial. These act by inhibiting the enzyme in malignant cells and parasites, respectively. The small size of DHFR (18–20 kDa) makes it amenable to structural studies and there have been numerous complexes determined using both X-ray and NMR methods. The focus here will be on a recent illustrative example of the structure of a new complex of DHFR with trimetrexate (**17**). Trimetrexate was initially investigated as an antimalarial agent but has subsequently been found to have antineoplastic activity against breast, neck, and head cancers. It has also been used as an antibacterial for the treatment of *Pneumocystis carinii* pneumonia in AIDS patients. As seen from the following structures, trimetrexate combines some of the features of trimethoprim and methotrexate:



(15)



(16)



(17)

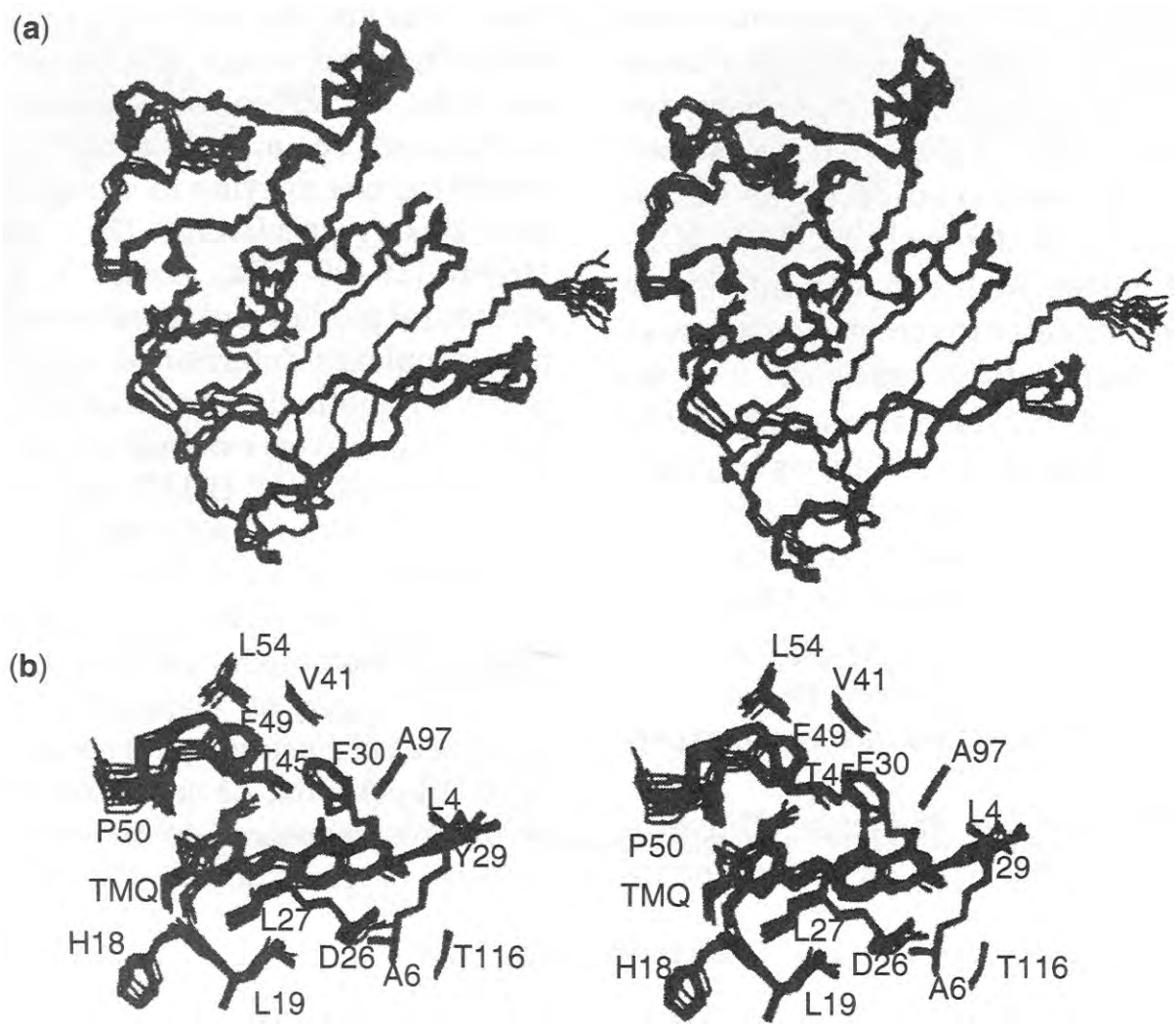


Figure 12.27. Stereoview of a superposition over the backbone atoms (N, Cα, and C) of residues **1–162** of the final **22** structures of the DHFR-trimetrexate complex. (a)View of the protein backbone and the trimetrexate heavy atoms. (b) View of trimetrexate in the binding site of enzyme. (c) Conformation of trimetrexate in the binding site of enzyme. The orientation of trimetrexate is identical for (a)-(c) and only its heavy atoms are shown. (Reprinted with permission from Ref. **132.** Copyright 1999 The Protein Society.)

The three-dimensional structure of the complex of DHFR with trimetrexate was determined using about **2000** distance restraints, **300** angle restraints, and **100** hydrogen-bonding restraints **(132).** Simulated annealing calculations produced a family of **22** structures consistent with the constraints. Several intermolecular protein-ligand NOEs were obtained by using a novel approach that monitored temperature effects of NOE signals resulting from dynamic processes in the bound ligand. At low temperature (5°C) the trimethoxy ring of bound trimetrexate flips sufficiently slowly to give narrow signals i slow exchange, which give good NOE cros peaks. At higher temperature these broade and their NOE cross peaks disappear, thus a lowing the signals in the lower temperatu spectrum to be identified as NOEs involvin ligand protons. Figure **12.27** shows the stru ture of the complex, including the orientatio of the ligand in the binding site.

The binding site for trimetrexate is well defined and was compared with the binding sites in related complexes formed with methotrexate and trimethoprim. No major conforma-

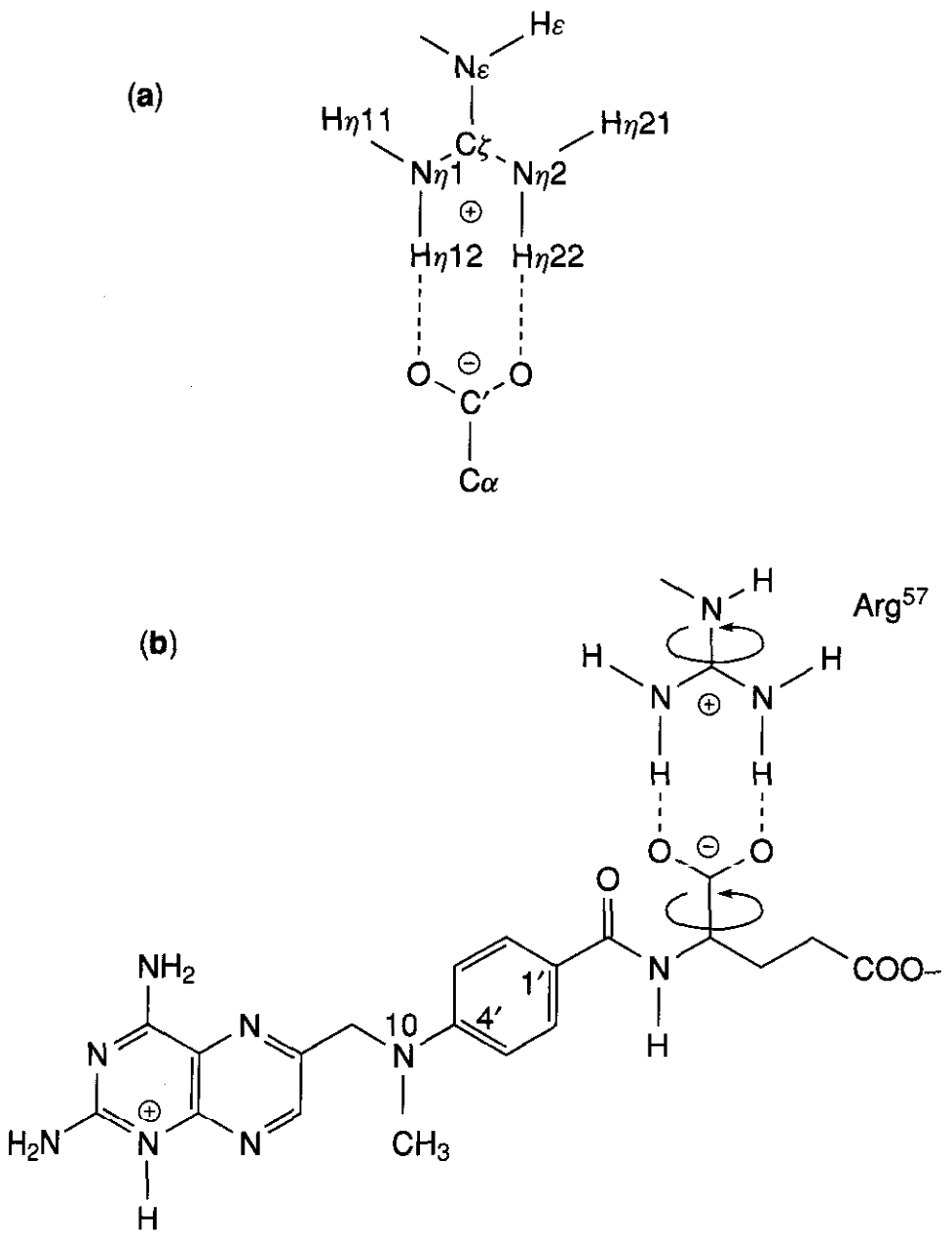


Figure 12.28. Correlated motions of a carboxylate group from methotrexate and Arg57 of DHFR detected by NMR. (a) Structure of an arginine-carboxylate complex formed with symmetrical end-on interactions and (b) structure of methotrexate showing its interaction with the guanidine group of Arg57 of **DHFR.** (Reprinted with permission from Ref. **133.)**

tional differences were detected between the different complexes. The 2,4-diaminopyrimidine-containing moieties in the three drugs bind essentially in the same binding pocket and the remaining parts of their molecules adapt their conformations such that they can make effective van der Waals interactions with essentially the same set of hydrophobic amino acids. The side-chain orientations and local conformations are not greatly changed in the different complexes.

The ring flipping of the trimethoxy aromatic ring mentioned above was detected by variable-temperature studies of the spectral line shape. The presence of such dynamics processes involving the ligand appear to be not uncommon in macromolecule-ligand complexes and the ability of NMR methods to detect such phenomena represents one distinct advantage of NMR over X-ray methods of structure determination. Relaxation measurements were also used to probe dynamics of the protein and no large amplitude motions were found, apart from that at the C-terminus (132). The power of NMR methods for studying dynamics of complexes is further illustrated by an earlier study of the complex of DHFR with methotrexate (133). In this case a correlated dynamic rotation of a carboxylate group on the ligand and Arg^{57} of the protein was detected, as illustrated in Fig. 12.28.

3.2.4.5 HIV ***Protease.*** Because of its essential role in the HIV life cycle, H N protease is a major target for structure-based design of anti-AIDS drugs. There are now more than 100 structures of HIV protease and protease inhibitor complexes in the HIV-protease structure database (134–136) and the availability of this wealth of high resolution structural information has been the driving force behind numerous structure-based design programs (134, 135, 137). Most of the high reso-

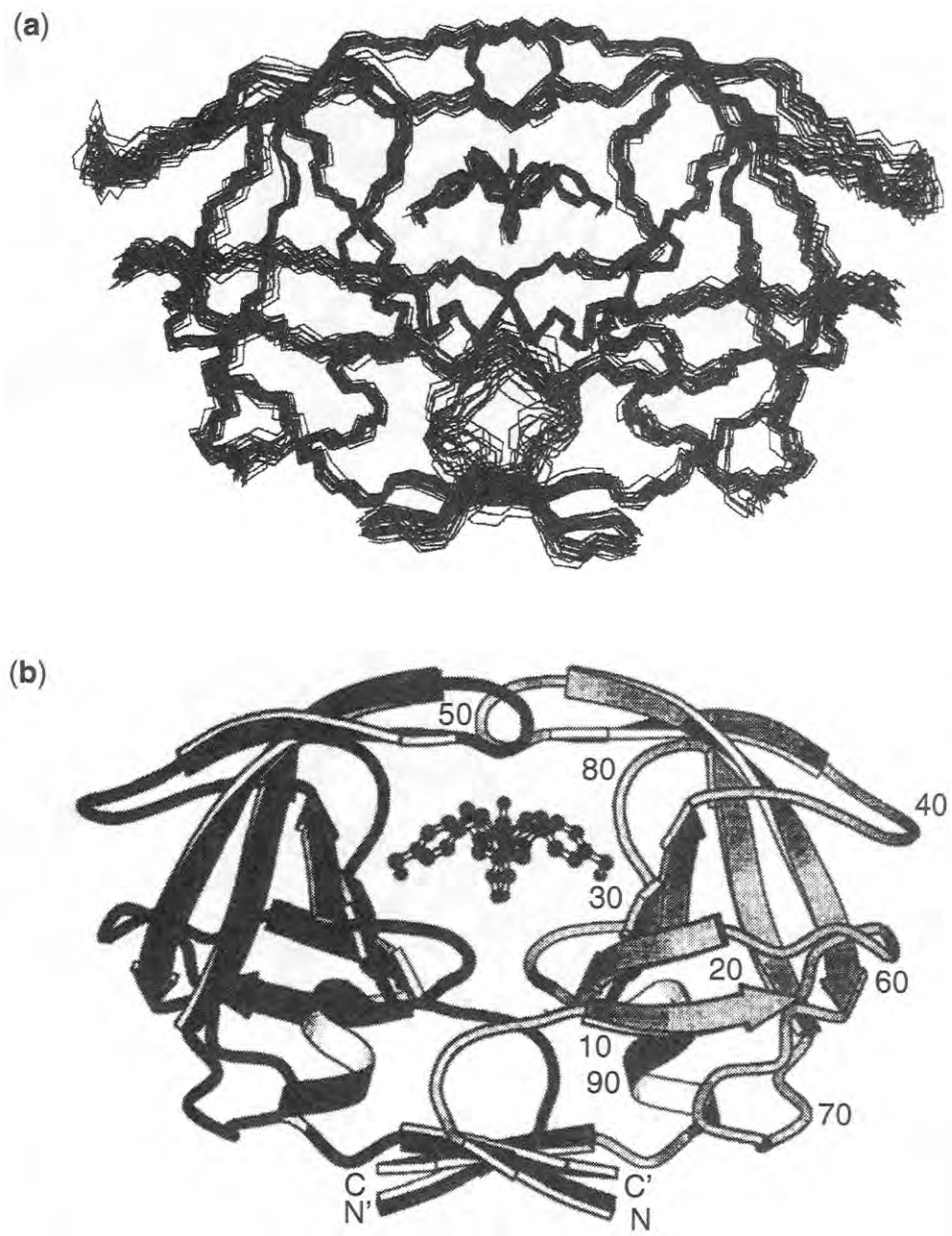
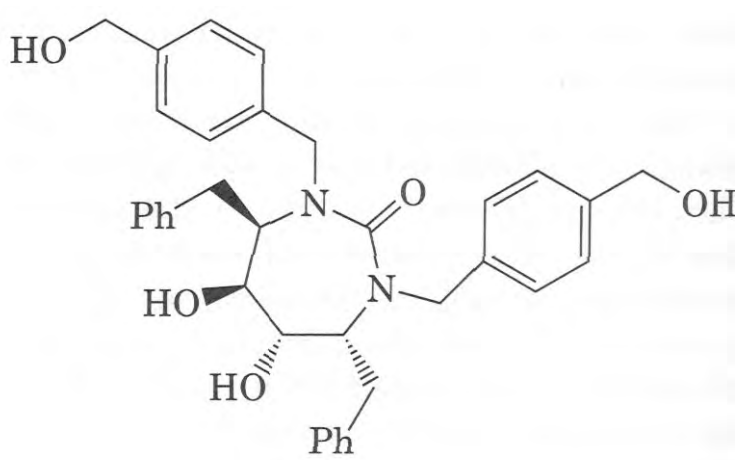


Figure 12.29. (a) View of the superimposed heavy atom (N, Ca, C) of the ensemble of structures of the HIV-1 proteases/DMP323 complex. (b) Ribbon diagram of the minimized average structure of the complex. (Reprinted with permission from Ref. 138. Copyright 1996 The Protein Society.)

lution structural information on HIV protease has been obtained from X-ray crystallography data (136). Although there are relatively few examples of HIV protease/inhibitor complexes that have been determined by use of NMR spectroscopy, the NMR data, taken together with the structural data from X-ray experiments, have contributed to an understanding of protease-inhibitor recognition and dynamics. Indeed, studies of HIV proteasefinhibitor complexes are a powerful example of the way in which complementary information obtained from X-ray crystallography and NMR spectroscopy can be used to facilitate structure-based drug design.

HIV protease/inhibitor complexes have a molecular weight of approximately 22 kDa. Although NMR spectroscopy is well suited to determination of the structure of molecules in this size range, efforts to determine the solution structure of the complex were hampered by the fact that the protease undergoes rapid autocatalysis in solution. It required the development of potent inhibitors before NMR studies of the complex became feasible. The first solution structure (Fig. 12.29) of HIV protease bound to the cyclic urea inhibitor DMP-323 (18)was reported in 1996 (138).



(18)

The protease exists as a homodimer. Each 99-residue monomer contains 10 β-strands and the dimer is stabilized by a four-stranded antiparallel β-sheet formed by the N- and C-terminal strands of each monomer. The active site of the enzyme is formed at the interface, where each monomer contributes a catalytic triad (Asp^{25}-Thr^{26}-Gly^{27}) that is responsible for cleavage of the protease substrates. The "flap region" is located above the reactive site and is formed by a hairpin from each monomer of two antiparallel β-strands joined by a β-turn. There is little difference between the solution and crystal structures of protease-inhibitor complexes, except in those regions where the polypeptide chain is disordered. However, experiments in solution have allowed access to parameters that are not directly accessible from crystal data. These parameters, such as the amplitude and frequency of backbone dynamics, the protonation states of the catalytic aspartate residues, and the rate of monomer interchange, are essential in understanding the interaction of HIV protease with potent inhibitors.

The cyclic urea inhibitor DMP-323 was designed by analysis of crystal structures of HIV proteaselinhibitor complexes. A feature common to many of the complexes of HIV protease is a buried water molecule that bridges the inhibitor and Ile^{50} in the flaps. Interactions with this water molecule are thought to induce the fit of the flaps over the inhibitor (139). In contrast, mammalian aspartic-proteaselinhibitor complexes are unable to accommodate an equivalent water molecule (135). This observation led to the design of a series of cyclic urea-based inhibitors that are capable of displacing the buried water molecule (139). As well as improving the specificity of inhibitors to the viral protease, displacement of the water molecule was expected to increase the entropic contribution to inhibitor binding and thus enhance the affinity of complex formation. The cyclic urea inhibitors are highly potent and specific inhibitors of HIV protease (139) and for DMP-323 it has been shown in both the crystal structure (139) and in solution (140) that the urea moiety does indeed replace the buried water molecule.

Although DMP-323 replaces one buried water molecule, several others are observed in the crystal structure of the complex. A more recent NMR study investigated the role of these water molecules to determine whether any had a structural role in the formation of the HIV protease/DMP-323 complex (141). In favorable cases, NMR can be used to estimate the residence times of hydration water molecules (142), thus providing information about the timescale of the interaction of buried water with the bulk solvent. This analysis led to the identification of a symmetry-related pair of water molecules that may have a structural role in formation of the complex. Such information may prove useful in the design of future cyclic urea inhibitors. An interesting finding in this study was the fact that each of the hydroxyl protons of DMP-323 is in rapid exchange with solvent. This is a surprising result, given that two of these hydroxyl protons are completely buried and form a network of hydrogen bonds with the catalytic Asp^{25}/Asp^{125} side chains (143). Furthermore, the dissociation rate of DMP-323 is less than $1\,s^{-1}$ under the conditions of the experiment, which is too small to average the chemical shifts of the hydroxyl protons and the bulk water. The observation is ascribed to local fluctuations in the complex that allow solvent molecules to penetrate into the binding site. This conclusion is supported by the observation that the catalytic protons of the Asp^{25}/Asp^{125} side chains in the protease/DMP-323 complex undergo H-D exchange with solvent, even though they are buried and hydrogen bonded to the inhibitor (143). These studies highlight that even well-ordered structures such as the protease/DMP-323 complex may be flexible on the millisecond to microsecond timescale.

Interestingly, in the DMP-323 complex, both of the catalytic Asp^{25}/Asp^{125} side chains are protonated over the pH range 2–7 (143). The protonated Asp^{25}/Asp^{125} residues form a network of hydrogen bonds with the hydroxyl groups of DMP-323. In contrast it has been shown that in the complex with the asymmetric inhibitor KNI-272, the side chain of Asp^{25} is protonated, whereas that of Asp^{125} is not. A suggested explanation for this is that both oxygens of the Asp^{125} side chain are deprotonated to accept two hydrogen bonds, one from a bound water molecule and one from the inhibitor. In contrast the side chain of Asp^{25} is

protonated so that it can donate a hydrogen bond to the inhibitor (144). Consequently, the protonation state of the enzyme is influenced strongly by interaction with specific inhibitors and this knowledge is essential for a detailed understanding of the protease/drug interactions.

NMR has also been used to study the relationship between flexibility and enzymatic function for HIV protease. For the protease/DMP-323 complex, ^{15}N spin-relaxation studies determined that residues that are flexible correlate well with residues that are disordered in the NMR structure of the complex (145). For example, residues in poorly defined loops were found to undergo large-amplitude internal motions on the nanosecond-picosecond timescale. In contrast, two regions of the molecule were found to exhibit motions on the millisecond-microsecond timescale. The first of these is at the N-terminus of the protein around Thr^4-Leu^5. This is adjacent to the major site of autolysis of the protease and it has been suggested that the rate of cleavage may regulate HIV protease activity *in vivo* (146). Consequently, the observed flexibility may be important for regulation of protein function. The second region found to be undergoing millisecond-microsecond motion was the tips of the flaps around Ile^{50}-Gly^{51}. In crystal structures, this region of the protease is well ordered and not involved in crystal contacts, although its conformation varies from structure to structure. This motion is interpreted as a dynamic conformational exchange process, which is fast relative to the chemical-shift timescale. Thus when the protease is bound to a symmetric inhibitor in solution, this conformational exchange results in the chemical shifts of the flap residues in the two monomers being identical (138, 145). In contrast, when the protease is bound to an asymmetric inhibitor, such as KNI-272, crystal structures show that each monomer interacts with the inhibitor in a different way (144). This is reflected in the fact that the chemical shifts of the monomers are different when asymmetric inhibitors are bound (141, 147). Analysis of spectra from such an asymmetric complex has revealed that the inhibitor is capable of "flipping" its orientation with respect to the two monomers without dissociating from the complex (148). These data again highlight the importance of defining both the structural and dynamic aspects of binding to understand the requirements for potent interactions between HIV protease and its inhibitors.

The development of inhibitors of HIV protease represents a major success for structure-based drug design. When HIV was first identified in the early 1980s there were no known drugs effective for treatment of infection. A combination of X-ray crystallography, NMR spectroscopy, computer modeling, and chemical synthesis has resulted in the development of several effective HIV protease inhibitors. However, in common with other retroviruses, HIV has a high transcription error rate that results in a rapid mutational rate. One of the results of this is the production of a divergent population of viruses in which the sequence of the HIV protease produced may differ substantially (149, 150). As a consequence, drug-resistant strains of the virus emerge. Clearly, knowledge of the structural principles that govern inhibition of the protease and the mechanism by which the virus develops resistance will continue to be important in the development of effective new drugs.

4 NMR SCREENING

In the past, NMR was predominantly used in the design stage of drug discovery rather than the screening stages. Recently, new methods that make use of NMR to screen ligands for binding to a protein target have been developed and are proving to be a powerful tool in the discovery of new drug leads. This section gives an overview of the various experimental methods, summarized in Table 12.8, which can be used to screen mixtures of ligands for binding to a drug target. There will also be a brief discussion on the practical considerations that need to be made when designing an NMR screening program.

4.1 Methods

4.1.1 Chemical-Shift Perturbation. Chemical shift is a function of the chemical (and hence magnetic) environment that individual nuclei experience. Perturbations of chemical

Table 12.8 Summary of the Methods Available for NMR Screening and Their Respective Characteristics

Screening Methodology	Signals Observed	Protein Size Limit	Labeling	Binding Information Obtained	K_D Limit	K_D Determined	Suitable for HTS	Mixture Deconvolution Required?
Chemical shift perturbation (e.g., SAR by NMR)	Protein	<30 kDa	^{15}N protein	Location	10^{-3}–10^{-7} M	✓	✓	✓
STD	Ligand	None	None	Orientation	10^{-3}–10^{-7} M	x	✓	x
Diffusion-based (e.g., affinity NMR)	Ligand	None	^{2}H protein for isotope editing	None	10^{-3}–10^{-7} M	✓	✓	x
Relaxation-based	Ligand	None	None	None	10^{-3}–10^{-7} M	✓	✓	x
trNOE	Ligand	None	None	Bound conformation	10^{-3}–10^{-7} M	x	✓	x
NOE pumping	Ligand or protein (reverse)	None	None	Bound conformation	10^{-3}–10^{-7} M	x	✓	x ✓[a]
Spin labeling	Ligand or protein	None	Spin label for either ligand or protein	Orientation, simultaneous binding	10^{-3}–10^{-6} M	x	✓[b]	✓[b]

[a]For reverse NOE pumping.
[b]For primary screening if the protein is spin-labeled or for second-site screening if the first-site ligand is spin-labeled.

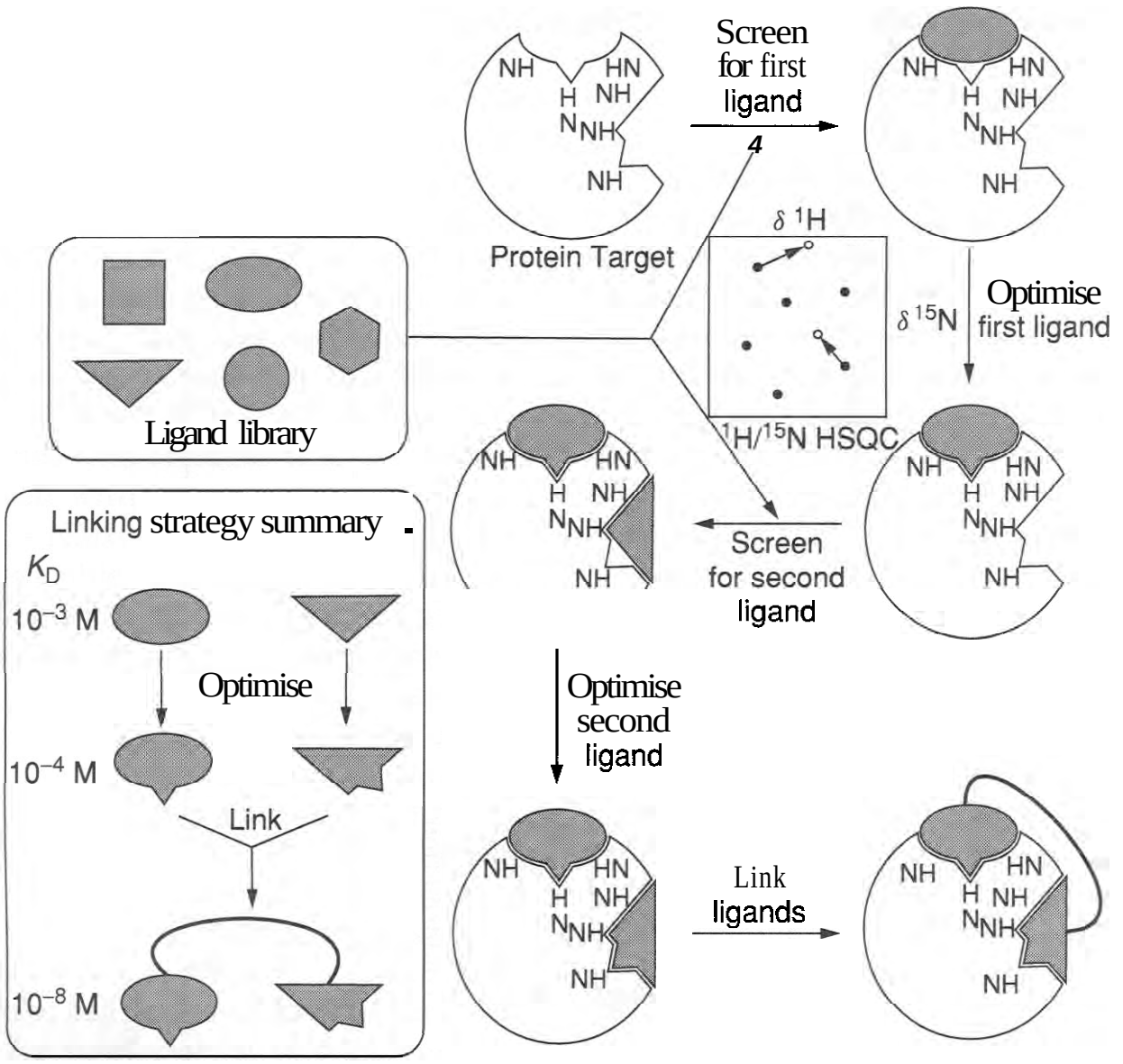


Figure 12.30. Summary of the SAR by NMR drug discovery methodology. A protein target is screened against a library consisting of small organic molecules by use of the $^{1}H/^{15}N$ HSQC experiment. When two ligands that bind in close proximity are identified, they are linked to form a composite ligand with an increased affinity for the target.

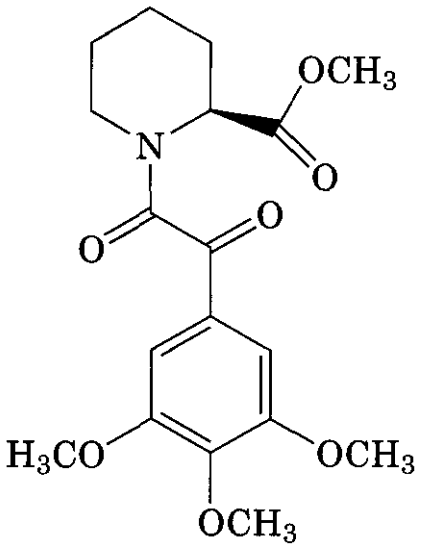
shifts can be used to detect binding of a ligand to a protein target. When a ligand binds to a protein the local chemical environment is changed, and this is reflected by a change in the chemical shifts of nuclei in close proximity to the ligand-binding site. The most common experiment used in this screening methodology is the $^{1}H/^{15}N$ HSQC that generates a discrete signal for each amide group within the protein. A reference $^{1}H/^{15}N$ HSQC spectrum, which is acquired in the absence of potential ligands, is compared to a spectrum recorded in the presence of ligands and any changes in the amide chemical shifts are indicative of a ligand binding to a location close to the corresponding amide groups. The major advantage of this technique is that, if the NMR assignment of the amide resonances is known, then the site of binding for each ligand can be determined. This is a valuable piece of information in the development of more potent second-generation drug leads. Binding affinities can also be determined by measuring the change in chemical shift as a function of ligand concentration

One technique that utilizes this screening method for drug design is "SAR-by-NMR," developed by Fesik and coworkers (1, 4, 151-155). SAR-by-NMR is a fragment-based drug design approach in which a potent drug candidate is derived by chemically linking two or more small low affinity ligands for a target. In theory, the binding energy of the linked compounds will be the sum of the binding energies of the two individual compounds plus contributions to binding energy attributed to linkage. Therefore, it is possible to generate a drug lead with a nanomolar dissociation constant (K_D) from two milli- to micromolar fragments.

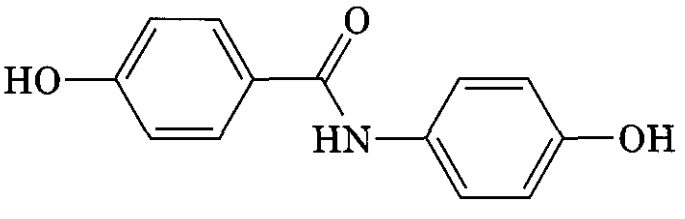
The first step in this process (Fig. 12.30) involves screening a library of ligands (typically with a MW < 400) in mixtures of up to 10 for binding to a protein target by comparing the $^{1}H/^{15}N$ HSQC spectrum of a ^{15}N-enriched protein in both the presence and the absence of ligands. Any ligand-induced changes in the chemical shift of the nitrogen and amide proton signals indicate binding of one or more ligands in the mixture to the protein target. The mixture containing the binding ligand(s) is deconvoluted and each individual compound screened to identify the individual ligand(s) responsible for the observed chemical-shift perturbations. Once a binding ligand is identified analogs can be screened to optimize binding.

A second ligand, which binds at a proximal site, is then identified either from the original screen or by repeating the library screening with the first ligand site bound to the protein. This ligand is then optimized and the structure of the ternary complex determined by use of either NMR or X-ray crystallography. The ternary complex structure provides information on the conformation and orientation of the bound ligands, which facilitates the synthesis of hybrid molecules where the two ligands are joined by a suitable linking moiety.

There are several examples that illustrate the potential of SAR by NMR. As noted earlier, FK506 binding protein (FKBP) inhibits calcineurin and blocks T-cell activation when complexed to the immunosuppressant FK506. This protein was used as a target for SAR by NMR screening and subsequently two ligands, (19) and (**20**), were identified with K_D values

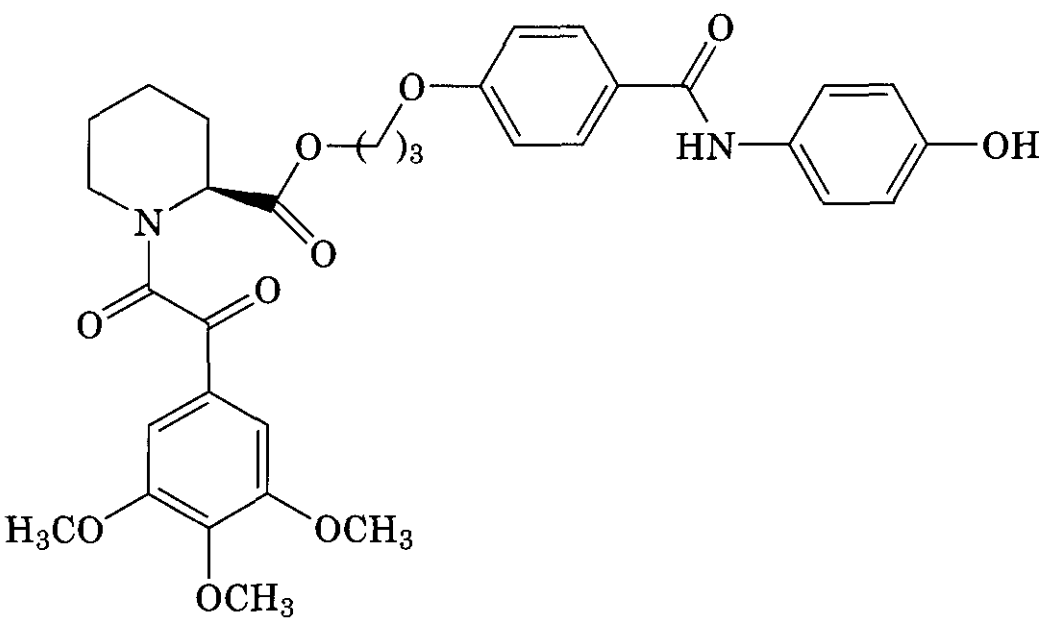


(19)



(20)

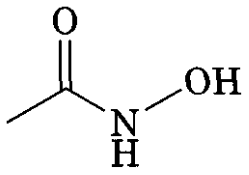
of 2 μM and 0.1 mM, respectively. A model of the ternary complex between the protein and both ligands was produced, which indicated that the methyl ester of (19) was close to the benzoyl hydroxyl group in (20). These two groups were linked with alkyl chains of various lengths, with the most active compound



(21)

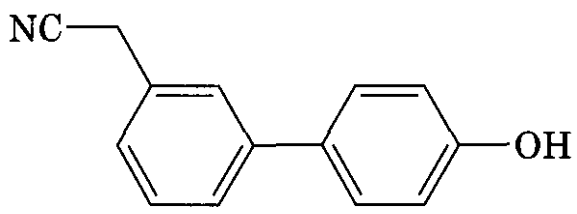
(21) having a three-carbon linker and a K_D value of 19 nM (151).

Inhibitors of the matrix metalloproteinase (MMP) stromelysin have also been designed through the use of the SAR-by-NMR screening methodology. As mentioned in section 3.2.5.3, MMPs are involved in matrix degradation and tissue remodeling, with overexpression of these enzymes being associated with arthritis and tumor metastasis. Acetoxyhydroxamate (22) was used as one ligand be-



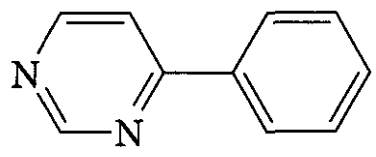
(22)

cause it was known previously that MMP inhibitors contain a hydroxamate moiety. The K_D value of (22) was determined to be 17 mM. To identify a second ligand the protein was screened against a ligand library in the presence of saturating amounts of (22). The library was biased for hydrophobic compounds, given that stromelysin demonstrates a substrate preference for a hydrophobic amino acids and structural studies had identified a hydrophobic binding pocket supporting this observation. From the library screen a series of biphenyl compounds were identified and analogs of these compounds were synthesized. A biphenyl derivative (23) was produced with a

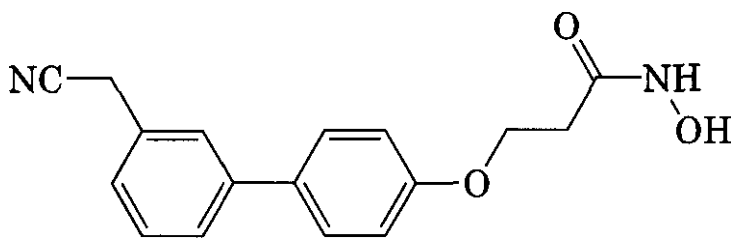


(23)

K_D value of 0.02 mM. The NMR structure of a ternary complex, consisting of stromelysin (22) and the biaryl derivative (24) (chosen for its superior aqueous solubility), was determined and indicated that the methyl group of (22) was in close proximity to the pyrimidine ring of (24). With this information (22) and (23) were subsequently linked by different



(24)



(25)

length linkers and the most active compound produced, (25), had a K_D value of 15 nM (154).

A variation of SAR-by-NMR is to optimize binding or improve the pharmacological properties of known drug leads generated by other methods (e.g., natural products isolation or combinatorial chemistry). A compound can be fragmented into individual subunits and then alternative fragments identified through use of $^1H/^{15}N$ HSQC screening. These fragments can then be incorporated into the molecular structure in the hope of improving the binding and/or pharmacological properties of the parent compound (Fig. 12.31). The alternative fragment must bind in the same location as the corresponding section of the original molecule, making $^1H/^{15}N$ HSQC screening method ideal as it provides information on the binding site of ligands.

In a demonstration of this fragmentation method, an antagonist of the interaction between leukocyte function-associated antigen 1 (LFA-1) and intracellular adhesion molecule 1 (ICAM-1) was used as a starting molecule. This interaction plays a role in the inflammatory response and specific T-cell immune responses, and inhibitors have applications in the treatment of inflammation and organ transplant rejection. The p-arylthio cinnamide antagonist (26) had an IC_{50} value of 44 nM; however, it was envisaged that the molecule's activity and physical properties could be improved by replacing the isopropyl phenyl group with a more hydrophilic moiety. Screening of a 2500-compound library provided several hits, and analogs of (26) were made that

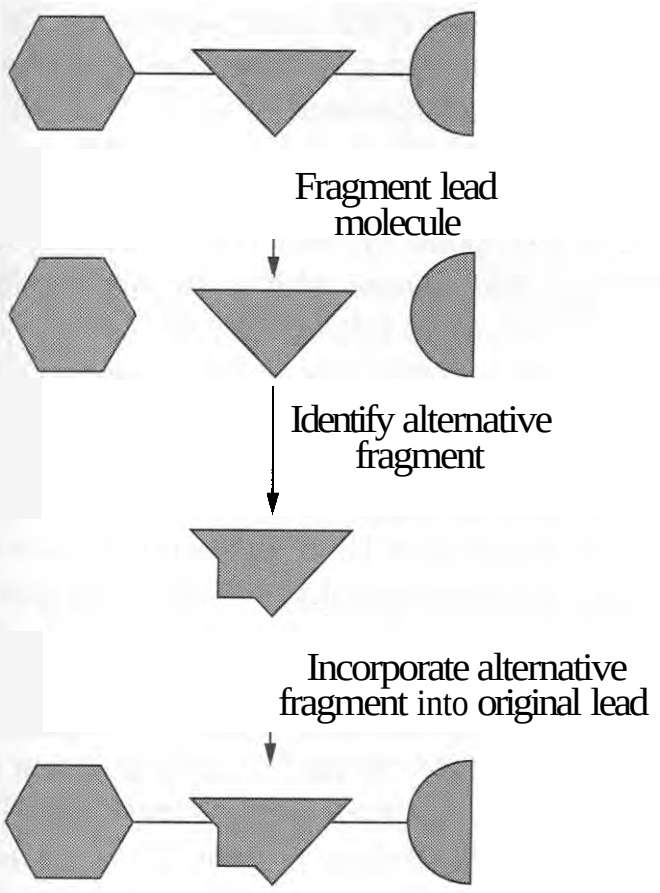
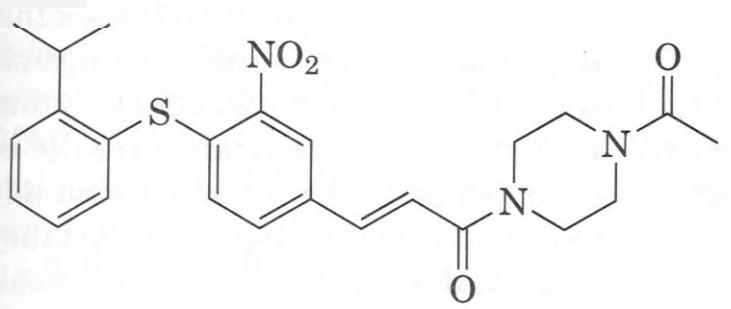
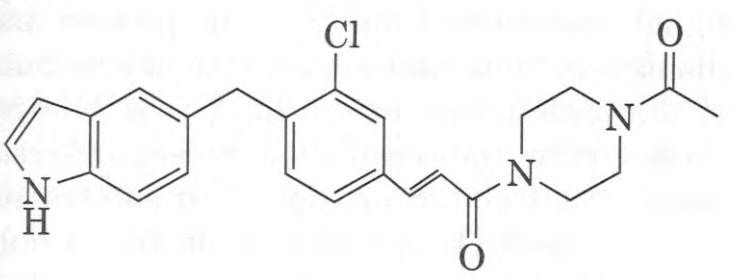


Figure 12.31. The fragment optimization approacfi developed from SAR by NMR. A known ligand of a protein is broken into fragments and small molecules based on the fragments are screened for binding. Any molecules that are found to bind can then be incorporated into the original lead compound with the hope of improving its binding and/or physicochemical properties.

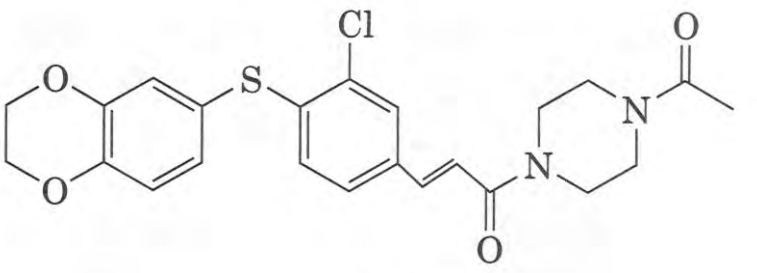


(26)

incorporated these ligands in place of the isopropyl phenyl group. Compounds (**27**) and (28)had both improved aqueous solubility and pharmacokinetic profiles, with similar or im-



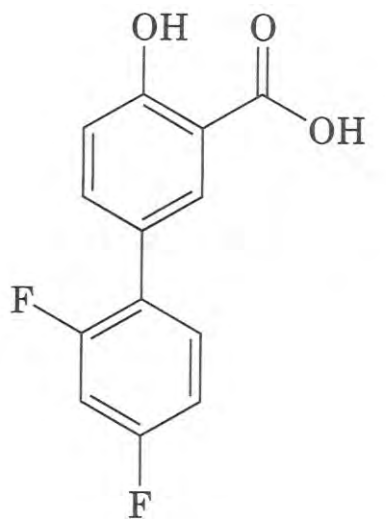
(27)



(28)

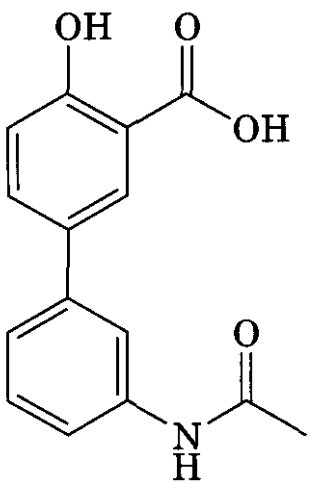
proved activity (IC_{50} values of 20 and 40 nM, respectively) when compared to that of the parent compound (26) (156).

Many compounds bind to human serum albumin (HSA), which significantly reduces their in *vivo* activity and hence their potential as a drug lead. The fragmentation method has recently been used to find analogs of diflusinal (29) that have a reduced affinity toward HSA

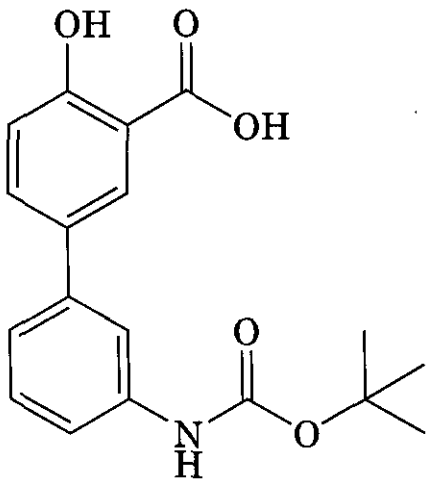


(29)

(157). Diflusinal is a nonsteroidal anti-inflammatory that is more potent, longer acting, and is more tolerated in *vivo* than aspirin. However, **99%** of diflusinal is bound to albumin in plasma and, as a result, high doses are required for it to be effective. Structural studies of the diflusinal/HSA-III complex indicated that, by introducing polar functionality to the difluorophenyl moiety, binding affinity to HAS may be reduced without affecting activity. A series of organic compounds analogous to the difluorophenyl fragment were screened using the $^{1}H/^{15}N$ HSQC chemical-shift perturbation method and several alternative fragments were identified. These were incorporated into the diflusinal structure, resulting in a number of analogs (e.g., **30** and **31)** with reduced affinity for HSA but still maintaining some activity. It was predicted that the next



(30)



(31)

generation of compounds would incorporate the fluorine atoms present in (29) into leads such as (30) and (31) because the presence of fluorine increased activity without increasing the affinity for HSA.

Overall, SAR-by-NMR is already showing great promise as a major new tool in drug design, but it has a few limitations. First, the screening of large ligand libraries and any subsequent deconvolution steps can be expensive, in that a large amount of ^{15}N-labeled protein is needed. However, as already noted, recent advances in cryoprobe technology have reduced the amount of protein required. The use of smaller, more druglike libraries such as that described in the SHAPES ideology (7) could also be used to reduce the amount of protein required for screening. The second limitation is that the NMR assignments for the protein must be known or determined. This limits the size of the protein target to 30 kDa or less, although this value will presumably increase as TROSY-based NMR technologies for structure determination advance. The method also requires the comparison of ^{1}H/^{15}N HSQC of the protein in both the absence and the presence of ligands. Changes in solvent conditions such as pH, polarity, salt concentration, and viscosity may cause shifts in amide resonances, leading to false positives (158).

Recently, a protocol was described that screens ligands using the ^{1}H/^{15}N HSQC experiment but includes a mass spectrometry prescreening step. Ligand mixtures are added to a protein target and then subjected to size-exclusion chromatography, which separates ligand/protein complexes from free ligands. The identity of the bound ligands can then be determined using MS data and, once identified, screened by ^{1}H/^{15}N HSQC to determine the location and specificity of binding (159). This MS/NMR methodology reduces the amount of ^{15}N-labeled protein required because only a fraction of the library is screened by NMR and there is no deconvolution step.

4.1.2 Magnetization Transfer Experiments. Proteins are composed of a large network of dipole-dipole interactions, resulting in the efficient transfer of magnetization throughout the molecule. The saturation transfer difference (STD) experiment (Fig. 12.32) uses this phenomenon to detect the binding of ligands to a protein. It relies on the fact that saturation of a single protein resonance results in saturation of all protein resonances and any ligands that bind to the protein, provided they are not affected directly by the selective saturation pulse (160). The STD experiment is able to detect the binding of ligands with a K_D between 10^{-3} and 10^{-8} M.

An STD experiment consists of irradiating an isolated protein resonance (either at low or high field) with a series of pulses that saturate the entire protein and any binding ligands. This results in a spectrum containing reduced signal intensities from both the protein and the ligands that bind to it. A second spectrum of the protein and ligand library is then recorded with the saturation pulses off-resonance. Subtraction of these two spectra results in the STD spectrum that shows only those ligands that bind to the protein (residual protein resonances are removed using a T_2 re-

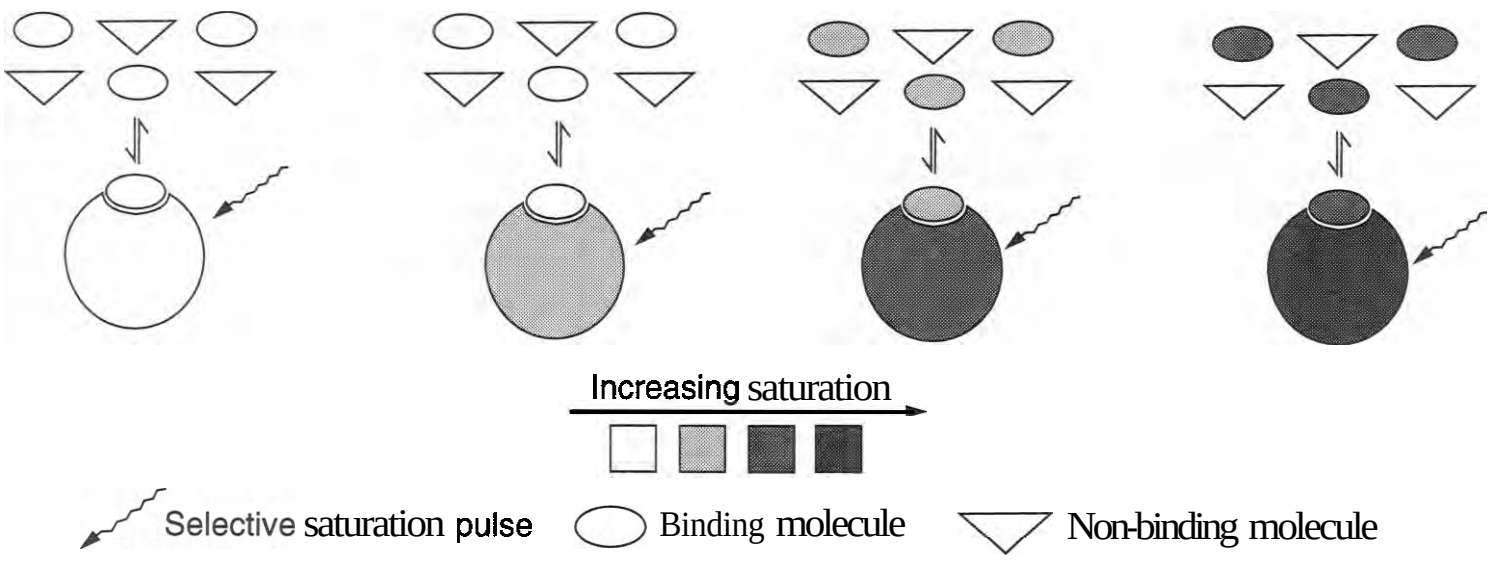
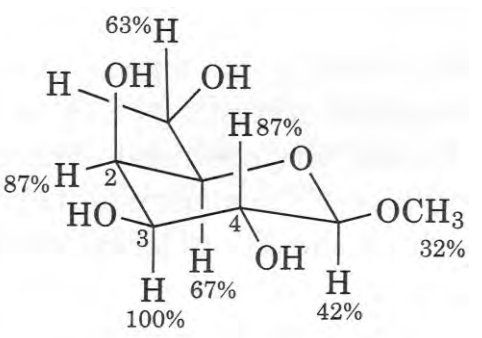


Figure 12.32. A schematic representation of the saturation transfer effect. The protein resonances are saturated (indicated by shading) by a selective pulse by spin diffusion. Resonances of nonbinding ligands (triangles) are not affected by this pulse but ligands that are interacting with the protein (ellipses) will also become saturated. These interacting ligands are transferred to solution through chemical exchange where they are detected. (Reprinted with permission from Ref. 16. Copyright 1999, Wiley-VCH.)

laxation filter). This subtraction occurs internally through phase cycling after every scan, to reduce artifacts attributed to temperature or magnetic field variations (160, 161).

An STD can be added to most forms of NMR experiments including COSY, TOCSY, NOESY, and inversely detected ^{13}C or ^{15}N spectra. A high resolution magic angle spinning (HR-MAS) STD experiment has been developed to study the binding of ligands to a protein immobilized on a solid support. HR-MAS STD NMR provides a way of obtaining ligand-binding information for proteins that are difficult to work with in solution attributed to either poor solubility or conformational changes (161). In addition to screening ligands for binding to a protein (160, 162, 163), the binding epitope of a molecule can also be determined by examining the intensities of ligand resonances (164–166). The proton signals having the strongest signals will correspond to those that are part of the ligand's binding epitope. For example, it was shown that when methyl β-D-galactoside (32) bound to *Ricinus communis* agglutin *I* the H2, H3, and H4 were saturated to the highest degree (values on structure indicate relative signal strengths) and hence were in close proximity to the protein protons. This analysis was subsequently extended to the decasaccharide NA, (33) and demonstrated that the Gal-6' and

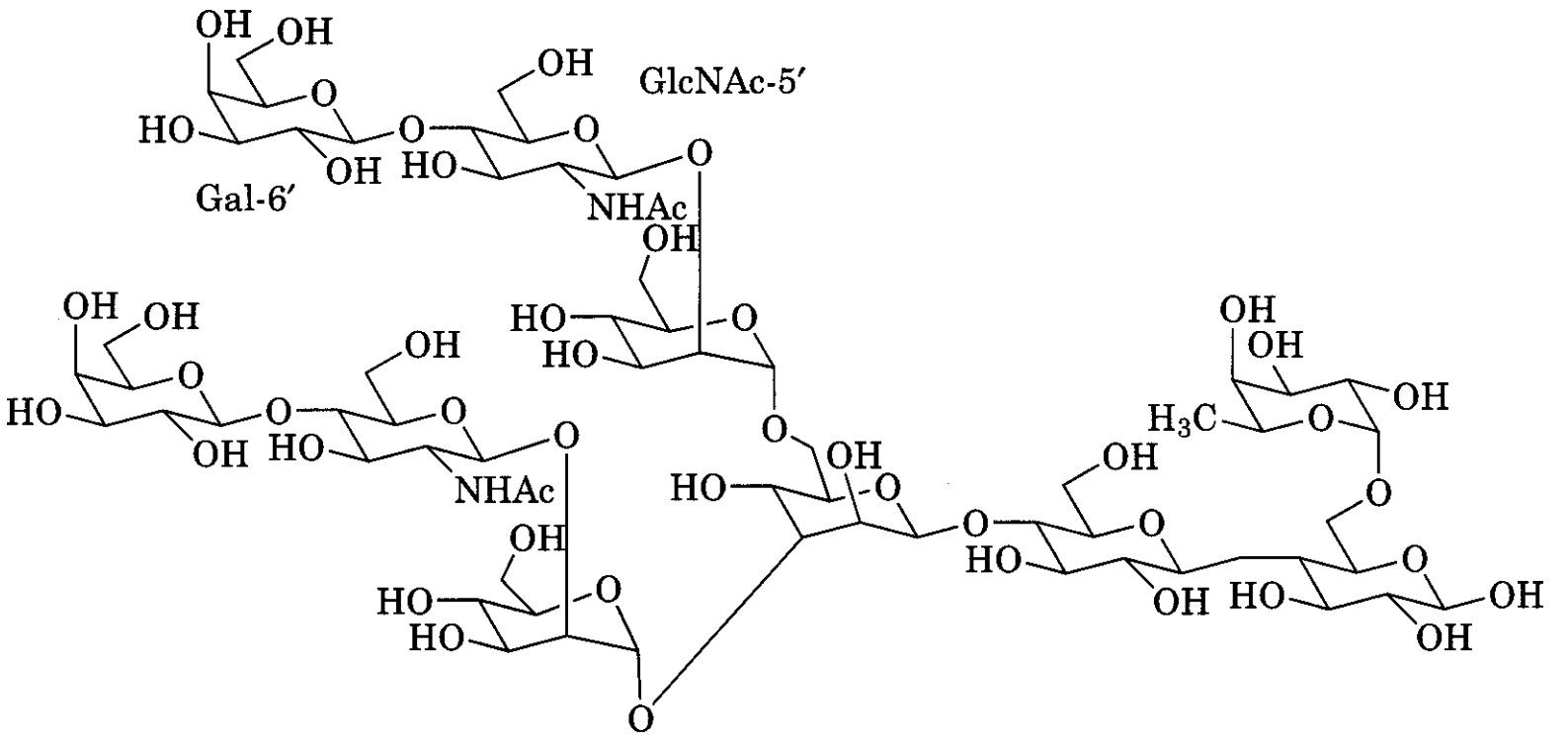


(32)

GlcNAc-5′ residues bind edge-on to the protein, with the binding contribution of the terminal galactose residue being the greater (165).

STD NMR studies have also been performed on membrane-bound receptors by embedding the protein in the phospholipid bilayer of a liposome (166).

There are a number of advantages in using the STD experiment for the detection of ligand binding. The saturation transfer effect is an efficient process, which results in high sensitivity, and hence only small quantities of protein are required (nanomolar concentrations of a protein with MW > 10 kDa) (160,165). In addition, protein size is noncritical; in fact, as the protein becomes larger, the saturation transfer effect becomes more efficient. The acquisition time for each experiment is also quite short and, because the experiment is ligand observed, no deconvolution of mixtures



(33)

is required, making this a good technique for high throughput screening of large ligand libraries. Unlike the chemical-shift perturbation techniques, STD experiments provide no information on the site of ligand binding.

A second variation of saturation transfer experiment has been devised by Dalvit and coworkers that uses the transfer of magnetization from the water (167). Water is intimately associated with proteins being bound either within or on the surface of the macromolecular structure. Saturation of the water resonance will lead to protein saturation through a variety of mechanisms, including saturation of the αH resonances, saturation of exchanging protein resonances, and NOE interactions between water and the protein. If a compound is bound to the protein it will also become saturated, and this effect can be used as an indication of ligand binding (167).

4.1.3 Molecular Diffusion. Molecules can be distinguished based on their diffusion coefficients, which are related to molecular size. Large macromolecules, such as proteins, diffuse more slowly than small molecules and it is this size difference that can be exploited to screen for ligand binding. If a small molecule binds to a protein target its diffusion coefficient is altered to a value more like that of the protein. Therefore, by utilizing a diffusion filter, resonances generated by small molecules that do not bind to the protein can be removed from the spectrum.

Diffusion editing is achieved with the use of a pair of gradient pulses. If field homogeneity is ignored, then all spins experience an identical magnetic field despite having different positions throughout the sample. The application of a field gradient has the effect of making field strength dependent on position. Under the influence of the gradient pulse, the phase of individual spins become dependent on their position within the sample and hence the spins are spatially "encoded." If diffusion does not occur, this spatial encoding is fully reversible by a second gradient of inverse polarity and no loss of NMR signal will occur. However, the second gradient pulse will be unable to "decode" the spins that have undergone diffusion and the resulting NMR signal will be reduced. Acquiring spectra of a sample with and without the diffusion filter and then subtracting them allows the ligands binding to the protein to be identified. This filtering method can be used for both 1D and 2D experiments and can be "tuned" by altering the strength and duration of the gradients.

Because the ligand signals are being observed in this screening method, no convolution of the ligand mixture is required, given that any signals can be assigned directly to

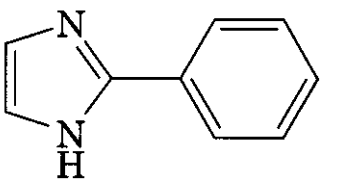
individual compounds within the mixture. However, signals from the protein are always present, which can pose a problem in interpreting spectra. An isotope-edited version of the diffusion experiment has been designed to avoid this problem, although labeled protein is required (168). Generally, there is no requirement for labeling of the protein target or for the protein resonances to be assigned and thus, in theory, there is no size limit on the proteins that can be screened by use of this method, although no information is obtained on the location of ligand binding. However, if the protein is large, then the transverse relaxation time may be too short to observe the bound ligands in the diffusion-edited spectrum (169). Only one sample, containing protein and ligands, is used to obtain both reference and screening data and therefore differences between the sample and reference spectra caused by addition of the ligands (pH, salt concentration, etc.) are avoided.

Diffusion-filtered NMR screening requires that there is a significant difference in observed translational diffusion between the free and bound states. The ligands are in fast exchange on the diffusion timescale and as a consequence the observed diffusion coefficient for binding ligands is an average between the free and bound diffusion values. Free ligands diffuse at a much faster rate than those in the bound state and thus only a small amount of free ligand has a considerable effect on the observed average diffusion coefficient. This effect may be significant enough to reduce the difference between binding and nonbinding ligands, making it more difficult to interpret results (169). It has also been demonstrated that chemical exchange and NOE can affect the interpretation of diffusion experiments and that these factors need to be taken into consideration (170, 171).

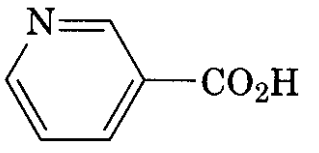
Shapiro and coworkers developed a methodology based on diffusion filtering, named "affinity NMR," that they have used to screen for binding (172–175). Diffusion-edited NMR experiments were able to identify two known binding tetrapeptide ligands of vancomycin from a mixture of 10 peptides (176). Hajduk et al. demonstrated the application of diffusion-editing experiments by differentiating ligands of stromelysin from a mixture containing nonbinding compounds (177).

4.1.4 Relaxation. Like diffusion, the transverse relaxation time (T_2) of molecules is also dependent on molecular size. Large molecules, such as proteins, have a short T_2 and hence exhibit broad NMR signals, whereas small molecules have a longer T_2 and hence narrower line widths. Therefore, if a small molecule ligand binds to a protein, its T_2 value will decrease and a line-broadening effect of bound ligand signals can be observed. Alternatively, a relaxation filter can be used to remove signals from molecules with a short T_2 value. Subtraction from a reference spectrum will result in a spectrum containing only those ligands that bind to the protein.

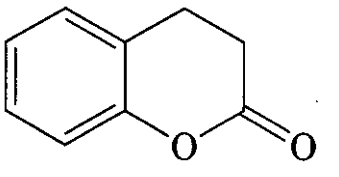
The ability to identify binding ligands using relaxation filters has been demonstrated using FKBP. A mixture of nine compounds consisting of one known ligand of FKBP, 2-phenylimidazole (**34**), and eight nonbinding compounds (e.g., 35–37 were screened and only signals from (**34**) were observed (177).



(34)

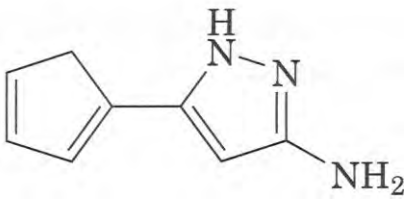


(35)



(36)

4.1.5 NOE. NOE experiments can also be used to identify ligands that bind to protein targets (178–180).Small molecules have a fast



(37)

tumbling rate and, as a consequence, generally exhibit small positive NOEs. In contrast, large molecules such as proteins generate strong negative NOEs because of their slow tumbling time. When a small molecule binds to a protein, its tumbling rate is slowed to that of the protein and it exhibits strong negative NOEs. On dissociation, these are transiently retained and are known as transfer NOEs (TrNOEs) (Fig. 12.33). TrNOEs and those arising directly from the free ligand can be distinguished by the rate of signal build up. Transfer NOEs accumulate significantly faster and therefore can be selected for by use of shorter mixing times in the NOE experiment (179).

In practice, a 2D NOESY spectrum of the mixture of potential ligands in the absence of protein is recorded and all molecules exhibit small positive NOEs. The experiment is then repeated in the presence of protein and molecules that bind display negative TrNOEs. Subtraction of the two spectra provides signals arising from only those compounds that bind. These TrNOEs can be interpreted to provide information regarding the bound conformation of the active ligands. However, when analyzing the conformational data care must be taken to ensure that the ligands are in fast exchange and that the observed TrNOEs are not affected by contributions from spin-diffusion (179). Relative binding affinities between ligands can also be determined by comparison of TrNOE signal strength but, again, the fast-exchange regime and spin-diffusion effects need to be taken into account (178,179). If all ligands are in fast exchange, the stronger binding ligands occupy more binding sites and thus give larger TrNOE intensities. Because of the need for an averaging effect, brought about by fast chemical exchange, TrNOE experiments are limited to those ligands with a K_D value from 10^{-3} to 10^{-7} M. The spectral properties of excess ligand in solution are evoked by small fractions of bound molecules, greatly enhancing sensitivity (178).

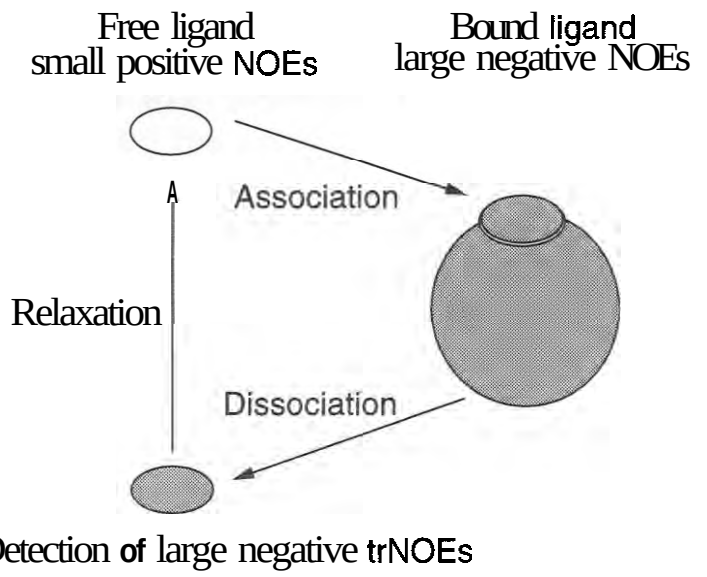
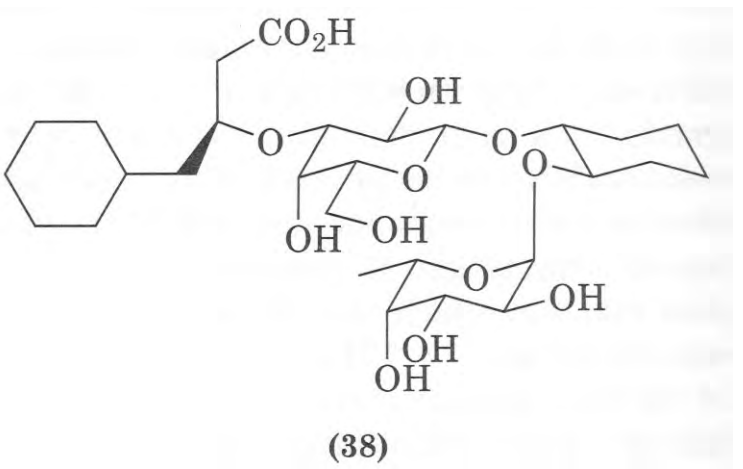


Figure 12.33. A schematic representation of the TrNOE experiment used to detect ligand binding (180). The free ligand (white ellipse) exhibits only small positive NOEs, although binding to the large protein target results in the generation of large negative TrNOEs. The appearance of these large negative TrNOE signals can be used to identify ligands within a mixture that are binding to the protein and also provide some information on the bound conformation of the ligand.

Transfer NOE experiments have been used to identify a bioactive disaccharide from a library of 15 mono- and disaccharides that bound to ***Aleuria aurantia*** agglutinin (179). Another study has described the identification of a silalyl Lewis mimetic **(38)** that binds to



(38)

E-selectin from a library of 10 compounds (178). As well as being used to detect binding, TrNOEs may also be used to determine bound ligand conformations, as described earlier in this chapter.

A second technique that uses NOEs to detect binding is NOE pumping. This method was designed to alleviate some of the problems associated with the diffusion-edited screening methods (169). Signals from ligand molecules are removed using a diffusion filter and then transfer of signal from the protein to bound ligands by NOE occurs. The inverse of this is possible (known as reverse NOE pumping), which uses a relaxation filter to attenuate the protein resonances, after which the signal is transferred to the protein by NOE. Ligands may lose signal either by relaxation (for a free ligand) or through relaxation and NOE transfer (for a bound ligand). Therefore by subtracting spectra (which is done internally to reduce subtraction artifacts) from experiments with and without NOE pumping to the protein, the binding ligands can be detected (181).

The ability of NOE and reverse NOE pumping to identify ligands has been demonstrated through the use of human serum albumin (HSA) and several known binding and nonbinding compounds (169, 181).

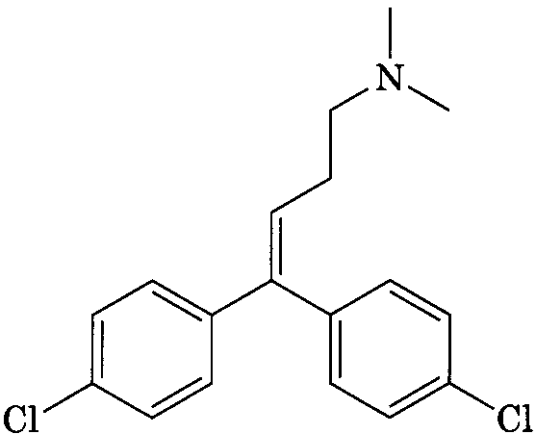
4.1.6 Spin **Labels.** Spin-spin relaxation rates are proportional to the product of the squares of the gyromagnetic ratios of the involved spins. The gyromagnetic ratio of an unpaired electron is significantly larger than that of a proton and therefore any spins influenced by this electron will have substantially shortened relaxation times. The resonances of protons that are within 15–20 Å from the unpaired electron will experience this effect and be significantly broadened. The introduction of a short spin-lock period will significantly reduce the intensity or quench these signals.

The spin-label method can be used as either a primary screening method or to identify a second ligand-binding site. The primary screening method requires residues around the binding pocket of the target to be spin labeled. Residues suitable for this labeling include lysine, cysteine, histidine, glutamate, aspartate, tyrosine, and methionine. Any ligands that bind to the protein in close proximity to the spin-labeled residues will be able to be identified. To screen for second-site ligand binding, the known first-site binding ligand is spin labeled. A reduced signal will be observed for any ligands that bind simultaneously and in close proximity to the first ligand-binding site. In addition, the degree of reduction in signal intensity gives an indication of the orientation of the second ligand in relation to the first, given that the effect of the spin label is inversely proportional to the distance separating the electron and proton. This information is valuable in the design of linkers to join the two ligands.

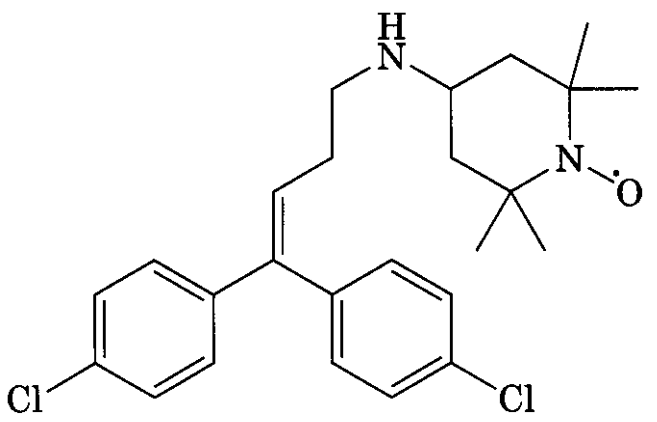
There are several advantages to using the spin-label screening method. Currently, it is the only method that can detect ligands that bind to the protein simultaneously, unlike other methods that can produce false positives if the first ligand-binding site is not fully saturated. The concentration of protein required for screening is relatively small ($\sim 10 \mu M$) because of the substantial enhancement of the relaxation rate by the spin label. The protein can also be unlabeled and partially purified and there is no molecular weight limit. The spin labels also quench protein signals, making interpretation of spectra easier. The experiment is easy to set up and analyze, making it amenable to automation. It is also insensitive to small changes in solvent conditions that can generate false positives in other methods. The information obtained on the orientation of ligands is also valuable and makes it an alternative to the chemical-shift perturbation methods when the proteins are large and NMR assignments have not been made.

A disadvantage of the method is the requirement for spin-labeled proteins and ligands. In addition, any ligands with slow dissociation rates will show no averaging of relaxation rates and therefore tightly binding compounds ($K_D < 10^{-6} M$) will produce false negatives. Protein spin labeling must occur adjacent but not within the binding site to minimize alteration of its binding properties.

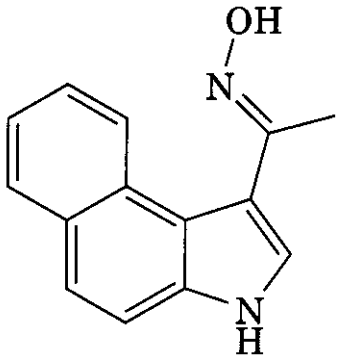
The antiapoptopic protein Bcl-xL is responsible for the reduced susceptibility of cancer cells to undergo apoptosis and is therefore a target for the development of new anticancer agents. The structure of a previously identified ligand for Bcl-xL **(39)** was modified to incorporate a TEMPO spin label (40). By use of spin-labeled (**40**), an eight-compound library was screened for simultaneous binding to Bcl-xL. From this library an aromatic ketoxime



(39)



(40)



(41)

(41) was identified as binding simultaneously with and in the vicinity of (40). Analysis of relaxation enhancements revealed that the protons around the indole ring were closest to the spin label.

4.2 Practical Considerations

4.2.1 Screening Approach. The first step in using NMR screening is to select a suitable screening method for the target protein being used. Table 12.8 lists the characteristics of each NMR screening method. The choice of experiment will be determined by the characteristics of the protein target and the information that is desired from the screen. For example, the SAR-by-NMR method is suitable only for small, easily expressed proteins because the NMR assignments for the target need to be known so the location of binding can be determined and a large amount of ^{15}N-enriched protein is required. If a simple "yes/no" answer on ligand binding is wanted, then the shorter, less resource intensive ligand-observed experiments (e.g., STD, diffusion-edited, or TrNOE) may suffice.

It is also important to determine the correct NMR solvent conditions for the screening procedure. These should facilitate good solubility, with little precipitation or aggregation, and acquisition of good quality data; maintain protein structure and activity; and provide a sufficient buffering effect to allow for ligands to be added. Two methods that permit the screening of a range of solvent conditions to determine without the need for a large amount of protein are the microdialysis button test (182) and the microdrop-screening method (183). A review on this subject has recently been published (184), to which the reader is referred for a more in-depth discussion on the subject of solvent conditions for NMR.

4.2.2 Library Design. Effective design and management of the ligand library to be used for screening is essential if successful results are to be obtained. The major considerations in library design are only briefly described here. There are a number of reviews that provide a more in-depth discussion of library design (15, 185).

4.2.2.1 ***Ligand Properties.*** Diversity of ligands is an important factor to consider in the design of a library for NMR screening and there are a number of factors to take into consideration. Although it would seem logical to maximize diversity, this may not always be the most efficient approach. If the system being studied exhibits neighborhood behavior, then maximizing diversity is a good option. Neighborhoods are regions of multidimensional molecular space defined by a set of molecular descriptors. By the choice of a molecule that is in the center of a neighborhood, it is possible, in theory, to represent all molecules within that molecular space. By spreading out the mole-

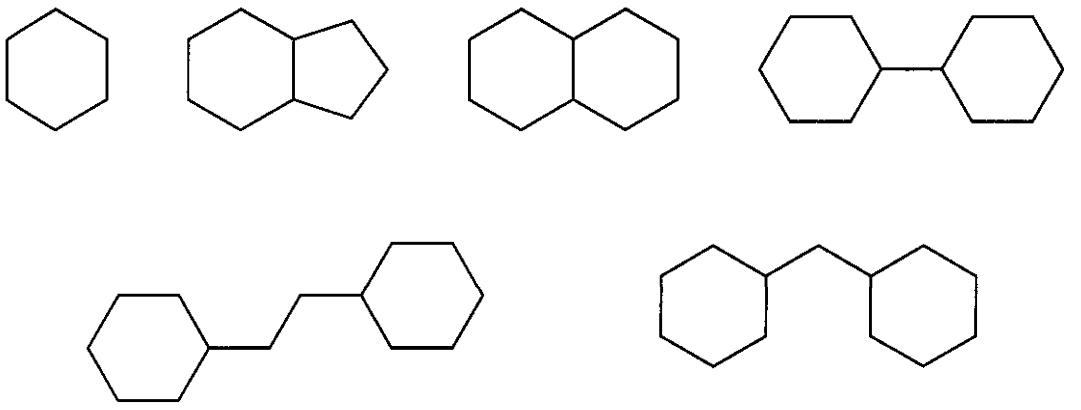


Figure 12.34. Examples of molecular frameworks from the SHAPES library.

cules that are selected for the library so that each neighborhood does not overlap, diversity is maximized.

However, if neighborhoods are only small then compound libraries must be very large so that the neighborhoods overlap and hence all molecular space is covered. In addition, some systems do not exhibit neighborhood behavior and relatively small changes to the structure of a compound may lead to large changes in its binding affinity for the target. Maximizing diversity may also be inefficient because many molecules do not possess physicochemical characteristics that are suitable as the basis for a drug. In practice, the more that is known about the drug target, the less diverse and more focused the library can be. However, if the library is too focused then some outlying "new" ligand type for the target being screened may be missed.

One strategy for library design is to select compounds that have druglike characteristics. A simple set of rules, determined by Lipinski and coworkers, for determining whether a compound is druglike is known as "the rule of 5." According to this set of criteria, the majority of orally available drugs have five or fewer hydrogen bond donors, 10 or fewer hydrogen bond acceptors, a log P of less than 5, and a molecular weight less than 500 (186). Additional factors that can be taken into consideration include the number of heavy atoms, rotatable bonds, and ring systems (187–189). Another study has revealed that there are a number of frameworks and side-chains that commonly occur in many drugs. Drug molecules, from the comprehensive medicinal database, were broken down into systems consisting of frameworks (Fig. 12.34) and side-chains. Analysis of these two structural features revealed that approximately 50% of all known drugs could be represented by only 32 different frameworks. When atom type and bond order were included in the analysis, 41 frameworks were found to describe 24% of all drugs (190). A similar analysis of side-chain frequency indicated that approximately 70% of all side chains present in the compound database analyzed were from the top 20 occurring side chains (191).

The presence of these common frameworks and side-chains has been exploited in the SHAPES methodology (7) for NMR screening. This strategy employs a small focused library based on these common frameworks and side chains to screen against protein targets through the use of relaxation and NOE experiments. The advantages of this approach are that the library is small and hence only relatively small amounts of protein are required and any hits from the library will possess druglike characteristics. However, a disadvantage of the method is that it is unlikely to yield new drug types, given that the library is based on known drug frameworks.

Diversity of molecular type is not the only factor that must be taken into account when designing a library to be used in an NMR screening program. Because the screening occurs in an aqueous solution, the organic compounds chosen for the library must demonstrate reasonable solubility in the aqueous conditions used. In general, compounds are dissolved in DMSO and then added to the protein solution at the appropriate concentration. Currently, there are no good methods for determining the solubility of a wide range of compounds before screening commences. A simple method is to dilute the DMSO solution in buffer and observe whether any precipitation or aggregation occurs. However, this method will not be suitable for compounds

that precipitate or aggregate over several hours, and solutions that appear clear may still contain high MW aggregates, which will cause false positives in experiments such as the diffusion, relaxation, and TrNOE methods (15, 185).

It is also preferable to choose ligands that are synthetically accessible and/or possess suitable moieties to build upon or link to other fragments. This is especially important in the SAR-by-NMR screening methodology because this relies on the ability to link individual fragments to form a more potent drug lead. If the ligands to be linked are not synthetically accessible or do not possess suitable linking functional groups then this process is severely hindered.

4.2.2.2 ***Mixture Design.*** The optimal number of compounds per mixture is dependent on the screening method. For ligand-observed experiments the limiting factor for the number of compounds in a mixture is spectral overlap. Ligands need to be chosen so that spectral overlap is minimized, making interpretation of the resulting data far simpler. In theory, protein-observed experiments could have a large number of compounds per mixture that would both minimize screening time and the requirement for large amounts of protein. However, because the experiments are protein observed then deconvolution of the mixtures and rescreening of each individual compound are required to identify any hits. Therefore, the number of compounds per mixture is dependent on the hit rate in the screening procedure, given that the greater the hit rate, the more deconvolution steps required and consequently more protein and spectrometer time are needed. The number of experiments required is at a minimum when the number of compounds is equal to $1/(\text{hit rate})^{1/2}$; thus, with a hit rate of 10% the optimal number of compounds per mixture is three (185). In addition to these factors, if the hit rate is high then it is likely that several compounds within a mixture containing a large number of compounds may compete for the same binding pocket, which may lead to false negatives.

In mixtures of organic compounds the possibility of interactions between compounds, such as reactions or ion pairing, is also possible and should be taken into consideration, especially when one uses large numbers of compounds per mixture. It has been demonstrated that in random mixtures of 10 compounds in DMSO, the probability of a reaction occurring between two of the mixture's components is 26%. This value can be reduced by careful selection of mixture components (e.g., separating acids from bases) to approximately 9% (192).

4.2.3 Hardware and Automation. Automation is a requirement if libraries containing a large number of compounds are to be screened. Technology has been developed that allows the automation of almost all steps of the NMR screening process from sample preparation through to data analysis (193).

The general setup for NMR screening consists of a robot for just-in-time preparation of each sample, which is then transferred to the magnet either through a flow system or as discrete samples on a rail system. There are several disadvantages in using a flow system, including the possibility of contamination of samples by previously screened compounds, the capillary line can be blocked if the protein or ligands precipitate or form aggregates, recovery of the sample is more laborious because it has been diluted, and cryoprobe technology (discussed later in this section) is not yet available in the flow system. Many of these problems can be overcome by using the discrete samples with the rail system.

Data acquisition is easily automated and there are several software packages that will automate data processing for 2D spectra. The processing of 1D spectra automatically is reported to be less reliable because of the large solvent signal and usually require manual adjustment of phasing (193). One of the most laborious tasks in NMR screening is the analysis and comparison of the resulting spectra. For 1D ligand-observed experiments difference methods (e.g., STD) provide the most reliable method for interpretation of results, in that the presence of signals in the spectra will correspond to the ligands that are binding. In 2D protein-observed experiments (e.g., $^{1}H/^{15}N$ HSQC) a more statistically rigorous analysis of changes in chemical shift is required and a discussion of this is beyond the scope of

this chapter. A more in-depth account of data analysis is provided by Ross and Senn (193).

Currently, approximately 50–100 samples can be screened per day and if mixtures contain 10 compounds each this provides a substantial throughput. This throughput rate will increase as technology improves, as has been demonstrated by the use of cryoprobes. Cryogenic NMR probes, where the preamplifier and radio frequency coils are cooled to low temperatures, can significantly increase the signal-to-noise ratio of an NMR spectrum. By use of these probes NMR data can be obtained in much faster times and by use of lower protein concentrations, which subsequently increases throughput, the total amount of protein needed to screen a library is reduced. Hajduk and coworkers (21) demonstrated the substantial improvements made through the use of a CryoProbe instead of a conventional probe in $^{1}H/^{15}N$ chemical-shift perturbation screening. Stromelysin (50 μM) was screened against mixtures of 100 compounds (50 μM each), facilitating the screening of more than 10,000 compounds in one day. The use of lower concentrations of both protein and ligands increases the stringency levels for the binding strength of ligands. At a protein/ligand concentration of 0.5 mM, ligands with dissociation constants in the millimolar range can be detected, although at a protein/ligand concentration of 50 μM this dissociation constant limit is reduced to approximately 0.15 mM. Although using higher protein/ligand concentrations can be advantageous when screening libraries containing small low affinity ligands, a higher stringency is required when screening large libraries, to reduce the number of hits obtained to a manageable number (21).

5 CONCLUSIONS

In this chapter we have given an overview of the two major approaches used in NMR and drug discovery, structure-based design and NMR-based screening. Both areas are flourishing and, together with more traditional uses of NMR, they demonstrate the versatility of NMR as a tool in medicinal chemistry. The power of NMR has been dramatically enhanced over the last decade by developments in both instruments and methodology. On the instrumental side, increases in magnetic field strengths and the development of cryoprobes have greatly increased sensitivity. Linkages of NMR to LC and MS have increased versatility. On the methods front there have been a range of new approaches discovered that will enhance the study of larger molecular complexes. Advances in protein expression and labeling have played a major role in stimulating the development of new NMR pulse sequences to extract information from such complexes.

6 ACKNOWLEDGMENTS

Work in our laboratory on NMR in drug design and development is supported by the Australian Research Council. D.J.C. is an ARC Professorial fellow. We thank Norelle Daly for assistance with some of the figures and Robyn Craik, Shaiyena Williams, and David Ireland for help in preparation of the manuscript.

REFERENCES

1. P. J. Hajduk, R. P. Meadows, and S. W. Fesik, *Science,* 278,497 (1997).
2. S. W. Fesik, *J. Biomol. NMR,* 3,261 (1993).
3. B. Stockman, *Prog. Nucl. Magn. Reson. Spectrosc.,* 33, 109 (1998).
4. P. J. Hajduk, R. P. Meadows, and S. W. Fesik,. *Q. Rev. Biophys.,* 32, 211 (1999).
5. G. C. Roberts, *Curr. Opin. Biotechnol.,* **10,** 42 (1999).
6. J. M. Moore, *Curr. Opin. Biotechnol.,* **10,** 54 (1999).
7. J. Fejzo, C. A. Lepre, J. W. Peng, G. W. Bemis, Ajay, M. A. Murcko, and J. M. Moore, *Chem. Biol.,* **6,** 755 (1999).
8. P. A. Keifer, *Curr. Opin. Biotechnol.,* **10,** 34 (1999).
9. G. C. Roberts, *Drug Discovery Today,* **5,** 230 (2000).
10. D. C. Fry and S. D. Emerson, *Drug Des. Disco ~.17,* 13 (2000).
11. D. J. Craik and M. J. Scanlon in G. A. Webb, Ed., Annual Reports on NMR Spectroscopy, Vol. 42, Academic Press, San Diego, 2000, pp. 115–173.
12. T. Diercks, M. Coles, and H. Kessler, *Curr. Opin. Chem.* Biol., 5,285 (2001).
13. R. P. Hicks, *Curr. Med. Chem.,* 8, 627 (2001).
14. M. Shapiro, *Farmaco,* **56,** 141 (2001).

15. J. W. Peng, C. A. Lepre, J. Fejzo, N. Abdul-Manan, and J. M. Moore, *Methods Enzymol.*, 338,202 (2001).

16. D. J. Craik, Ed., *NMR in DrugDesign*, CRC Press, Boca Raton, FL, 1996, pp. 1–476.

17. U. Holzgrabe, I. Wawer, and B. Diehl, *NMR Spectroscopy in Drug Development and Analysis*, Wiley-VCH, Weinheim, Germany, 1999, pp. 1–299.

18. A. E. Derome, *Modern NMR Techniques for Chemistry Research*, Pergamon, New York, 1987.

19. H. Gunther, *NMR Spectroscopy—An Introduction*, John Wiley & Sons, Chichester, UK, 1980, pp. 1–436.

20. K. Pervushin, R. Riek, G. Wider, and K. Wuthrich, *Proc. Natl. Acad. Sci. USA*, 94, 12366 (1997).

21. P. Hajduk, T. Gerfin, J. Boehlen, M. Haberli, D. Marek, and S. Fesik, *J. Med. Chem.*, 42, 2315 (1999).

22. P. A. Keifer, *Prog. Drug Res.*, 55, 137 (2000).

23. K. Wuthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York, 1986, pp. 1–292.

24. C. E. Heading, *Drugs*, 4,339 (2001).

25. D. S. Wishart, B. D. Sykes, and F. M. Richards, *J. Mol. Biol.*, 222,311 (1991).

26. D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes, *J. Biomol. NMR*, 5, 67 (1995).

27. K. J. Nielsen, L. Thomas, R. J. Lewis, P. F. Alewood, and D. J. Craik, *J. Mol. Biol.*, 263, 297 (1996).

28. L. K. MacLachlan, D. A. Middleton, A. J. Edwards, and D. G. Reid in D. G. Reid, Ed., Protein NMR Techniques, Vol. 60, Humana Press, Totowa, NJ, 1997, pp. 337–362.

29. D. S. Wishart, B. D. Sykes, and F. M. Richards, *Biochemistry*, 31, 1647 (1992).

30. F. Dasgupta, A. K. Mukherjee, and N. Gangadhar, *Curr. Med. Chem.*, 9, 549 (2002).

31. D. J. Craik, K. J. Nielsen, and K. A. Higgins in G. A. Webb, Ed., Annual Reports on NMR Spectroscopy, Vol. 32, Academic Press, San Diego, 1995, pp. 143–213.

32. S. R. Krystek Jr., D. A. Bassolino, J. Novotny, C. Chen, T. M. Marschner, and N. H. Andersen, *FEBS Lett.*, 281, 212 (1991).

33. N. H. Andersen, C. P. Chen, T. M. Marschner, S. R. Krystek Jr., and D. A. Bassolino, *Biochemistry*, 31, 1280 (1992).

34. V. Saudek, J. Hoflack, and J. T. Pelton, *FEBS Lett.*, 257, 145 (1989).

35. S. Endo, H. Inooka, Y. Ishibashi, C. Kitada, E. Mizuta, and M. Fujino, *FEBS Lett.*, 257, 149 (1989).

36. R. G. Mills, S. I. O'Donoghue, R. Smith, and G. F. King, *Biochemistry*, 31,5640 (1992).

37. S. Munro, D. Craik, C. McConville, J. Hall, M. Searle, W. Bicknell, D. Scanlon, and C. Chandler, *FEBS Lett.*, 278, 9 (1991).

38. M. D. Reily and J. B. Dunbar Jr., *Biochem. Biophys. Res. Commun.*, 178,570 (1991).

39. H. Tamaoki, Y. Kobayashi, S. Nishimura, T. Ohkubo, Y. Kyogoku, K. Nakajima, S. Kumagaye, T. Kimura, and S. Sakakibara, *Protein Eng.*, 4,509 (1991).

40. A. Aumelas, L. Chiche, S. Kubo, N. Chino, H. Tamaoki, and Y. Kobayashi, *Biochemistry*, 34, 4546 (1995).

41. A. Aumelas, L. Chiche, E. Mahe, D. Le-Nguyen, P. Sizun, P. Berthault, and B. Perly, *Int. J. Pept. Protein Res.*, 37, 315 (1991).

42. D. C. Dalgarno, L. Slater, S. Chackalamannil, and M. M. Senior, *Int. J. Pept. Protein Res.*, 40, 515 (1992).

43. Y. Boulanger, E. Biron, A. Khiat, and A. Fournier, *J. Pept. Res.*, 53,214 (1999).

44. K. Arvidsson, T. Nemoto, Y. Mitsui, S. Ohashi, and H. Nakanishi, *Eur. J. Biochem.*, 257, 380 (1998).

45. C. M. Hewage, L. Jiang, J. A. Parkinson; R. Ramage, and I. H. Sadler, *J. Pept. Sci.*, 3,415 (1997).

46. B. A. Wallace, R. W. Janes, D. A. Bassolino, and S. R. Krystek Jr., *Protein Sci.*, 4, 75 (1995).

47. M. Coles, V. Sowemimo, D. Scanlon, S. L. Munro, and D. J. Craik, *J. Med. Chem.*, 36, 2658 (1993).

48. D. J. Detlefsen, S. E. Hill, S. H. Day, and M. S. Lee, *Curr. Med. Chem.*, 6, 353 (1999).

49. W. C. Patt, J. J. Edmunds, J. T. Repine, K. A. Berryman, B. R. Reisdorph, C. Lee, M. S. Plummer, A. Shahripour, S. J. Haleen, J. A. Keiser, M. A. Flynn, K. M. Welch, E. E. Reynolds, R. Rubin, B. Tobias, H. Hallak, and A. M. Doherty, *J. Med. Chem.*, 40, 1063 (1997).

50. W. C. Patt, X. M. Cheng, J. T. Repine, C. Lee, B. R. Reisdorph, M. A. Massa, A. M. Doherty, K. M. Welch, J. W. Bryant, M. A. Flynn, D. M. Walker, R. L. Schroeder, S. J. Haleen, and J. A. Keiser, *J. Med. Chem.*, 42, 2162 (1999).

51. S. L. Munro, P. R. Andrews, D. J. Craik, and D. J. Gale, *J. Pharm. Sci.*, 75, 133 (1986).

52. B. M. Duggan and D. J. Craik, *J. Med. Chem.*, **39, 4007 (1996).**

53. B. M. Duggan and D. J. Craik, **J. *Med. Chem.*, 40, 2259** (1997).

54. M. G. Casarotto and D. J. Craik, *J. Pharm. Sci.*, **90, 713 (2001).**

55. R. Abseher, L. Horstink, C. W. Hilbers, and M. Nilges, ***Proteins: Struct., Funct., Genet.*, 31, 370 (1998).**

56. J. Gehrmann, P. F. Alewood, and D. J. Craik, *J. Mol. Biol.*, **278, 401 (1998).**

57. J. Balbach, S. Seip, H. Kessler, M. Scharf, N. Kashani-Poor, and J. W. Engels, ***Proteins*, 33, 285 (1998).**

58. D. J. Craik, B. M. Duggan, and S. L. A. Munro in M. I. Choudary, Ed., Biological Inhibitors, Vol. **2,** Harwood Academic, Amsterdam, **1996,** pp. **255302.**

59. R. **L.** Wagner, J. W. Apriletti, M. E. McGrath, B. L. West, J. D. Baxter, and R. J. Fletterick, *Nature*, **378, 690** (1995).

60. D. J. Gale, D. J. Craik, and R. T. C. Brownlee, ***Magn. Reson. Chem.*, 26, 275 (1988).**

61. W. Bourguet, M. Ruff, P. Chambon, H. Gronemeyer, and D. Moras, ***Nature*, 375, 377 (1995).**

62. J. P. Renaud, N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer, and D. Moras, ***Nature*, 378, 681 (1995).**

63. J. Feeney and B. Birdsall in G. C. K. Roberts, Ed., ***NMR of Macromolecules: A Practical Approach,*** Oxford University Press, Oxford, UK, **1993,** pp. **181-215.**

64. J. Feeney in I. Bertini, H. Molinari, and N. Niccolai, Eds., ***NMR and Biomolecular Structure,*** VCH, New York, **1991,** pp. **189-205.**

65. J. Feeney, ***Biochem. Pharmacol.*, 40, 141 (1990).**

66. **K.** J. Nielsen, D. Adams, L. Thomas, T. Bond, P. F. Alewood, D. J. Craik, and R. J. Lewis, *J. Mol. Biol.*, **289, 1405** (1999).

67. A. P. Campbell and B. D. Sykes, ***Annu. Rev. Biophys. Biomol. Struct.*, 22,** 99 (1993).

68. **B.** D. Sykes, ***Curr. Opin. Biotechnol.*, 4, 392 (1993).**

69. D. J. Craik and K. A. Higgins in G. A. Webb, Ed., Annual Reports on NMR Spectroscopy, Vol. **22,** Academic Press, London, **1990,** pp. 61–**138.**

70. G. Bertho, J. Gharbi-Benarous, M. Delaforge, and J. P. Girault, ***Bioorg. Med. Chem.*, 6, 209** (1998).

71. R. E. Hubbard, ***Curr. Opin. Biotechnol.*, 8, 696** (1997).

72. G. Wider and K. Wuthrich, ***Curr. Opin. Struct. Biol.*, 9, 594 (1999).**

73. G. M. Clore and A. M. Gronenborn, ***Curr. Opin. Chem. Biol.*, 2, 564** (1998).

74. **N.** Tjandra and A. Bax, ***Science*, 278, 1111** (1997).

75. G. F. King and J. P. Mackay in D. J. Craik, Ed., ***NMR in Drug Design,*** CRC Press, Boca Raton, FL, **1996,** pp. **101-200.**

76. N. Tjandra, A. M. Garrett, A. M. Gronenborn, A. Bax, and G. M. Clore, ***Nat. Struct. Biol.*, 4, 443** (1997).

77. A.M. Edwards, C. H. Arrowsmith, D. Christendat, A. Dharamsi, J. D. Friesen, J. F. Greenblatt, and M. Vedadi, *Nat. **Struct. Biol.*, 7** (Suppl.), **970** (2000).

78. D. Christendat, **A** Yee, A. Dharamsi, Y. Kluger, M. Gerstein, C. H. Arrowsmith, and A. M. Edwards, *Prog. **Biophys. Mol. Biol.*, 73, 339** (2000).

79. L. E. Kay, ***Nat. Struct. Biol.*, 5, 513 (1998).**

80. F. M. Marassi and S. J. Opella, ***Curr. Opin. Struct. Biol.*, 8, 640 (1998).**

81. A. Watts, ***Curr. Opin. Biotechnol.*, 10, 48 (1999).**

82. L.-Y. Lian and G. C. K. Roberts in G. C. K. Roberts, Ed., ***NMR of Macromolecules: A Practical Approach,*** Oxford University Press, Oxford, UK, **1993.**

83. L. Dugad and J. T. Gerig, ***Biochemistry*, 27, 4310 (1988).**

84. E. I. Hyde, B. Birdsall, G. ***C.*** Roberts, J. Feeney, and A. S. V. Burgen, ***Biochemistry*, 19, 3746 (1980).**

85. J. Feeney, J. G. Batchelor, J. P. Albrand, and G. C. K. Roberts, **J. *Magn. Reson.*, 33, 519 (1979).**

86. S. Pavlopoulos, M. Rose, G. Wickham, and D. J. Craik, ***Anticancer Drug Des.*, 10, 623 (1995).**

87. G. J. Pelton and D. E. Wemmer, *Proc. **Natl. Acad. Sci. USA*, 86, 5723 (1990).**

88. W. Leupin, W. J. Chazin, S. Hyberts, W. A. Denny, and K. Wuthrich, ***Biochemistry*, 25, 5902 (1986).**

89. S. M. Chen, W. Leupin, M. Rance, and W. J. Chazin, ***Biochemistry*, 31, 4406 (1992).**

90. R. E. Klevit, D. E. Wemmer, and B. R. Reid, ***Biochemistry*, 25, 3296 (1986).**

91. **D.** J. Patel and L. Shapiro, *J. **Biol. Chem.*, 261, 1230** (1986).

92. M. S. Searle and K. J. Embrey, ***Nucleic Acids Res.*, 18, 3753 (1990).**

93. S. W. Fesik, J. R. Luly, J. W. Erickson, and C. Abad-Zapatero, Biochemistry, **27,8297 (1988).**
94. C. Zwahlen, P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat, and L. E. Kay, *J. Am.* Chem. *Soc.*, **119,6711 (1997).**
95. M. J. Gradwell and J. Feeney, J. Biomol. NMR, **7**, 48 **(1996).**
96. K. D. Harshman and P. B. Dervan, Nucleic Acids Res., **13, 4825 (1985).**
97. J. G. Pelton and D. E. Wemmer, Biochemistry, **27,8088 (1988).**
98. J. G. Pelton and D. E. Wemmer, Proc. Natl. *Acad.* Sci. USA, **86,5723 (1989).**
99. M. Coll, J. Aymami, G. A. van der Marel, J. H. van Boom, A. Rich, and A. H. Wang, Biochemistry, **28,310 (1989).**
100. P. E. Pjura, K. Grzeskowiak, and R. E. Dickerson, *J.* Mol. Biol., **197,257 (1987).**
101. M. K. Teng, N. Usman, C. A. Frederick, and A. H. Wang, Nucleic Acids Res., **16, 2671 (1988).**
102. M. A. Carrondo, M. Coll, J. Ayrnami, A. H. Wang, G. A. van der Marel, J. H. van Boom, and A. Rich, Biochemistry, **28, 7849 (1989).**
103. J. R. Quintana, A. A. Lipanov, and R. E. Dickerson, Biochemistry, **30, 10294 (1991).**
104. J. A. Parkinson, J. Barber, K. T. Douglas, J. Rosamond, and D. Sharples, Biochemistry, **29,** 10181 **(1990).**
105. K. J. Embrey, M. S. Searle, and D. J. Craik, Eur. J. Biochem., **211,437 (1993).**
106. A. Fede, A. Labhardt, W. Bannwarth, and W. Leupin, Biochemistry, **30, 11377 (1991).**
107. A. Fede, M. Billeter, W. Leupin, and K. Wuthrich, Structure, **1, 177 (1993).**
108. T. Taga, H. Tanaka, T. Goto, and S. Tada, *Acta Crystallogr.*, **C43**, **751 (1987).**
109. B. E. Bierer, P. K. Somers, T. J. Wandless, S. J. Burakoff, and S. L. Schreiber, Science, **250, 556 (1990).**
110. P. Karuso, H. Kessler, and D. F. Mierke, *J. Am.* Chem. *Soc.*, **112,9434 (1990).**
111. S. W. Fesik, R. T. Gampe Jr., T. F. Holzman, D. A. Egan, R. Edalji, J. R. Luly, R. Simmer, R. Helfrich, V. Klahore, and D. H. Rich, Science, **250, 1406 (1990).**
112. S. W. Fesik, R. T. Gampe Jr., H. L. Eaton, G. Gemmecker, E. T. Olejniczak, P. Neri, T. F. Holzman, D. A. Egan, R. Edalji, R. Simmer, R. Helfrich, J. Hochlowski, and M. Jackson, Biochemistry, **31,6574 (1991).**
113. C. Weber, G. Wider, K. von Freyberg, R. Truber, W. Braun, H. Widner, and K. Wuetrich, Biochemistry, **30, 6564 (1991).**
114. M. K. Rosen, R. F. Standaert, A. Galat, M. Nakatsuka, and S. L. Schreiber, Science, **248,863 (1990).**
115. T. J. Wandless, S. W. Michnick, M. K. Rosen, M. Karplus, and S. L. Schreiber, J. Am. Chem. *Soc.*, **113,2339** (1991).
116. A. M. Petros, R. T. Gampe Jr., G. Gemmecker, P. Neri, T. F. Holzman, R. Edalji, J. Hochlowski, M. Jackson, J. McAlpine, J. R. Luly, et al., J. Med. Chem., **34,2925 (1991).**
117. G. D. van Duyne, R. F. Standaert, M. Karplus, S. L. Schreiber, and J. Clardy, Science, **252, 839** (1991).
118. S. W. Michnick, M. K. Rosen, T. J. Wandless, M. Karplus, and S. L. Schreiber, Science, **252, 836 (1991).**
119. J. M. Moore, D. A. Peattie, M. J. Fitzgibbon, and J. A. Thomson, Nature, **351,248 (1991).**
120. R. P. Meadows, D. G. Nettesheim, R. X. Xu, E. T. Olejniczak, A. M. Petros, T. F. Holzman, J. Severin, E. Gubbins, H. Smith, and S. W. Fesik, Biochemistry, **32, 754 (1993).**
121. J. W. Cheng, C. A. Lepre, S. P. Chambers, J. R. Fulghum, J. A. Thomson, and J. M. Moore, Biochemistry, **32,9000 (1993).**
122. J. W. Cheng, C. A. Lepre, and J. M. Moore, Biochemistry, **33,4093 (1994).**
123. P. R. Gooley, B. A. Johnson, A. I. Marcy, G. C. Cuca, S. P. Salowe, W. K. Hagmann, C. K. Esser, and J. P. Springer, Biochemistry, **32, 13098 (1993).**
124. P. R. Gooley, J. F. O'Connell, A. I. Marcy, G. C. Cuca, S. P. Salowe, B. L. Bush, J. D. Hermes, C. K. Esser, W. K. Hagmann, J. P. Springer, and B. A. Johnson, *Nat.* Struct. Biol., **1, 111** (1994).
125. P. R. Gooley, J. F. O'Connell, A. I. Marcy, G. C. Cuca, M. G. Axel, C. G. Caldwell, W. K. Hagmann, and J. W. Becker, *J.* Biomol. NMR, **7**, 8 (1996).
126. S. R. Van Doren, A. V. Kurochkin, Q.-Z. Ye, L. L. Johnson, D. J. Hupe, and E. R. P. Zuiderweg, Biochemistry, **32,13109 (1993).**
127. S. R. Van Doren, A. V. Kurochkin, W. Hu, Q.-Z. Ye, L. L. Johnson, D. J. Hupe, and E. R. Zuiderweg, Protein Sci., **4,2487** (1995).
128. M. A. McCoy, M. J. Dellwo, D. M. Schneider, T. M. Banks, J. Falvo, K. J. Vavra, A. M. Mathiowetz, M. W. Qoronfleh, R. Ciccarelli, E. R. Cook, T. A. Pulvino, R. C. Wahl, and H. Wang, *J.* Biomol. NMR, **9, 11** (1997).
129. F. J. Moy, M. R. Pisano, P. K. Chanda, C. Urbano, L. M. Killar, M. L. Sung, and R. Powers, *J. Biomol.* NMR, **10, 9 (1997).**

130. E. J. Jacobsen, M. A. Mitchell, S. K. Hendges, K. L. Belonga, L. L. Skaletzky, L. S. Stelzer, T. J. Lindberg, E. L. Fritzen, H. J. Schostarez, T. J. O'Sullivan, L. L. Maggiora, C. W. Stuchly, A. L. Laborde, M. F. Kubicek, R. A. Poorman, J. M. Beck, H. R. Miller, G. L. Petzold, P. S. Scott, S. E. Truesdell, T. L. Wallace, J. W. Wilks, C. Fisher, L. V. Goodman, P. S. Kaytes, et al., ***J. Med. Chem.*, 42, 1525** (1999).

131. N. C. Gonnella, R. Bohacek, X. Zhang, I. Kolossvary, C. G. Paris, R. Melton, C. Winter, S. I. Hu, and V. Ganu, ***Proc. Natl. Acad. Sci. USA*, 92,462 (1995).**

132. V. I. Polshakov, B. Birdsall, T. A. Frenkiel, A. R. Gargaro, and J. Feeney, ***Protein Sci.*, 8, 467 (1999).**

133. P. M. Nieto, B. Birdsall, W. D. Morgan, T. A. Frenkiel, A. R. Gargaro, and J. Feeney, ***FEBS Lett.*, 405, 16 (1997).**

134. A. Wlodawer, M. Miller, M. Jaskolski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. Kent, ***Science*, 245,616 (1989).**

135. A. Wlodawer and J. W. Erickson, ***Annu. Rev. Biochem.*, 62,543 (1993).**

136. A. Wlodawer and J. Vondrasek, ***Annu. Rev. Biophys. Biomol. Struct.*, 27, 249 (1998).**

137. D. J. Kempf and H. L. Sham, ***Curr. Pharm. Des.*, 2,225 (1996).**

138. T. Yamazaki, A. P. Hinck, Y. X. Wang, L. K. Nicholson, D. A. Torchia, P. Wingfield, S. J. Stahl, J. D. Kaufman, C. H. Chang, P. J. Domaille, and P. Y. Lam, ***Protein Sci.*, 5, 495 (1996).**

139. P. Y. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, et al., ***Science*, 263,380 (1994).**

140. S. Grzesiek, A. Bax, L. K. Nicholson, T. Yamazaki, P. T. Wingfield, S. J. Stahl, C. J. Eyermann, D. A. Torchia, C. N. Hodge, P. Y. Lam, P. K. Jadhav, and C. H. Chang, ***J. Am. Chem. Soc.*, 116,1581 (1994).**

141. Y. X. Wang, D. I. Freedberg, S. Grzesiek, D. A. Torchia, P. T. Wingfield, J. D. Kaufman, S. J. Stahl, C. H. Chang, and C. N. Hodge, ***Biochemistry*, 35, 12694 (1996).**

142. G. Otting, E. Liepinsh, and K. Wuthrich, ***Science*, 254,974 (1991).**

143. T. Yamazaki, L. K. Nicholson, P. Wingfield, S. J. Stahl, J. D. Kaufman, P. J. Domaille, and D. A. Torchia, ***J. Am. Chem. Soc.*, 116, 10791 (1994).**

144. E. T. Baldwin, T. N. Bhat, S. Gulnik, B. Liu, I. A. Topol, Y. Kiso, T. Mimoto, H. Mitsuya, and J. W. Erickson, ***Structure*, 3,581 (1995).**

145. L. K. Nicholson, T. Yamazaki, D. A. Torchia, S. Grzesiek, A. Bax, S. J. Stahl, J. D. Kaufman, P. T. Wingfield, P. Y. Lam, P. K. Jadhav, et al., ***Nat. Struct. Biol.*, 2,274 (1995).**

146. J. R. Rose, R. Salto, and C. S. Craik, ***J. Biol. Chem.*, 268,11939 (1993).**

147. D. I. Freedberg, Y. X. Wang, S. J. Stahl, J. D. Kaufman, P. Wingfield, Y. Kiso, and D. A. Torchia, ***J. Am. Chem. Soc.*, 120, 7916** (1998).

148. E. Katoh, T. Yamazaki, Y. Kiso, P. Wingfield, S. J. Stahl, J. D. Kaufman, and D. A. Torchia, ***J. Am. Chem. Soc.*, 121,2607 (1999).**

149. D. L. Winslow, S. Stack, R. King, H. Scarnati, A. Bincsik, and M. J. Otto, ***AIDS Res. Hum. Retroviruses*, 11, 107 (1995).**

150. J. H. Condra, W. A. Schleif, O. M. Blahy, L. J. Gabryelski, D. J. Graham, J. C. Quintero, A. Rhodes, H. L. Robbins, E. Roth, M. Shivaprakash, et al., ***Nature*, 374,569 (1995).**

151. S. B. Shuker, P. J. Hajduk, R. P. Meadows, and S. W. Fesik, ***Science*, 274, 1531 (1996).**

152. E. Olejniczak, P. Hajduk, P. Marcotte, D. Nettesheim, R. Meadows, R. Edalji, T. Holzman, and S. Fesik, ***J. Am. Chem. Soc.*, 119, 5828 (1997).**

153. P. J. Hajduk, J. Dinges, J. M. Schkeryantz, D. Janowick, M. Kaminski, M. Tufano, D. J. Augeri, A. Petros, V. Nienaber, P. Zhong, R. Hammond, M. Coen, B. Beutel, L. Katz, and S. W. Fesik, ***J. Med. Chem.*, 42,3852 (1999).**

154. P. Hajduk, G. Sheppard, D. Nettesheim, E. Olejniczak, S. Shuker, R. Meadows, D. Steinman, G. Carrera, P. Marcotte, J. Severin, K. Walter, H. Smith, E. Gubbins, R. Simmer, T. Holzman, D. Morgan, S. Davidsen, J. Summers, and S. Fesik, ***J. Am. Chem. Soc.*, 119, 5818 (1997).**

155. P. Hajduk, M. Zhou, and S. Fesik, ***Bioorg. Med. Chem. Lett.*, 9,2403 (1999).**

156. G. Liu, J. R. Huth, E. T. Olejniczak, R. Mendoza, P. DeVries, S. Leitza, E. B. Reilly, G. F. Okasinski, S. W. Fesik, and T. W. von Geldern, ***J. Med. Chem.*, 44,1202 (2001).**

157. H. Mao, P. J. Hajduk, R. Craig, R. Bell, T. Borre, and S. W. Fesik, ***J. Am. Chem. Soc.*, 123, 10429 (2001).**

158. A. Ross, G. Schlotterbeck, W. Klaus, and H. Senn, ***J. Biomol. NMR*, 16, 139 (2000).**

159. F. J. Moy, K. Haraki, D. Mobilio, G. Walker, R. Powers, K. Tabei, H. Tong, and M. M. Siegel, ***Anal. Chem.*, 73, 571 (2001).**

160. M. Mayer and B. Meyer, *Angew. Chem. Znt. Ed. Engl.,* 38, 1784 (1999).
161. J. Klein, R. Meinecke, M. Mayer, and B. Meyer, *J. Am. Chem. Soc.*, 121, 5336 (1999).
162. W. Hellebrandt, T. Haselhorst, T. Kohli, E. Baurnl, and T. Peters, *J. Carbohydr. Chem.,* 19, 769 (2000).
163. M. Vogtherr and T. Peters, *J. Am. Chem. Soc.*, 122, 6093 (2000).
164. H. Maaheimo, P. Kosma, L. Brade, H. Brade, and T. Peters, *Biochemistry,* 39,12778 (2000).
165. M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 123, 6108 (2001).
166. R. Meinecke and B. Meyer, *J. Med. Chem.,* 44, 3059 (2001).
167. C. Dalvit, P. Pevarello, M. Tato, M. Veronesi, A. Vulpetti, and M. Sundstrom, *J. Biomol. NMR,* 18, 65 (2000).
168. N. Gonnella, M. Lin, M. J. Shapiro, J. R. Wareing, and X. Zhang, *J. Magn. Reson.*, 131, 336 (1998).
169. A. Chen and M. J. Shapiro, *J. Am. Chem. Soc.*, 120,10258 (1998).
170. A. Chen, C. S. Johnson Jr., M. Lin, and M. J. Shapiro, *J. Am. Chem. Soc.*, 120,9094 (1998).
171. A. Chen and M. J. Shapiro, *J. Am. Chem. Soc.*, 121,5338 (1999).
172. A. Chen and M. J. Shapiro, *Anal. Chem.,* 71, 669A (1999).
173. M. Lin and M. J. Shapiro, *J. Org. Chem.,* 61, 7617 (1996).
174. M. Lin, M. J. Shapiro, and J. R. Wareing, *J. Am. Chem. Soc.,* 119,5249 (1997).
175. M. Lin, M. J. Shapiro, and J. R. Wareing, *J. Org. Chem.,*62,8930 (1997).
176. K. Bleicher, M. Lin, M. J. Shapiro, and J. R. Wareing, *J. Org. Chem.,*63,8486 (1998).
177. P. J. Hajduk, E. T. Olejniczak, and S. W. Fesik, *J. Am. Chem. Soc.*, 119, 12257 (1997).
178. D. Henrichson, B. Ernst, J. L. Magnani, W. Wang, B. Meyer, and T. Peters, *Angew. Chem. Int. Ed. Engl.,* 38, 98 (1999).
179. B. Meyer, T. Weimar, and T. Peters, *Eur. J. Biochem.,* 246, 705 (1997).
180. M. Mayer and B. Meyer, *J. Med. Chem.,* 43, 2093 (2000).
181. A. Chen and M. J. Shapiro, *J. Am. Chem. Soc.*, 122,414 (2000).
182. S. Bagby, K. I. Tong, D. Liu, J. R. Alattia, and M. Ikura, *J. Biomol. NMR,* 10, 279 (1997).
183. C. Lepre and J. Moore, *J. Biomol. NMR,* 12, 493 (1998).
184. S. Bagby, K. I. Tong, and M. Ikura, *Methods Enzymol.,* 339, 20 (2001).
185. C. A. Lepre, *Drug Discovery Today,* 6, 133 (2001).
186. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, *Adv. Drug Delivery Rev.,* 46, 3 (2001).
187. A. K. Ghose, V. N. Viswanadhan, and J. J. Wendoloski, *J. Comb. Chem.,* 1, 55 (1999).
188. T. I. Oprea, J. Gottfries, V. Sherbukhin, P. Svensson, and T. C. Kuhler, *J. Mol. Graph. Model.,* 18,512 (2000).
189. J. Xu and J. Stevenson, *J. Chem. Inf. Comput. Sci.,* 40, 1177 (2000).
190. G. Bemis and M. Murcko, *J. Med. Chem.,* 39, 2887 (1996).
191. G. Bernis and M. Murcko, *J. Med. Chem.,* 42, 5095 (1999).
192. M. Hann, B. Hudson, X. Lewell, R. Lifely, L. Miller, and N. Ramsden, *J. Chem. Inf. Comput. Sci.,* 39,897 (1999).
193. A. Ross and H. Senn, *Drug Discovery Today,* 6, 583 (2001).