CHAPTER

# 10

# SEQUENTIAL ASSIGNMENT, STRUCTURE DETERMINATION, AND OTHER APPLICATIONS

The first nine chapters of this book have focused on the theory of NMR spectroscopy and on the experimental methods used for multidimensional NMR spectroscopy of proteins. Analyses of the information contained in the NMR spectra of proteins are the subjects of this final chapter. Exhaustive coverage of these topics would comprise an entire additional text; instead, this chapter has three aims: (*i*) to indicate the types of analyses commonly performed using NMR spectra of proteins, (*ii*) to rationalize usage of the experiments described in Chapters 6–9 for a thorough investigation of proteins by NMR spectroscopy, and (*iii*) to provide a bibliography of the primary literature to augment the discussions of aims (*i*) and (*ii*).

## 10.1 Resonance Assignment Strategies

In the initial stage of any investigation by NMR spectroscopy, each resonance in the NMR spectrum must be associated with a specific nucleus in the molecule under investigation. Resonance assignments must be *sequence specific*: each resonance must be assigned to a spin in a particular amino acid residue in the protein sequence. NMR spectroscopy provides three types of information useful for spectral assignments: through-bond interactions (*via* scalar couplings), through-space interactions (*via* dipolar couplings), and chemical environment (*via* isotropic chemical shifts). The strategies employed for resonance assignments depend upon whether only homonuclear $^1$H NMR spectra are available (unlabeled proteins) or whether $^{13}$C and $^{15}$N heteronuclear correlation spectra are available (isotopically labeled proteins).

### 10.1.1 $^1$H Resonance Assignments for Unlabeled Proteins

The procedures for obtaining $^1$H resonance assignments are based upon the following critical observation: with few exceptions, correlations resulting from $^1$H–$^1$H scalar couplings normally are only observed between $^1$H nuclei separated by two or three bonds in proteins. Cross-peaks in $^1$H homonuclear correlation NMR spectra occur between $^1$H spins within the same amino acid residue or *spin system*. Cross-peaks do not occur between $^1$H spins in different residues, because the inter-residue $^4J_{H^N_{i+1}H^\alpha_i}$ coupling constant is negligible. Therefore, scalar correlation experiments, such as COSY (Section 6.2), MQF-COSY (Section 6.3), MQ spectroscopy (Section 6.4), and TOCSY (Section 6.5), are used to identify resonance positions within each amino acid spin system, and the NOESY experiment (Section 6.6) is used to sequentially connect the amino acid spin systems. Two-dimensional NOESY and TOCSY experiments also can be combined to yield homonuclear 3D experiments, as described in Section 6.7.

Initially, $^1$H resonances are categorized on the basis of their chemical shifts. As indicated in Fig. 10.1, for the vast majority of residues in a protein, backbone amide $^1H^N$ spins resonate between 10.0 and 7.0 ppm, aromatic $^1$H spins resonate between 8.0 and 6.5 ppm (with the exception of Trp and His $^1H^{\varepsilon 1}$ spins), backbone $^1H^\alpha$ spins resonate between 6.0 and 3.5 ppm, aliphatic side chain methine and methylene $^1$H spins resonate between 3.5 and 1.0 ppm, and methyl $^1$H spins (other than Met $^1H^\varepsilon$ resonate at chemical shifts less than 2.0 ppm. Random coil $^1$H chemical shifts have been determined for unstructured peptides (*1–3*).

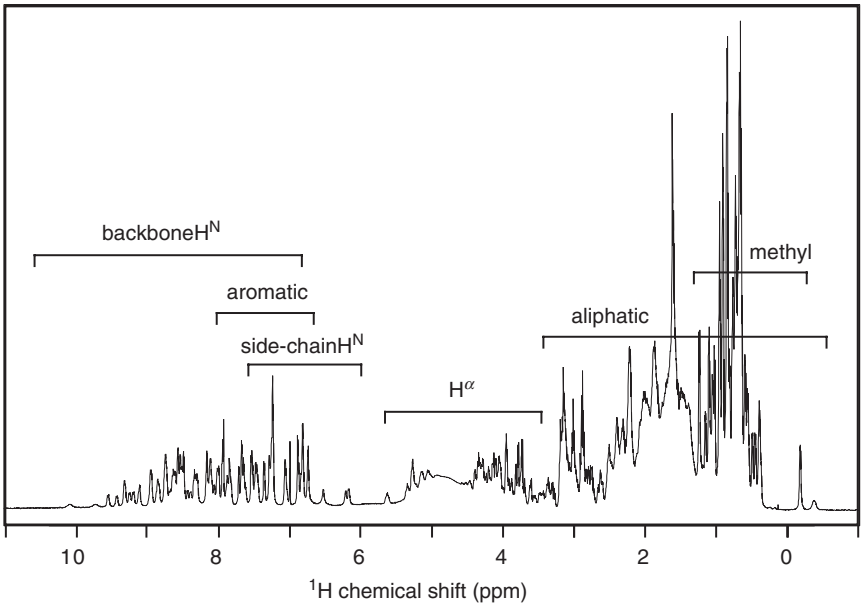



FIGURE 10.1 Chemical shift ranges observed for the various types of $^1H$ resonances in ubiquitin. The spectrum was recorded in $H_2O$ solution using presaturation and a Hahn echo (Section 3.6.4.2) following the acquisition pulse. The residual water resonance was removed from the FID by convolution (Section 3.5.4) prior to Fourier transformation.

The dependence of $^1H$ chemical shifts on protein secondary structure is discussed elsewhere (*4–11*). Figure 10.2 plots the distribution of $^1H$ chemical shifts (averaged over all observed conformations) for proteins in a database compiled by BioMagResBank (*12*).

The backbone amide $^1H^N$ signals are usually the best resolved set of resonances within a protein $^1H$ NMR spectrum; thus, the $^1H^\alpha$ and side chain resonance positions are most readily determined in scalar correlation experiments by the observation of direct and relayed cross-peaks to the backbone amide spins (*13, 14*). In theory, most $^1H$ 2D spectra are symmetric about the $F_2 = F_1$ diagonal (multiple-quantum spectra being the principal exceptions); hence, a particular cross-peak involving an amide $^1H^N$ spin is observable above ($F_2 > F_1$) or below ($F_2 < F_1$) the diagonal. In real spectra, this symmetry often is broken due to experimental artifacts, the most obvious of which is usually a streak of noise at the $F_2$ frequency of the water resonance that

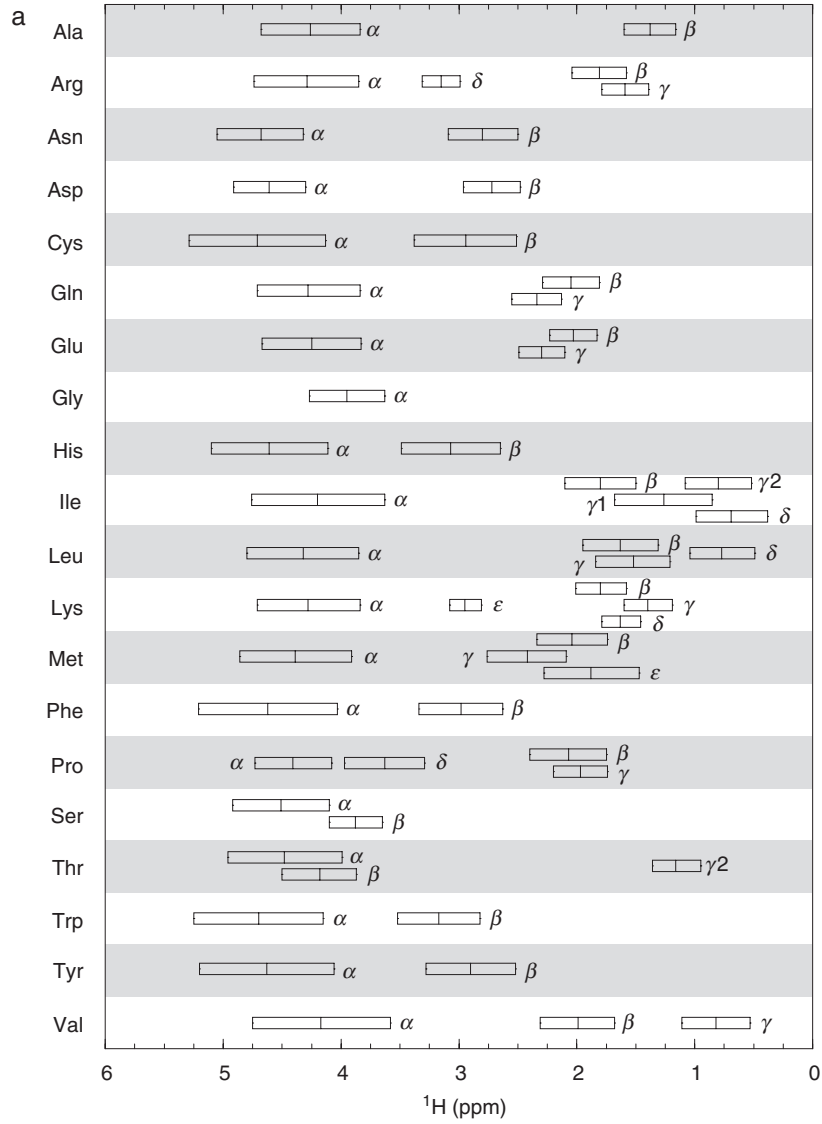



FIGURE 10.2 (a) Aliphatic and (b) nonaliphatic side chain $^1$H chemical shifts. The mean and standard deviations for each $^1$H nucleus type were obtained from the database of chemical shifts at BioMagResBank (www.bmrb.wisc.edu). The bars extend one standard deviation in either direction from the mean value. The chemical shifts of backbone amide $^1H^N$ spins show limited dependence on amino acid type [sequence-dependent random coil chemical shifts are given in Schwarzinger et al. (*1*) and Braun et al. (*143*)] and have a mean of 8.3 ppm and a standard deviation of 0.7 ppm.

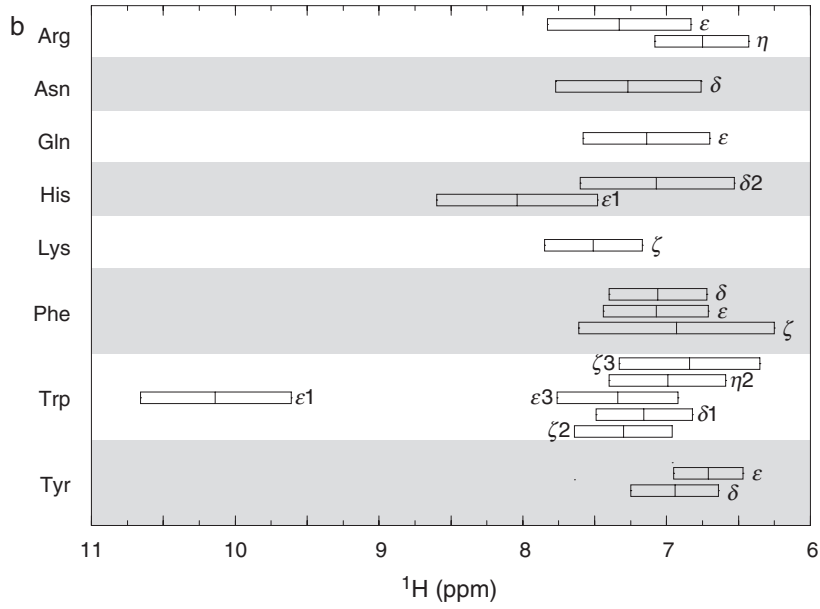



FIGURE 10.2—*Continued*

arises from incomplete solvent suppression. In unfavorable circumstances, ridges emanating from the central stripe extend 1.0 ppm or more from the water resonance and obscure many cross-peaks in this region. The distortions at the $F_1$ frequency of the water resonance are rarely comparable in width or intensity to the $F_2$ distortions. This asymmetry means that cross-peaks involving amide protons are most readily observed in the "above diagonal" region. Concentrating on these cross-peaks also has the advantage of combining the high digital resolution of the $F_2$ dimension with the high chemical shift dispersion of the amide $^1H^N$ resonances. For these reasons, the discussion of R.COSY, DR.COSY, and TOCSY experiments performed in $H_2O$ in Chapter 6 focused on the side chain to amide proton correlation cross-peaks occurring above the diagonal.

The principal process of determining $^1H$ resonance assignments is known as the *sequential assignment strategy*, and has been developed by Wüthrich and co-workers (*15*). This strategy is summarized in Fig. 10.3. The first stage of analysis makes use of $^1H$–$^1H$ scalar couplings to establish sets of $^1H^N$, $^1H^\alpha$, and aliphatic side chain resonances that belong to the same amino acid residue spin system. A protein of $N$ residues has $N$ distinct backbone-based spin systems. The $N$ spin

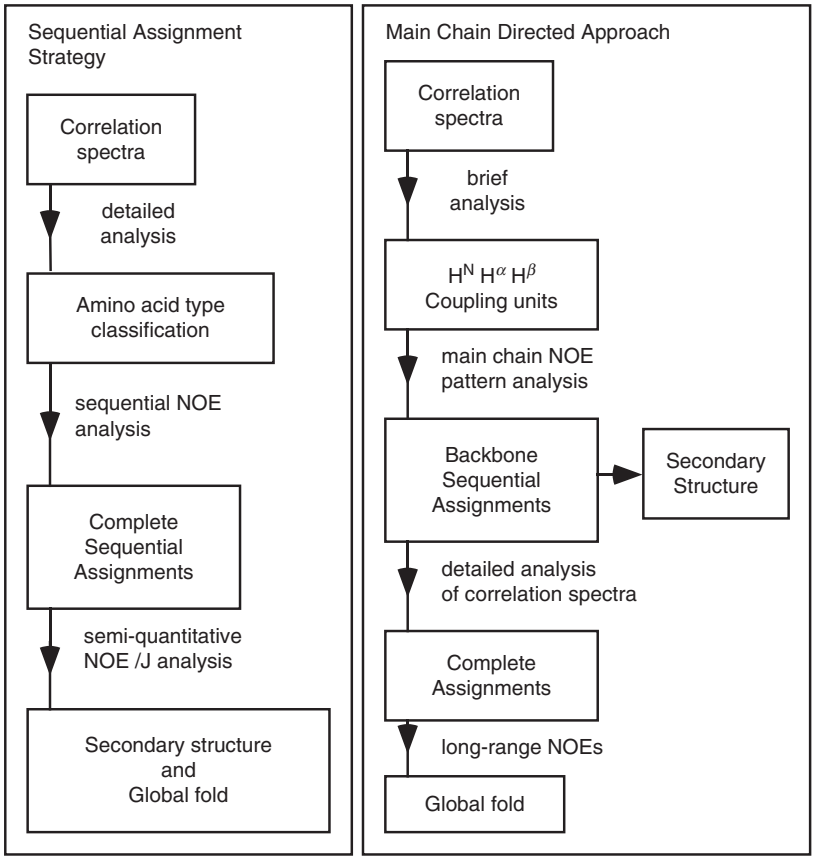



FIGURE 10.3 Comparison of the steps involved in the sequential assignment method (left) and the main chain directed method (right) of $^1$H resonance assignment for unlabeled proteins. Both schemes require knowledge of the primary sequence and make use of the same experimental data, although this information is utilized at different stages of the assignment process.

systems are assigned an amino acid type (or one of several possible types) based on the coupling topology and resonance chemical shifts, as described in Fig. 10.4.

Aromatic $^1$H spins of tyrosine, phenylalanine, tryptophan, and histidine residues; side chain amide protons of glutamine and asparagine residues; the side chain guanidinium group of arginine residues; and the methyl group of methionine residues are not scalar coupled to the remainder of the side chain and consequently comprise distinct spin

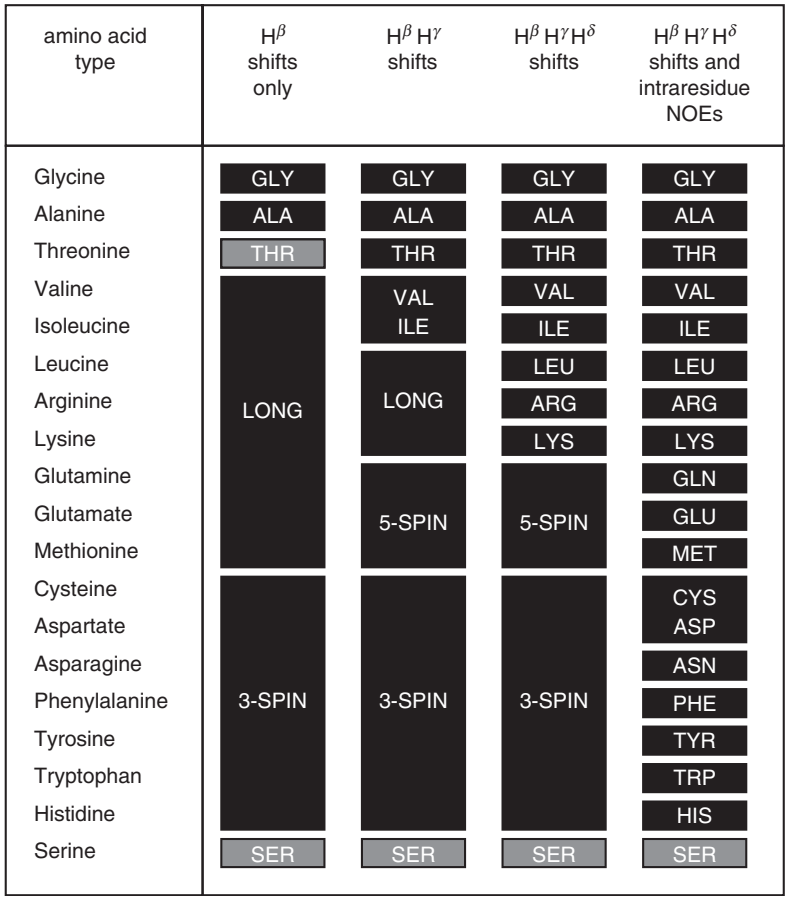
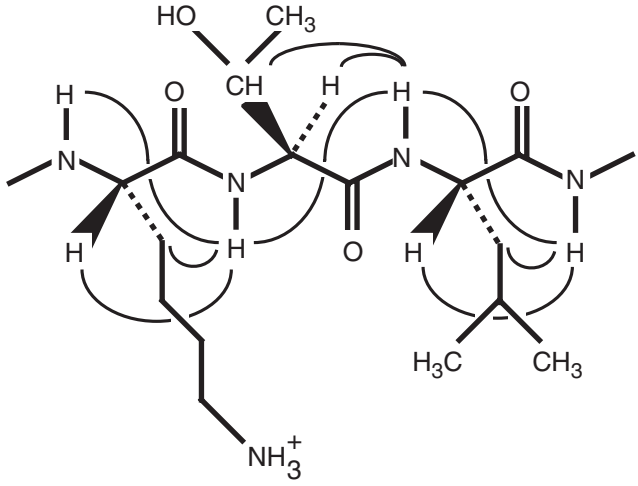



FIGURE 10.4 Categorization of spin systems based on knowledge of chemical shifts and spin topology. Each box indicates a unique category containing the amino acid types shown on the left. The four columns of categories indicate how spin systems can be more finely subdivided as chemical shifts further along the side chain are included in the analysis. The shading indicates that assignment of threonine and serine spin systems may be ambiguous due to the partial overlap of the $^{1}H^{\beta}$ chemical shifts of these residues with $^{1}H^{\beta}$ of three-spin systems. The first three category columns indicate categorizations possible using correlation spectra only; intraresidue NOEs are needed to perform the most complete categorization shown in the fourth column. Categorization in this fashion assumes that correlations to methyl groups can be differentiated from other interactions on the basis of chemical shift, intensity, or cross-peak fine structure.

systems. Association of the side chain and backbone resonances of these amino acid residues has to be made on the basis of intraresidue NOE correlations. For example, the most intense NOE correlations to $^1H^{\delta}$ spins of tyrosine or phenylalanine residues are generally from intraresidue $^1H^{\beta}$ spins, because these nuclei are always less than 2.8 Å apart (*15*). As a second example, the side chain amide $^1H$ nuclei trans to the side chain carbonyl oxygen ($^1H^{N\delta 2}$ in asparagine and $^1H^{N\varepsilon 2}$ in glutamine) are always closer to the other side chain protons than are the cis $^1H$ nuclei ($^1H^{N\delta 1}$ in asparagine and $^1H^{N\varepsilon 1}$ in glutamine). Therefore, the $^1H$ resonance with the more intense intraresidue NOEs is assigned to the trans position (*16*).

In the second stage of the assignment process, every spin system is assigned to a particular residue within the polypeptide chain by using through-space dipolar coupling (NOE) interactions to sequentially connect the spin systems identified from scalar correlations. Statistical analysis of hydrogen atom locations inferred from *x*-ray crystal structures of proteins has shown that a majority of the short distances between $^1H^N$, $^1H^{\alpha}$, and $^1H^{\beta}$ nuclei are between residues adjacent in the primary sequence (*17*). Thus, identification of intense NOEs from $^1H^N$, $^1H^{\alpha}$, and/or $^1H^{\beta}$ of one spin system to $^1H^N$ of a second spin system suggests that the two spin systems are adjacent in the primary sequence, with the first spin system nearer to the N-terminus of the protein (i.e., the spin systems correspond to residue $i$ and residue $i+1$, respectively). The sequential NOEs commonly are given the short-hand notation $d_{NN}$, $d_{\alpha N}$, and $d_{\beta N}$, respectively (*15, 18*). Identification of a series of sequential NOE interactions places several spin systems in the order $i, i+1, i+2, \ldots, i+n$, as illustrated for ubiquitin in Fig. 10.5. As more spin systems are connected, the sequence of spin systems eventually will match a unique section of the primary amino acid sequence of the protein (which must be known *a priori*); at this point, the spin systems are assigned *sequence specifically*. If the spin system types are well characterized (i.e., the majority of side chain resonance positions have been identified), then alignment of four or five spin systems usually is sufficient to achieve sequence-specific assignment.

The observation of $^1H^N$–$^1H^N$, $^1H^{\alpha}$–$^1H^N$, or $^1H^{\beta}$–$^1H^N$ NOEs is not limited to sequential interactions, and may also occur between non-sequential residues as a result of secondary or tertiary structure in the protein (*17*). The resulting ambiguity in the assignment process is reduced by the identification of sequential $^1H^N$–$^1H^N$, $^1H^{\alpha}$–$^1H^N$, *and* $^1H^{\beta}$–$^1H^N$ NOEs. Additionally, the sequential ordering of spin systems must match the primary sequence; $^1H^N$–$^1H^N$, $^1H^{\alpha}$–$^1H^N$, or $^1H^{\beta}$–$^1H^N$ NOEs between spin systems that are never adjacent in the




| | | |
|---|---|---|
| Lys6 | Thr7 | Leu8 |
| Ile13 | Thr14 | Leu15 |
| Arg54 | Thr55 | Leu56 |
| Ser65 | Thr66 | Leu67 |

FIGURE 10.5 The use of backbone–backbone NOEs to obtain sequence-specific $^1$H resonance assignments. In the top portion of the figure, the $d_{NN}$, $d_{\alpha N}$, and $d_{\beta N}$ connecting the three spin systems together are shown by the curved lines. Observation of a threonine prior to a leucine residue occurs four times in ubiquitin, hence such pairs of spin systems cannot be assigned sequence specifically. Once a lysine residue preceding the threonine is identified, the tripeptide segment can only arise from residues 6, 7, and 8. Due to the similarity of $^1H^{\beta}$ and $^1H^{\gamma}$ shifts of lysine, isoleucine, and arginine residues, $^1H^{\delta}$ chemical shifts must be identified in the first residue of this tripeptide to distinguish between Lys6, Ile13, and Arg54, as described in Fig. 10.4.

primary sequence must result from longer range contacts. In the limit, the assignments encompass *all* spin systems, and self-consistency is the best measure of the validity of the results. Tabulations of the relative intensities of the sequential NOEs customarily are presented graphically to demonstrate that complete, self-consistent sequential assignments have been achieved, as shown in Fig. 10.6.

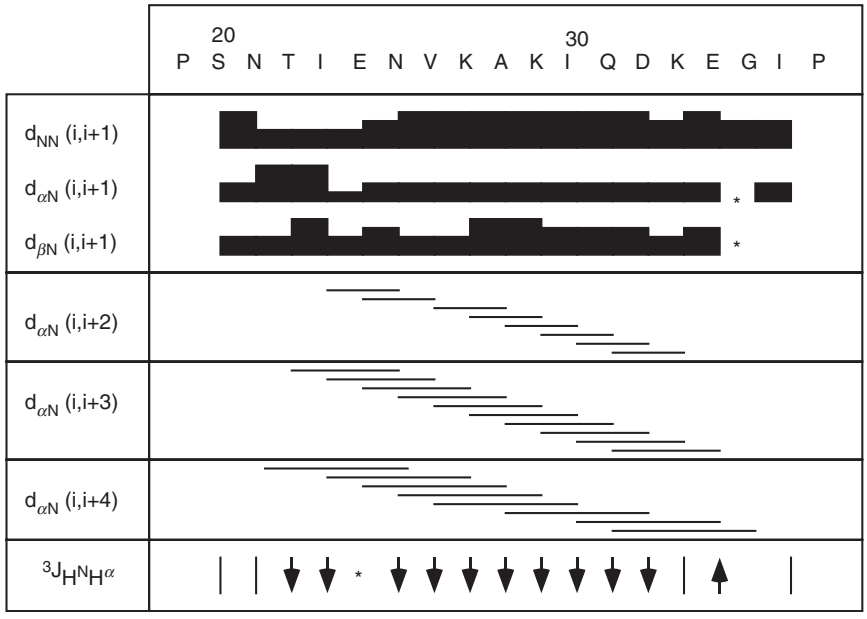



FIGURE 10.6 Summary of sequential NOEs observed for a portion of ubiquitin. The sequence is displayed along the top of the box, and the bars connect residues between which sequential NOEs are observed. The thickness of the bars indicates the intensity of the NOE. Commonly, medium-range NOEs and coupling constant data are included in such tabulations to help identify regions of regular secondary structure. Downward-pointing arrows, vertical lines, and upward-pointing arrows indicate $^3J_{H^N H^\alpha} < 6.0$ Hz, 6.0 Hz $< {}^3J_{H^N H^\alpha} < 8.0$ Hz, and $^3J_{H^N H^\alpha} > 8.0$ Hz, respectively. Asterisks indicate data that were not observed due to resonance overlap. This particular section of ubiquitin contains a helix from residues Thr22 to Glu34, as indicated by the intense $d_{NN}$ NOEs, medium-range NOEs, and small coupling constants.

An alternative strategy, known as the *main chain directed* (MCD) approach, has been developed by Englander and Wand (*19, 20*). In the MCD approach, scalar coupling connectivities are used initially to identify $^1H^N$–$^1H^\alpha$–$^1H^\beta$ units only. Assignment of the spin systems by amino acid type is not attempted. Next, the $^1H^N$–$^1H^\alpha$–$^1H^\beta$ units are aligned sequentially by systematically searching the NOESY spectrum for patterns of sequential NOEs. Different elements of secondary structure give rise to specific patterns of NOEs (*15*), and a search is made for these motifs in the following order: helix, antiparallel sheet, parallel sheet, turns, and loops. Once all of the backbone coupling

units have been aligned sequentially and categorized by secondary structural element, determination of the amino acid type of several side chains permits the defined elements of secondary structure to be aligned with the primary sequence. The sequential and MCD methods are compared schematically in Fig. 10.3.

The advantage of the MCD approach is that the sometimes-troublesome identification of amino acid type from scalar coupling data is not attempted initially, so that at least sequential backbone assignments can be made. Eventually, however, complete assignments of the side chains are needed because NOEs to all resonances must be assigned to determine precise three-dimensional structures. Frequently, elements of both the sequential and the MCD assignment approaches are combined. The initial identification of spin systems in the sequential approach is rarely complete due to problems of resonance degeneracy, spectral artifacts, and absent cross-peaks (due to small coupling constants). However, even limited knowledge of amino acid types can assist in the MCD sequential assignment process: once several residues have been connected sequentially, identification of the next residue is trivial if only one of the candidates is the correct type of amino acid. Further, sequential assignment of the backbone resonances facilitates assignment of the side chain resonances because the expected spin system topologies (amino acid types) are known.

Computer automated analysis of $^1$H 2D NMR spectra of proteins to obtain resonance assignments has been discussed for some years. Both fully automated (*21–24*) and semiautomated (*25–28*) approaches have been described. Resonance degeneracy and incomplete correlations in the $^1$H spectra are the main obstacles to automation. Homonuclear three-dimensional spectra potentially can circumvent resonance overlap (*29–32*). However, as discussed in Section 10.1.2, most recent efforts to automate the resonance assignment process have made use of the increased resolution afforded by heteronuclear NMR spectroscopy of isotopically labeled proteins. [For reviews of automation in NMR spectroscopy, see Baran et al. (*33*), Altieri and Byrd (*34*), Güntert (*35*), and Gronwald and Kalbitzer (*36*).]

The largest proteins for which nearly complete $^1$H sequential assignments can be obtained by homonuclear methods alone are usually in the range of 10–12 kDa [for examples, see Chazin et al. (*37*) and Williamson et al. (*38*)]. The upper limit depends somewhat on topology because proteins with substantial $\beta$-sheet content usually display better dispersion of $^1$H resonances than do predominantly $\alpha$-helical proteins. Homonuclear $^1$H assignments of ubiquitin have been obtained by both the sequential assignment and the MCD strategies (*20, 39*).

### 10.1.2 HETERONUCLEAR RESONANCE ASSIGNMENTS FOR ISOTOPICALLY LABELED PROTEINS

The conventional sequential or main chain directed assignment strategies also are used to obtain sequence-specific assignments from 3D heteronuclear NMR spectra. The principal advantage of using 3D heteronuclear-edited NOESY and TOCSY (Section 7.2), rather than homonuclear 2D experiments, for resonance assignments is the significant reduction in cross-peak overlap. The 3D $^{1}H$–$^{15}N$ NOESY–HSQC experiment is used to identify sequential through-space $d_{\alpha N}$, $d_{\beta N}$, and $d_{NN}$ connectivities. Amino acid spin systems are identified by observation of direct and relayed through-bond connectivities between the $^{1}H^{N}$ spins and the $^{1}H^{\alpha}$ and aliphatic side chain protons using the 3D $^{1}H$–$^{15}N$ TOCSY–HSQC experiment (Section 7.2.2). Alternatively, complete side chain assignments also can be obtained from HCCH–COSY and HCCH–TOCSY experiments (Section 7.3). In this case, the side chain spin systems are connected with the backbone $^{1}H^{N}$ and $^{15}N$ resonances via $^{1}H^{N}$–$^{15}N$–$^{1}H^{\alpha}$ and other correlations observed in an $^{1}H$–$^{15}N$ TOCSY–HSQC spectrum, or by using correlations observed in one or two of several triple-resonance experiments (Section 7.4) [e.g., HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HBHA(CBCA)NH, or HBHA(CBCACO)NH]. The heteronuclear-edited NOE-based sequential assignment method has been successfully applied to a number of proteins, with molecular masses up to ~20 kDa (*40–46*).

The triple-resonance experiments introduced in Section 7.4 offer an alternative to the NOE-based strategy for sequentially assigning $^{1}H^{N}$, $^{15}N$, $^{13}CO$, $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{1}H^{\beta}$, and $^{13}C^{\beta}$ resonances. Using these experiments, sequential correlations are established via the relatively uniform and well-resolved heteronuclear one-bond and two-bond couplings, without any prior knowledge of spin system types. Side chain assignments are completed using the TOCSY–HSQC, HCCH–COSY, and HCCH–TOCSY experiments. Potential errors that arise from misassignment of sequential and long-range connectivities in the NOE-based procedures are avoided because assignments are based solely on predictable through-bond scalar correlations. Triple-resonance approaches were developed by Bax and co-workers to overcome difficulties in obtaining resonance assignments for calmodulin by homonuclear techniques (*47*). Calmodulin (16.7 kDa) is largely $\alpha$-helical and has very narrow chemical shift distributions for both $^{1}H^{N}$ and $^{1}H^{\alpha}$ spins.

For example, the combined use of the HNCA (Section 7.4.1) experiment, which provides intraresidue (together with some

sequential) correlations between $^{1}H^{N}$, $^{15}N$, and $^{13}C^{\alpha}$ resonances, and the HN(CO)CA experiment (Section 7.4.2), which gives solely inter-residue correlations between the $^{1}H^{N}$ and $^{15}N$ resonances of one residue and the $^{13}C^{\alpha}$ resonance of the preceding residue, provides an obvious route to sequential assignment. Each $^{13}C^{\alpha}$ resonance is linked to both its intraresidue and sequential $^{1}H^{N}$ and $^{15}N$ resonances. Ambiguities caused by chemical shift degeneracy are solved by using additional experiments that provide alternative correlations. For instance, the HNCO (Section 7.4.4.1) and HN(CA)CO (Section 7.4.4.2) experiments correlate the $^{1}H^{N}$ and $^{15}N$ resonances with both the intraresidue and the sequential $^{13}CO$ signals (rather than the $^{13}C^{\alpha}$ spins). To "align" the backbone sequential assignments with the protein amino acid sequence, side chain amino acid spin systems are identified from HCCH–COSY, HCCH–TOCSY, and $^{1}H$–$^{15}N$ TOCSY–HSQC experiments.

A different set of experiments, CBCA(CO)NH [or HN(CO)CACB] and HNCACB (or CBCANH), together with the closely related HBHA (CBCACO)NH and HBHA(CBCA)NH experiments (Section 7.4.5), offers an alternative sequential assignment strategy for proteins with $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ chemical shift degeneracy. Using these experiments, the $^{1}H^{N}$ and $^{15}N$ resonances are correlated with the intraresidue and sequential $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ (or $^{1}H^{\alpha}$ and $^{1}H^{\beta}$) resonances. Information regarding amino acid type can be obtained from the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts (see later). The HCC(CO)NH–TOCSY (*48–52*) and CC(CO)NH–TOCSY (*53*) experiments, discussed in Sections 9.1.6 and 9.1.7, also provide sequential and spin system assignment information. These experiments are similar in principle to the CBCA(CO)NH and HNCACB experiments, except that $^{13}C$ isotropic mixing is used instead of COSY-type transfer to relay aliphatic side chain magnetization to the $^{13}C^{\alpha}$ nucleus. Sequential and intraresidue correlations are obtained between backbone $^{1}H^{N}$ and $^{15}N$ resonances and the side chain (either $^{13}C$ or $^{1}H$) resonances. Finally, banks of experiments have been designed to achieve resonance assignments for particular side chain topologies, such as aromatic amino acids (*54, 55*), side chain $^{1}H$–$^{15}N$ groups (*56*), arginine amino acids (*57*), and proline amino acids (*58, 59*).

Increasingly, resonance assignments of proteins with molecular masses of $>20$ kDa are obtained using $^{2}H/^{13}C/^{15}N$ triply labeled proteins to reduce losses due to efficient $^{1}H$–$^{13}C$ dipole–dipole relaxation, as discussed in Section 9.1. Triple-resonance experiments for backbone resonance assignments discussed in Section 9.1.5 are based on out-and-back transfers from the $^{1}H^{N}$ spin. Assignments are obtained

primarily from three pairs of three-dimensional experiments: HNCA and HN(CO)CA, HNCACB and HN(CO)CACB, and HNCO and HN(CA)CO experiments. These experiments also have four-dimensional variants, such as HNCOCA (*60*). Assignments of aliphatic and aromatic side chains using combined analysis of $^{13}C/^{15}N$ doubly labeled and $^{2}H/^{13}C/^{15}N$ triply labeled proteins, random fractionally deuterated $^{2}H/^{13}C/^{15}N$ proteins, or selectively protonated proteins have been described in Sections 9.1.6, 9.1.7, and 9.1.9. Nearly complete backbone and methyl resonance assignments of malate synthase G, with 723 amino acid residues and a molecular mass of 81.4 kDa, have been achieved by Kay and co-workers (*61, 62*).

Information on amino acid type is obtained from $^{13}C$ chemical shift data. Random coil $^{13}C$ chemical shifts have been determined for unstructured peptides (*1, 2, 63*). The dependence of $^{13}C$ chemical shifts on protein secondary structure is discussed elsewhere (*11, 64–66*). The distributions of aliphatic and aromatic $^{13}C$ chemical shifts for different amino acid residues compiled in a database at BioMagResBank (*12*) are plotted in Fig. 10.7. The characteristic $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts of alanine, threonine, serine, and glycine residues (Fig. 10.7) allow ready identification of these amino acid types. Clearly, knowledge of other aliphatic $^{13}C$ chemical shifts can also be used to assign a given spin system to a unique or limited number of possible amino acid types. This information, coupled with the alignment of sequentially connected spin systems with the known amino acid sequence, leads to unambiguous assignment.

The relative simplicity and predictability of triple-resonance 3D and 4D spectra used for protein assignment purposes makes these experiments particularly amenable to automated or semiautomated analysis. Current efforts at automated assignment generally begin with automatic peak picking. The reduced resonance overlap in 3D and 4D triple-resonance spectra, relative to 2D homonuclear spectra, increases the reliability of this process; however, the final peak lists usually must be edited (by the spectroscopist) to distinguish "real" resonance peaks from spectral artifacts. The peak lists are searched automatically to find expected intraresidue and interresidue correlations, and the spin systems are identified and sequentially ordered according to these correlations. Information regarding spin system type, which may be obtained from $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts values, for instance, is incorporated and the ordered spin systems are aligned with the known amino acid sequence. Current state-of-the-art automatic assignment methods have been reviewed (*33–36, 67*).


FIGURE 10.7 (a) Aliphatic and (b) aromatic $^{13}C$ chemical shifts. The mean and standard deviations in the mean for each nucleus type were obtained from the database of chemical shifts at BioMagResBank (www.bmrb.wisc.edu). The bars extend one standard deviation in either direction from the mean value. Two ranges are included for $^{13}C^{\beta}$ of cysteine, due to the significant effect that disulfide bond formation has on these chemical shifts. $^{13}CO$ chemical shifts have a mean and standard deviation of 176.1 ppm and 2.3 ppm, respectively.

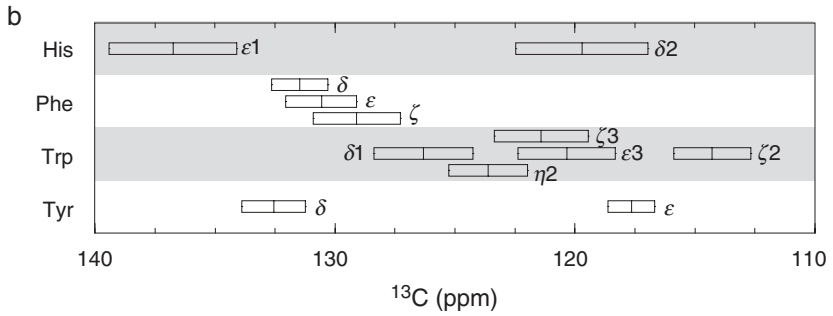



FIGURE 10.7—*Continued*

## 10.2 Three-Dimensional Solution Structures

Sequence-specific assignment of NMR resonances typically requires many different experiments and a substantial effort from the spectroscopist. However, mere achievement of resonance assignments usually does not reveal much of interest about the structure, dynamics, and interactions of the protein under investigation. Assignments are a necessary prerequisite to achieving more biochemical goals. This section discusses procedures for determining the three-dimensional structures of proteins with the aim of highlighting the information that can be obtained from the NMR experiments described in Chapters 6–9. Structure determination (*68*), characterization of protein dynamics (*69*), and assessment of protein–ligand interactions (*70*) by NMR spectroscopy have been reviewed.

### 10.2.1 NMR-Derived Structural Restraints

Essentially all parameters that can be measured by NMR spectroscopy are sensitive in some, more-or-less complex, manner to molecular conformation; therefore, quantification of these parameters permits structural analysis by NMR spectroscopy. At present, dipolar cross-relaxation (NOE) rate constants, scalar coupling constants, isotropic chemical shifts, and residual dipole–dipole coupling constants (RDCs) are the most commonly utilized parameters for protein structure determination. Structural restraints derived from NOE interactions and scalar coupling constants have formed the basis of protein structure determination by NMR spectroscopy, beginning with the report of the structure of proteinase inhibitor IIA from bull seminal plasma in 1985 (*71*).

In recent years, restraints derived from isotropic chemical shifts (*66*) and RDCs (*72*) have become extensively utilized. Other NMR parameters, including amide proton–solvent exchange protection factors (*73*), *trans*-hydrogen bond scalar coupling constants (*74*), and paramagnetic effects (*75–77*), also provide structural restraints in systems for which these data are available.

*10.2.1.1 NOE Distance Restraints* The most important NMR-observable parameter used in determining protein structure is the NOE. The dipolar cross-relaxation rate constant is proportional to the inverse sixth power of the distance between two interacting $^1$H spins (Section 5.5). In the initial rate approximation, NOE cross-peak intensities are proportional to the cross-relaxation rate constants. Thus, if one interproton distance, $r_{\text{ref}}$, is known (e.g., from covalent geometry), then another, unknown interproton distance, $r_i$, is determined by the relationship (ignoring internal mobility)

$$r_i = r_{\text{ref}}(S_{\text{ref}}/S_i)^{1/6}, \qquad [10.1]$$

in which $S_{\text{ref}}$ and $S_i$ are the cross-peak intensities. In practice, NOESY mixing times short enough to satisfy the initial rate approximation usually are impractical because the cross-peak intensities then have low signal-to-noise ratios. For longer mixing times, the intensities of NOESY cross-peaks are no longer directly proportional to the cross-relaxation rate constants between the interacting spins both because the time evolution of magnetization transfer even between two isolated spins is exponential and because magnetization is transferred among spins in multiple steps via spin diffusion (Section 6.6.1.2).

A majority of the restraints used to calculate the three-dimensional structure of a protein come from many hundreds of NOE cross-peaks, and are represented as bounds on the separation of pairs of $^1$H atoms. Without recourse to complex calculations (*78*), precise $^1$H–$^1$H separations cannot be determined from the NOE cross-peak intensities; instead, NOE cross-peaks typically are grouped on the basis of their intensities into several categories (e.g., strong, medium, and weak). Each category is associated with an upper bound separation between the interacting spins (e.g., 2.7, 3.3, and 5.0 Å). The cross-peak volume limits and the upper bound distances for the categories are estimated from the NOE intensities observed for protons of known covalent geometry (e.g., geminal methylene protons, or vicinal protons in aromatic rings) or between protons in regions of regular secondary structure (e.g., sequential $^1H^{\alpha}$–$^1H^{N}$, cross-strand $^1H^{\alpha}$–$^1H^{\alpha}$ and

cross-strand $^1H^{\alpha}$–$^1H^N$ NOEs in $\beta$-sheets, or $^1H^{\alpha}$–$^1H^N$ medium-range NOEs in helices). Conservative upper bound distances compensate for cross-peaks affected by spin diffusion or partial overlap; otherwise, failure to allow for such artifacts results in structures that have large violations of the input restraints, high energies, and artificially high precision (*79*). Usually, the lower bound separations for pairs of protons are set to the sum of the van der Waals radii.

The precision to which a structure can be determined is directly related to the number of experimental restraints used in the structure calculation. Structures of low resolution can be obtained with as few as five restraints per residue, while the most precise structures obtained from $^1H$ data alone have up to 15 restraints per residue. For the latter cases, the root-mean-square deviation (RMSD) of backbone atoms from the mean structure may be as low as 0.5 Å for the well-defined regions of secondary structure. The use of $^{15}N$ or $^{13}C$ labeling usually allows many more $^1H$–$^1H$ NOEs to be unambiguously identified, increasing the number of restraints per residue to between 20 and 25, and leading to the highest precision structures (backbone RMSD in the range of 0.3 to 0.5 Å). Stereospecific assignments of prochiral groups [especially $C^{\beta}H_2$ groups, and $(CH_3)_2CH$ groups of valine and leucine residues] are critical to obtaining structures of high precision (*80*). Coverage, defined as the fraction of theoretically observable NOEs that are actually utilized in the structure calculation, has been emphasized as an important determinant of structure quality (*81*).

Structure determination requires that the accuracy of the experimental NOE intensities must be high. NOESY spectra must have very flat base lines (even slight rolls or tilts in the baseline can have dramatic effects on cross-peak volumes) and be otherwise free of spectral artifacts; such topics are discussed in Chapters 6 and 7.

*10.2.1.2 Dihedral Angle Restraints from Scalar Coupling Constants* As was first described by Karplus (*82*), the magnitude of a $^3J$ scalar coupling constant is a function of the dihedral angle formed by the three covalent bonds:

$$^3J = A\cos^2\theta + B\cos\theta + C. \qquad [10.2]$$

The constants $A$, $B$, and $C$ depend upon the particular nuclei involved, and $\theta$ is the dihedral angle. Historically, dihedral angle restraints for $\phi$ and $\chi_1$ dihedral angles have been derived only from $^3J_{H^NH^{\alpha}}$ and $^3J_{H^{\alpha}H^{\beta}}$ coupling constants, respectively (*83–85*). Recently,

numerous experiments have been developed that allow measurement of $^{13}C$–$^{13}C$, $^{13}C$–$^{15}N$, $^{1}H$–$^{15}N$, and $^{1}H$–$^{13}C$ three-bond coupling constants in isotopically enriched proteins [Section 7.5; reviewed in Vuister et al. (*86*)].

Accurate parameterization of the Karplus equation is necessary in order to relate measured values of $^{3}J$ coupling constants to dihedral angle ranges. Numerous attempts have been made to refine the original Karplus approach to account for the chemical nature of the atoms involved in the coupling (*87*). However, such theoretical approaches lead to complex relationships that are not generally applicable. Instead, semiempirical methods of obtaining values for the constants $A$, $B$, and $C$ in [10.2] have been more successful. The most relevant values for the constants have been derived by correlation of observed $^{3}J$ values with the corresponding dihedral angles measured in protein structures determined by $x$-ray crystallography or NMR methods (*84, 88–92*). Recently, the availability of large data sets of coupling constants have allowed application of sophisticated mathematical and statistical methods to the optimization of the parameters of the Karplus equation (*93–95*).

Calibration of the Karplus equation and determination of dihedral angle restraints are complicated by relaxation-induced self-decoupling that reduces the apparent coupling constant in larger proteins (Section 5.4.2) (*96*) and by intramolecular motions that average the values of $^{3}J$ over distributions of dihedral angles (*97*). Motional averaging means that useful restraints sometimes cannot be derived from accurately measured coupling constants (*98*). For example, extensive backbone motion results in a $^{3}J_{H^{N}H^{\alpha}} \approx 7.0$ Hz in short, unstructured linear peptides. Thus, $^{3}J_{H^{N}H^{\alpha}}$ coupling constants observed in the range 6.0–8.0 Hz generally are not utilized as protein structural restraints, because they may reflect motional averaging of multiple conformations. The closer $^{3}J$ is to one of the extrema expected from the Karplus relationship, the less the likelihood of significant motional averaging and the more accurate the resulting restraint.

Examples of the Karplus parameters that have been reported in the literature for $^{3}J_{H^{N}H^{\alpha}}$ are given in Table 10.1 and plotted in Fig. 10.8a. Experimental data for ubiquitin are compared to the theoretical curves in Fig. 10.8b. Given the shape of the Karplus curve, shown in Fig. 10.8, as many as four different conformations can yield the same value of $^{3}J$, although some of the possible orientations may be sterically unfavorable. The degeneracy inherent in the Karplus relationship is alleviated if more than one scalar coupling constant sensitive to a given dihedral angle is measured. As shown in Table 10.2 and Fig. 10.8c, d,

TABLE 10.1
Parameterization of the Karplus Equation for $^3J_{H^NH^\alpha}$ [a]

| *A* | *B* | *C* | Ref. |
|---|---|---|---|
| 6.4 | −1.4 | 1.9 | (*84*) |
| 6.0 | −1.4 | 2.4 | (*88*) |
| 6.7 | −1.3 | 1.5 | (*89*) |
| 6.51 | −1.76 | 1.60 | (*90*) |
| 7.09 | −1.42 | 1.55 | (*137*) |
| 7.90 | −1.05 | 0.65 | (*93*) |
| 7.13 | −1.31 | 1.56 | (*94*) |

[a]The value of $\theta = \phi - 60^\circ$.

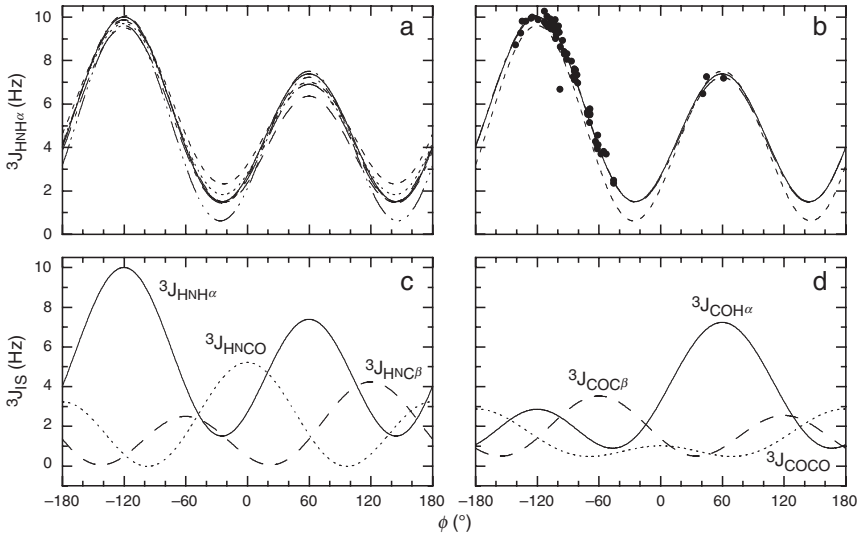



FIGURE 10.8 Karplus curves describing the variation of $^3J_{COCO}$, $^3J_{COH^\alpha}$, $^3J_{COC^\beta}$, $^3J_{H^NCO}$, $^3J_{H^NH^\alpha}$, and $^3J_{H^NC^\beta}$ with backbone dihedral angle $\phi$. Curves were calculated using [10.2]. (a) Curves for $^3J_{H^NH^\alpha}$ are calculated using parameter values given in Table 10.1 and are taken from various studies: (- - - -) (*84*), (– – –) (*88*), (– - –) (*89*), (— — —) (*90*), (– · –) (*137*), (– - - - –) (*93*), and (—) (*94*). (b) Curves for $^3J_{H^NH^\alpha}$ are calculated using parameter values given in Table 10.1 and are taken from three studies: (– - –) (*137*), (– – –) (*93*), and (—)(*94*). Solid points are experimental values taken from PDB file 1d3z (*129*). (c) and (d) Parameter values are given in Tables 10.1 and 10.2 and are taken from one study (*94*).

TABLE 10.2
Parameterization of the Karplus Equation for $\phi^a$

| Coupling constant | $A$ | $B$ | $C$ | $\Delta\phi$ | Ref. |
|---|---|---|---|---|---|
| $^3J_{COCO}$ | 1.36 | −0.93 | 0.60 | 0° | (*137*) |
| | $1.51 \pm 0.86$ | $-1.09 \pm 1.11$ | $0.52 \pm 0.39$ | | (*93*) |
| | $1.30 \pm 0.12$ | $-0.93 \pm 0.06$ | $0.64 \pm 0.03$ | | (*94*) |
| $^3J_{COH^\alpha}$ | 3.72 | −2.18 | 1.28 | +120° | (*137*) |
| | $3.76 \pm 1.05$ | $-1.63 \pm 0.56$ | $0.89 \pm 0.60$ | | (*93*) |
| | $3.84 \pm 0.14$ | $-2.19 \pm 0.10$ | $1.20 \pm 0.11$ | | (*94*) |
| $^3J_{COC^\beta}$ | 1.74 | −0.57 | 0.25 | −120° | (*137*) |
| | $2.72 \pm 0.80$ | $-0.31 \pm 0.52$ | $0.39 \pm 0.37$ | | (*93*) |
| | $2.52 \pm 0.33$ | $-0.49 \pm 0.33$ | $0.51 \pm 0.12$ | | (*94*) |
| $^3J_{H^NCO}$ | 4.29 | −1.01 | 0.00 | ±180° | (*137*) |
| | $4.41 \pm 0.81$ | $-1.36 \pm 1.03$ | $0.24 \pm 0.37$ | | (*93*) |
| | $4.19 \pm 0.30$ | $-0.99 \pm 0.18$ | $0.03 \pm 0.05$ | | (*94*) |
| $^3J_{H^NH^\alpha}$ | 7.09 | −1.42 | 1.55 | −60° | (*137*) |
| | $7.90 \pm 1.02$ | $-1.05 \pm 0.54$ | $0.65 \pm 0.58$ | | (*93*) |
| | $7.13 \pm 0.34$ | $-1.31 \pm 0.13$ | $1.56 \pm 0.34$ | | (*94*) |
| $^3J_{H^NC^\beta}$ | 3.06 | −0.74 | 0.13 | +60° | (*137*) |
| | $2.90 \pm 0.80$ | $-0.56 \pm 0.52$ | $0.18 \pm 0.37$ | | (*93*) |
| | $3.26 \pm 0.23$ | $-0.87 \pm 0.24$ | $0.10 \pm 0.08$ | | (*94*) |

[a]The value of $\theta = \phi + \Delta\phi$.

up to six $^3J$ scalar coupling constants depend upon the backbone $\phi$ dihedral angle.

In contrast to the abundance of scalar coupling constants that depend on $\phi$, only the relatively small $^3J_{H^\alpha N}$ (referring to the $^1H^\alpha_{i-1} - ^{15}N_i$ spin pair) depends on the backbone dihedral angle $\psi$. The Karplus parameters for this coupling constant are given in Table 10.3 and plotted in Fig. 10.9. Experiments for measuring relaxation interference have been proposed as an alternative approach for determining $\psi$ (*99*).

The side chain dihedral $\chi_1$ historically has been characterized using $^3J_{H^\alpha H^\beta}$ scalar coupling constants; as for $\phi$, the availability of isotopically enriched proteins has allowed measurement of many additional scalar coupling constants that depend on $\chi_1$. Consensus values of the parameters of the Karplus equation for scalar coupling constants sensitive to $\chi_1$ are given in Table 10.4 and the Karplus curves for $^3J_{H^\alpha H^{\beta 2}}$ and $^3J_{H^\alpha H^{\beta 3}}$ (the two $^1H^\beta$ spins for an amino acid with a $C^\beta$ methylene group)

TABLE 10.3
Parameterization of the Karplus Equation for $^3J_{H^\alpha N}$[a]

| *A* | *B* | *C* | Ref. |
|---|---|---|---|
| −0.88 | −0.61 | −0.27 | (*138*) |
| −1.0 | −0.65 | −0.15 | (*139*) |

[a]The value of $\theta = \psi - 120^\circ$.

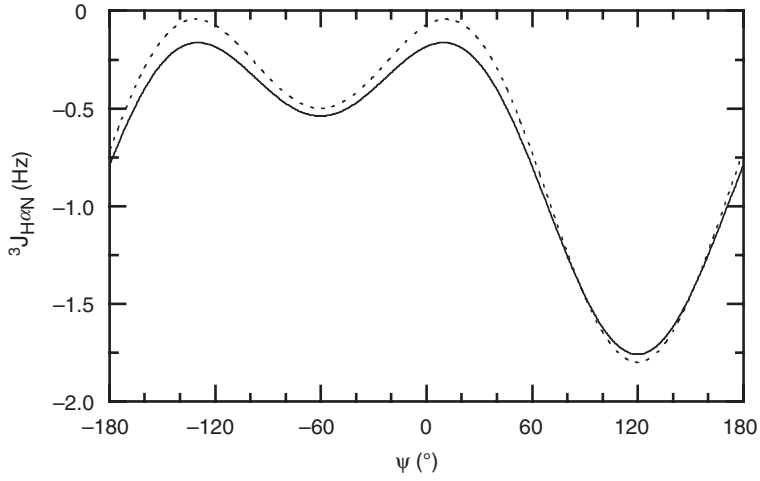



FIGURE 10.9 Karplus curve describing the variation of $^3J_{H^\alpha N}$ with backbone dihedral angle $\psi$. Curves were calculated using [10.2]. Parameter values are given in Table 10.3 and are taken from two studies: (—) (*138*) and (- - -) (*139*).

are plotted in Fig. 10.10. Dynamic averaging within and between rotameric states is more severe for side chain dihedral angles and the Karplus parameters for side chain dihedral angles are more dependent on the nature of the amino acid side chain than for the backbone dihedral angles (*95*).

As a result of the preceding considerations, dihedral angle restraints enforced during structural calculations must include all of the valid solutions to the Karplus curve and usually are defined conservatively (usually not restrained to ranges less than 60° or 120°) to allow for limited motional averaging of the observed coupling constant.

TABLE 10.4
Parameterization of the Karplus Equation for $\chi_1$[a]

| Coupling constant | $A$ | $B$ | $C$ | | Ref. |
|---|---|---|---|---|---|
| $^3J_{H^\alpha H^\beta}$ | 9.5 | −1.6 | 1.8 | −120°/0° | (*92*) |
| | 9.5 | −1.0 | 1.4 | | (*91*) |
| | 10.2 | −1.8 | 1.9 | | (*83*) |
| | 7.23 | −1.37 | 2.22 | | (*95*) |
| $^3J_{H^\alpha C^\gamma}$ | 7.1 | −1.0 | 0.7 | 120° | (*140*) |
| | 5.34 | −0.96 | 0.79 | | (*95*) |
| $^3J_{NH^\beta}$ | −4.4 | 1.2 | 0.1 | 120°/–120° | (*141*) |
| | −2.30 | 0.75 | −1.07 | | (*95*) |
| $^3J_{COH^\beta}$ | 7.2 | −2.0 | 0.6 | 0°/120° | (*142*) |
| | 4.02 | −1.58 | 1.32 | | (*95*) |
| $^3J_{NC^\gamma}$ | 1.29 | −0.49 | 0.37 | 0° | (*95*) |
| $^3J_{COC^\gamma}$ | 2.31 | −0.87 | 0.55 | 120° | (*95*) |

[a]The value of $\theta = \phi + \Delta\phi$. When two values of $\Delta\phi$ are given, the first corresponds to $^1H^{\beta 2}$ and the second to $^1H^{\beta 3}$.

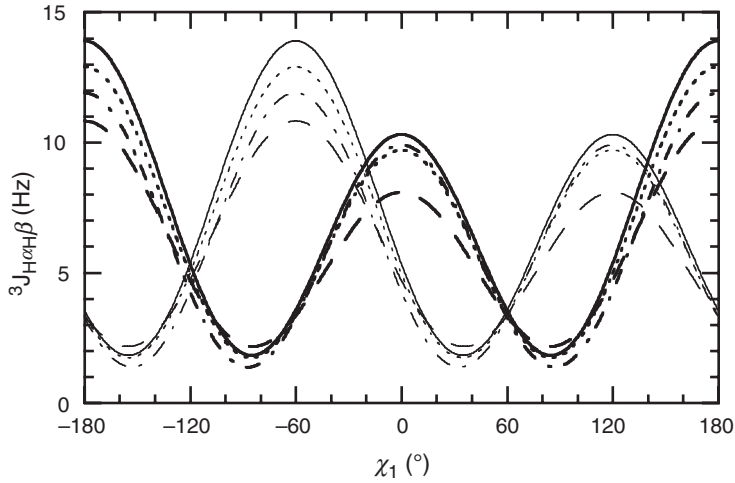



FIGURE 10.10 Karplus curves describing the variation of $^3J_{H^\alpha H^{\beta 2}}$ (thin lines) and $^3J_{H^\alpha H^{\beta 3}}$ (thick lines) with side chain dihedral angle $\chi_1$. Curves were calculated using [10.2]. Parameter values are given in Table 10.4 and are taken from various studies: (- - -) (*92*), (– – –) (*91*), (– - –) (*83*), and (—) (*95*).

*10.2.1.3 Dihedral Angle Restraints from Isotropic Chemical Shifts* Isotropic chemical shifts are exquisitely sensitive to local molecular conformation, but this extreme sensitivity also complicates the interpretation of chemical shifts in atomic detail. Fortunately, the dependence of chemical shifts of backbone nuclei, particularly $^{1}H^{\alpha}$, $^{13}CO$, $^{13}C^{\alpha}$, and $^{13}C^{\beta}$, on secondary structure is well-established (*11, 100*). Thus, the secondary chemical shift, defined as the observed value of the shift minus the value expected for the same residue in a random coil peptide [a compilation of random coil shifts is given by Schwarzinger and co-workers (*1*)], exhibits characteristic patterns for regular elements of secondary structure. This correlation forms the basis for the chemical shift index (CSI) method for identifying elements of secondary structure in proteins (*10, 64*). The TALOS program compares observed chemical shifts to a database of proteins with $^{1}H^{\alpha}$, $^{13}CO$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{15}N$ resonance assignments and high-resolution structures to obtain dihedral angle restraints for incorporation into structure calculations (*66*).

*10.2.1.4 Restraints from Residual Dipolar Coupling Constants* Recently, a new class of structural restraint has been introduced that is not strictly local in nature and which represents a major advance in NMR structural studies. These restraints are based upon the measurement of *residual dipolar couplings* (RDCs) between pairs of NMR active nuclei in partially aligned molecules. Theoretical origins of RDCs are discussed in Section 2.8 and experimental approaches for measuring RDCs are discussed in Section 7.6. For directly bonded pairs of nuclei, for which the bond length is known, RDCs depend only on the orientation of the bond vector in the molecular alignment frame, as shown by [2.326], and described by the polar angles $\{\theta, \phi\}$. Approaches for incorporating restraints from RDCs into structure calculations continue to be developed and the optimal approach is not established at present. Two problems arise in the interpretation of RDCs: (*i*) the axial and rhombic components of the alignment tensor, $A_a$ and $A_r$, and the Euler angles that define the orientation of the alignment frame in the laboratory frame must be determined and (*ii*) a single RDC can arise from more than one set of $\{\theta, \phi\}$ angles. Thus, a minimum of five measured RDC values are required to define the alignment tensor (alternatively, [2.318]–[2.321] show that five measurements are needed to define the five independent elements of the Saupe order matrix). In addition, a given RDC is consistent with eight combinations of the polar angles, corresponding to all combinations of $\{\theta, \pi - \theta\}$ and $\{\phi, -\phi, \pi - \phi, \pi + \phi\}$. Thus, a single measured RDC does not provide a unique restraint.

How these problems are solved has given rise to different methods for utilizing RDCs in structure calculations.

The most straightforward method for incorporating RDC data into structure calculation protocols is by direct refinement of the orientation of individual bond vectors against the measured values of the RDCs (*101*). In this case, the orientation of each individual bond vector is changed to satisfy dipolar couplings as structures are calculated. Direct refinement has been shown to improve the accuracy and precision of structures when used in conjunction with nearly complete sets of NOE, coupling constant, and chemical shift data (*101*). This approach can be used with limited sets of RDCs; for example, many applications use only the backbone N–H RDCs. However, direct refinement requires that relatively high-quality initial structures have been determined from NOE, scalar coupling, and other restraints, because many local minima are encountered when refining RDCs (*102, 103*). Bax and Grishaev discuss the difficulties that can result from refining against too limited a set of RDCs, such as the $^{1}H–^{15}N$ RDCs alone (*104*).

The second common approach treats the peptide plane as a fixed structural unit (*103*). For each plane the goal is to use at least five RDCs to define the Euler angles $(\alpha, \beta, \gamma)$ that transform from an initial frame to the peptide alignment frame. Each peptide alignment frame must coincide with the global alignment coordinate system, effectively rotating each peptide into its correct orientation in the overall structure. The five dipolar couplings measured with the HNCO-based experiments described in Section 7.6 are sufficient to solve for the transformation of each peptide plane into the alignment frame. A difficulty with this approach is that some method is needed to distinguish between the eight possible orientations of the peptide plane consistent with the RDCs. Trial structures calculated from NOE restraints can be used to serve this purpose.

*10.2.1.5 Hydrogen Bond Restraints from Amide Proton–Solvent Exchange* Slow rates of amide exchange (relative to the intrinsic rates for unstructured peptides; Chapter 3, Fig. 3.36) are associated with shielding of amide $^{1}H^{N}$ atoms from solvent, and most commonly result from hydrogen bonding interactions (*73*). Amide exchange rates are usually measured in one of two ways, depending on the rate of exchange. When the rate is comparable to or faster than the spin–lattice relaxation rate ($k_{ex} > 0.1\,s^{-1}$), the rate constant is most easily determined from a saturation transfer experiment (analogous to the transient NOE experiment discussed in Section 5.1.2) (*105, 106*). For slower rates

($k_{ex} < 0.01\,s^{-1}$), exchange usually is measured by rapidly transferring the protein from $H_2O$ into $D_2O$ solution, and repeatedly acquiring homonuclear TOCSY (Section 6.5.5) or $^1H$–$^{15}N$ HSQC (Section 7.1.1.2) spectra to observe the decrease in amide proton resonance intensities with time. Observation of a slow amide proton exchange rate implies that the $^1H^N$ atom may be involved in a hydrogen bond, but does not identify the atoms acting as hydrogen bond acceptors (and cannot exclude the possibility that the reduced exchange rate results from steric effects rather than hydrogen bonding). Hydrogen bond restraints have a large impact on the nature and precision of the resulting structures and are usually only enforced in well-defined regions of regular secondary structure, in which only one possible hydrogen bond acceptor is consistent with the NOE data.

*10.2.1.6 Hydrogen Bond Restraints from Trans-Hydrogen Bond Scalar Coupling Constants* Recently, *trans*-hydrogen bond scalar coupling interactions have been observed in nucleic acids (*107, 108*) and proteins (*74*) by NMR spectroscopy. In proteins, the N–H $\cdots$ O=C hydrogen bond between backbone amide and carbonyl moieties is characterized by $^{3h}J_{NC'}$, $^{2h}J_{HC'}$, and $^{3h}J_{HC^\alpha}$ *trans*-hydrogen bonds with absolute magnitudes of 0–1 Hz. These can be detected by modifications of conventional triple-resonance experiments. For example, evolution due to the $^{3h}J_{NC'}$ scalar coupling interaction can be detected in an HNCO experiment in which the $^{15}N_i$–$^{13}C'_{i-1}$ dephasing time is set to $2/J_{NC'} = 133$ ms (for a 15-Hz coupling constant) (*74*). The cross-peaks obtained from the usual one-bond scalar coupling between the $^{15}N_i$–$^{13}C'_{i-1}$ nuclei are suppressed and the cross-peaks from the *trans*-hydrogen bond coupling interactions are observed. Variants of these experiments have been performed on proteins up to 30 kDa in mass. The observation of *trans*-hydrogen cross-peaks unambiguously identifies the acceptor group for the hydrogen bond, not possible with amide exchange measurements, and the size of the coupling constant is a strong function of the geometry of the hydrogen bond (*109*).

## 10.2.2 STRUCTURE DETERMINATION

Details of the local backbone geometry can be obtained by an extension of the sequential assignment process; the relative intensities of $d_{NN}$, $d_{\alpha N}$, and $d_{\beta N}$ NOE cross-peaks and the measurement of the backbone $^3J_{H^NH^\alpha}$ are required. The observation of intense $d_{NN}$ NOEs and small $^3J_{H^NH^\alpha}$ coupling constants ($< 6.0$ Hz) are indicative of helical or turn sections of polypeptide (e.g., see Fig. 10.6); observation of

intense $d_{\alpha N}$, weak $d_{NN}$, and $d_{\beta N}$ NOEs and large $^3J_{H^N H^\alpha}$ coupling constants ($>8.0$ Hz) are indicative of extended $\beta$-strands of polypeptide (*15*). The combination of sequential NOE and $^3J_{H^N H^\alpha}$ coupling constant data with medium-range and a few long-range NOEs is capable of providing details of the regions of regular secondary structure within the protein. The elements of secondary structures can be connected together to give a crude view of the global fold by the identification of a few key long-range NOEs, as shown for ubiquitin in Fig. 10.11. Thus, without recourse to extensive calculations, important structural results (albeit of low absolute resolution) can be obtained in a straightforward manner.

A variety of methods have been developed to calculate atomic-resolution protein structures using restraints derived from experimental NMR data [for reviews, see Güntert (*35, 68*) and Grishaev and

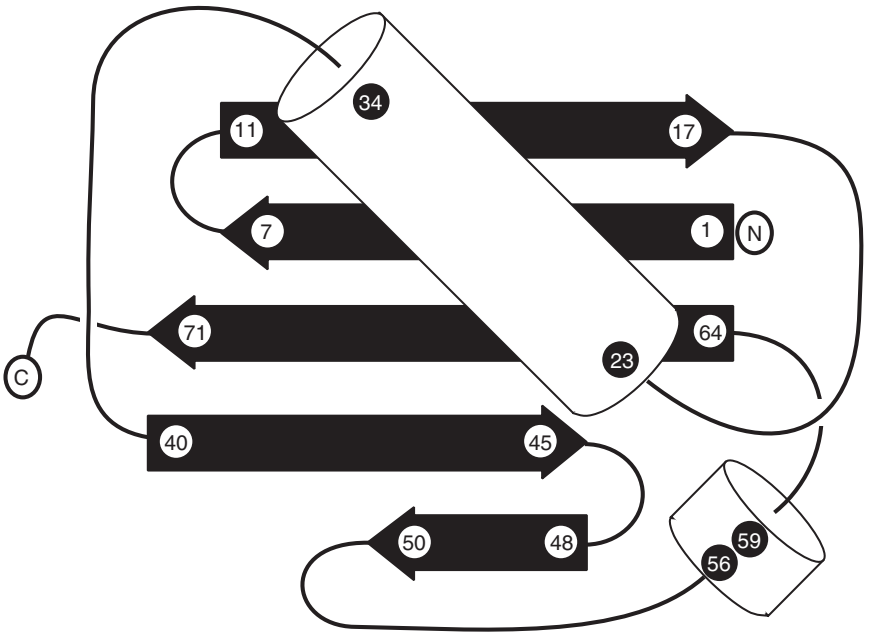



FIGURE 10.11 Definition of the global fold using long-range NOEs. Previously identified elements of regular secondary structure (see Fig. 10.6), can be arranged in a low-resolution global fold by the observation of a few key long-range NOEs involving protons at the termini of the strands and helices. This figure depicts the strands of $\beta$-sheet and sections of helix present in ubiquitin as black arrows and white cylinders, respectively.

Llinás (*67*)]. Importantly, NMR data do not uniquely define the three-dimensional structure of a protein or other biological macromolecule, because the restraints are included as ranges of allowed values, the data contain experimental uncertainties, and only a sparse subset of all possible restraints are observable. To increase the efficiency and accuracy of structure calculations, the experimentally derived restraints normally are supplemented by restraints specifically imposed to enforce proper covalent structure of the protein, including bond lengths, bond angles, and other elements of standard covalent geometry (chirality and the planarity of aromatic rings and peptide units). Protocols for structure determination aim to find coordinates for the protein atoms that will satisfy the input restraints in an unbiased fashion while exploring all of the regions of conformational space compatible with these restraints. Because of these considerations, structure calculations are repeated many times to determine an ensemble of (low energy) structures consistent with the input NMR data. Thus, a "good" ensemble of structures minimizes violations of the input restraints and maximizes the RMSD between members of the ensemble (*79, 110*).

The two most common approaches to generation of structures are distance geometry (DG) and restrained molecular dynamics (rMD). Historically, DG was the first approach utilized for structure determination; at an intermediate stage of development, DG frequently was used to generate initial structures for subsequent refinement by rMD methods. In modern approaches for structure determination, rMD has become the predominant technique. However, other approaches to structure determination continue to be pursued and future developments can be expected (*111*).

Popular implementations of DG use either the metric matrix algorithm (*112, 113*) or the variable target function approach (*114, 115*). Distance geometry determines ensembles of three-dimensional structures consistent with an incomplete set of distance restraints. The restraints are incomplete because not all distances can be characterized (the NOE is limited to distances less than approximately 5 Å) and because the distance restraints are not known precisely. The metric matrix algorithms in particular tend to be computationally expensive as the size of the protein increases.

Restrained molecular dynamics algorithms use either Cartesian or torsion-angle coordinate systems (*68*). Torsion-angle rMD has become the preferred method due to advances in computational algorithms. In either approach, molecular dynamics force fields are supplemented by pseudo-energy terms based on the NMR-derived restraints (*116, 117*).

These potentials drive the structure toward a conformation that will reduce the violation of the restraints during a forced heat-up and cool-down annealing cycle. The most computationally efficient implementations of the rMD method use a simplified force field in which bond length, bond angle, and repulsive van der Waals terms are retained (electrostatic and attractive van der Waals terms are ignored), and are referred to as dynamical simulated annealing (SA) (*118*). Due to advances in computational power, structures determined using simplified force fields now frequently are refined using complete force fields and including explicit or implicit solvent models (*119, 120*).

As has already been discussed (Sections 10.1.1 and 10.1.2), considerable efforts have been made to partially or fully automate the process of resonance assignments. Extensive efforts also are being made to automate the process of structure determination. Most automated structure calculation programs take as input a (sufficiently complete) list of resonance assignments and one or more lists of cross-peak positions and volumes from *n*D NOESY spectra. The programs then automatically assign the NOESY cross-peaks and calculate the three-dimensional structure of the protein. Current state-of-the-art methods for automated structure determination have been reviewed (*33–36, 67*). A comparison of "conventional" structure determination protocols with an optimized pipeline consisting of fast data acquisition (Section 9.3), automated resonance assignments, and automated structure calculation has been reported by Szyperski and co-workers (*121*).

Assessment of the resultant structures is an important aspect of any structure determination by NMR spectroscopy. The RMSD of the final ensemble of structures is a poor measure for judging the outcome of a structure calculation. Instead, quality of the calculated three-dimensional structures is assessed by two general measures: (*i*) how well the structures agree with the experimental restraint data and (*ii*) the geometrical reasonableness of the structures — for example, compared with expected stereochemistry derived from high-resolution *x*-ray crystal structures [for a review, see Spronk et al. (*122*)]. Consistency with experimental results commonly is judged by examining restraint pseudo-energies from the structure refinement, root-mean-square violations of restraints, and maximum consistent violations within the ensemble of structures. Stereochemical quality commonly is judged by quantifying the distributions of backbone and side chain dihedral angles, the number of van der Waals steric clashes, and packing of core residues. Programs such as PROCHECK (*123*), PROCHECK-NMR

(*124*), and WHATIF (*125*) are widely utilized for structure assessment. A number of new approaches for structure assessment have been developed using concepts from information theory (*126, 127*). Structures also can be validated independently against sets of experimental data that were not used in the structure calculation. Both RDCs (*128*) and anisotropic chemical shifts (*129*) have been suggested for this purpose by defining a figure of merit, or $Q$ factor.

A powerful approach for structure assessment, and potentially refinement, is known as complete relaxation matrix analysis (CORMA) (*78*), or back-calculation. Once three-dimensional structures have been calculated, interproton distances can be calculated from the atomic coordinates, the dipolar relaxation rate matrix can be calculated from the distances, and the theoretically expected NOE intensities can be calculated from the rate matrix using the results of Chapter 5 to generate theoretically expected NOESY spectra. After suitable scaling, the calculated NOE intensities are compared to the experimentally observed NOE intensities to judge how well the structures reproduce the experimental data. Usually, the agreement is judged by a figure of merit, or $R$ factor (*130*). The difference between observed and back-calculated NOE intensities also can be included as a pseudo-energy term during rMD calculations (*131, 132*). Thus, the structures are driven during the calculation to minimize the $R$ factor and maximize the agreement between the structures and the experimental NOE intensities. This process requires a complete relaxation matrix analysis to be performed at every step of the annealing cycle and therefore is computationally demanding. In addition, the dynamical model must be capable of reproducing the effects of molecular motion known to be present in proteins, such as fast rotation of methyl groups and 180° flips of aromatic side chains (Section 5.3) (*133*) and systematic deficiencies in the experimental data (*134*). However, back-calculation provides one of the few ways to independently assess the accuracy of protein structures derived from NMR data.

As an example, the structure of ubiquitin determined by NMR spectroscopy by Bax and co-workers is shown in Fig. 10.12. This structure was determined from 2727 NOEs, 98 dihedral angle restraints derived from homo- and heteronuclear $J$ couplings, and 372 RDC restraints (*129*). The high degree of agreement between experimental RDCs and values calculated from the final structure is shown in Fig. 10.13. These RDCs were used for the structure refinement and are not independent figures of merit for the stucture determination. The structure quality was assessed using the $Q$ factor calculated from the measured and calculated anisotropic carbonyl

chemical shifts (*129*),

$$Q = \left[\frac{\sum \left(\Delta\delta - \widehat{\Delta\delta}\right)^2}{\sum \Delta\delta^2}\right]^{1/2}, \qquad [10.3]$$

in which $\Delta\delta$ and $\widehat{\Delta\delta}$ are the measured and predicted differences in the $^{13}CO$ chemical shift measured in isotropic and anisotropic solution, and the summation extends over the backbone residues. The value of the $Q$ factor is 16%; for comparison, the $Q$ factor is 25% if the structure is calculated without inclusion of the RDCs and is 23% for the 1.8-Å crystal structure of ubiquitin (*135*).

Ubiquitin also has been the subject of many investigations of protein conformational dynamics by NMR spectroscopy, including studies of spin relaxation, RDCs, and amide proton–solvent exchange. As only one example of this voluminous literature, the backbone amide order parameters determined from $^{15}N$ spin relaxation rate constants are shown in Fig. 10.14 (*136*). As shown, and consistent with the ensemble of structures shown in Fig. 10.12, the backbone is highly constrained,

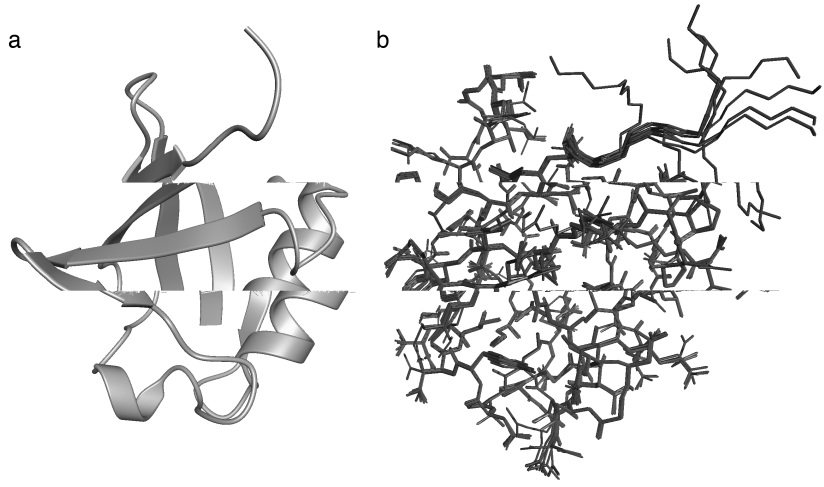



FIGURE 10.12 The NMR-derived structure of ubiquitin. (a) Backbone chain trace for the first member of the ensemble of 10 structures in PDB file 1d3z. (b) Superposition of the 10 members of the structure ensemble showing backbone and side chain heavy atoms. Side chain atoms are not shown for the disordered C-terminal residues 71–77 of ubiquitin for clarity.

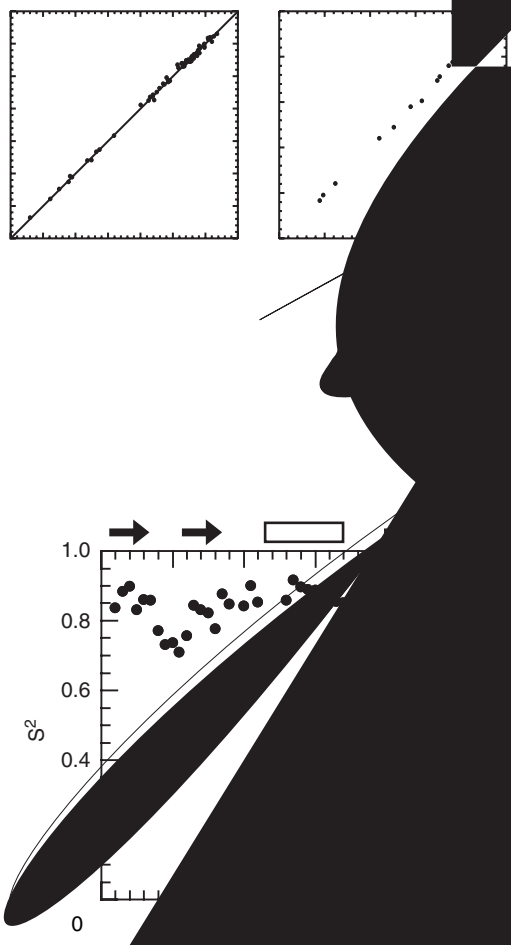



FIGURE 10.14 [illegible]
Experimental [illegible]
Secondary st[illegible]

with r[illegible]
C-te[illegible]
the[illegible]
i[illegible] [illegible]egio[illegible]
[illegible]nationa[illegible]

and disordered regions that arise from insufficient restraints in the structure calculation.

## 10.3 Conclusion

This text has two objectives: (*i*) presentation of the theory of NMR spectroscopy applicable to studies of biological macromolecules and (*ii*) explication of the types of NMR experiments utilized for assignments of resonances, determinations of three-dimensional structures, and investigations of dynamics and interactions of proteins. Knowledge of the three-dimensional structure of a protein, whether determined by *x*-ray crystallography or by NMR spectroscopy, is the beginning of structural biology and biophysics of protein function, not the endpoint. Proteins are not rigid isolated molecules, and understanding protein function depends upon a detailed knowledge of structure, dynamics, and interactions with other biomolecules. As indicated briefly in Chapters 8 and 9, NMR spectroscopy is a powerful approach for investigations of the intramolecular dynamical properties (fluctuations about the time-average structures) of proteins and of structural and kinetic aspects of intermolecular interactions.

The experiments discussed in this text have a certain universality (the topology of amino acid spin systems and secondary structure elements serve as unifying features) and are widely used in studies of proteins by NMR spectroscopy. In the interest of brevity, this text does not discuss the complete wealth of NMR experiments that have been developed and continue to be developed for investigations of proteins and other biological macromolecules. Nonetheless, the theoretical and experimental principles presented herein are equally valid and important for the understanding and optimization of all experiments for high-resolution NMR spectroscopy of biomolecules in solution.

Since the first reports of the observation of nuclear magnetic resonance absorption in condensed phase in 1946, nuclear magnetic resonance has developed into a major technique in biology, chemistry, and medicine. Yet, surveying the developments of the past decade suggests that the "Golden Age" of biological NMR spectroscopy shows no end. Continued advances in sample preparation, magnet and rf technology, NMR spectroscopic techniques, and computational methods promise to make the next 60 years as exciting as the last.

# References


1. S. Schwarzinger, G. J. A. Kroon, T. R. Foss, P. E. Wright, H. J. Dyson, *J. Biomol. NMR* **18**, 43–48 (2000).
2. S. Schwarzinger, G. J. Kroon, T. R. Foss, J. Chung, P. E. Wright, H. J. Dyson, *J. Am. Chem. Soc.* **123**, 2970–2978 (2001).
3. A. Bundi, K. Wüthrich, *Biopolymers* **18**, 285–297 (1979).
4. D. S. Wishart, B. D. Sykes, *Meth. Enzymol.* **239**, 363–392 (1994).
5. D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **222**, 311–333 (1991).
6. K.-H. Gross, H. R. Kalbitzer, *J. Magn. Reson.* **76**, 87–99 (1988).
7. D. C. Dalgarno, B. A. Levine, R. J. P. Williams, *Biosci. Rep.* **3**, 443–452 (1983).
8. A. Pastore, V. Saudek, *J. Magn. Reson.* **90**, 165–176 (1990).
9. M. P. Williamson, *Biopolymers* **29**, 1423–1431 (1990).
10. D. S. Wishart, B. D. Sykes, F. M. Richards, *Biochemistry* **31**, 1647–1651 (1992).
11. D. S. Wishart, D. A. Case, *Meth. Enzymol.* **338**, 3–34 (2001).
12. B. R. Seavey, E. A. Farr, W. M. Westler, J. L. Markley, *J. Biomol. NMR* **1**, 217–236 (1991).
13. W. J. Chazin, P. E. Wright, *Biopolymers* **26**, 973–977 (1987).
14. W. J. Chazin, P. E. Wright, *J. Mol. Biol.* **202**, 623–636 (1988).
15. K. Wüthrich, "NMR of Proteins and Nucleic Acids," pp. 1–292. Wiley, New York, 1986.
16. A. D. Kline, W. Braun, K. Wüthrich, *J. Mol. Biol.* **204**, 675–724 (1988).
17. M. Billeter, W. Braun, K. Wüthrich, *J. Mol. Biol.* **155**, 321–346 (1982).
18. K. Wüthrich, *Biopolymers* **22**, 131–138 (1982).
19. S. W. Englander, A. J. Wand, *Biochemistry* **26**, 5953–5958 (1987).
20. D. L. Di Stefano, A. J. Wand, *Biochemistry* **26**, 7272–7281 (1987).
21. G. J. Kleywegt, R. M. N. J. Lamerichs, R. Boelens, R. Kaptein, *J. Magn. Reson.* **85**, 186–197 (1989).
22. F. M. van de Ven, *J. Magn. Reson.* **86**, 633–644 (1990).
23. J. Xu, S. K. Strauss, B. C. Sanctuary, L. Trimble, *J. Magn. Reson., Ser. B* **103**, 53–58 (1994).
24. Z. Z. Chen, T. Jiang, G. H. Lin, J. J. Wen, D. Xu, Y. Xu, in "Lecture Notes in Computer Science" (R. Guigó, D. Gusfield, eds.), Vol. 2452, pp. 83–96. Springer-Verlag GmbH, Heidelberg, 2002.
25. P. J. Kraulis, *J. Magn. Reson.* **84**, 627–633 (1989).
26. M. Billeter, V. J. Basus, I. D. Kuntz, *J. Magn. Reson.* **76**, 400–415 (1988).
27. C. Eccles, P. Güntert, M. Billeter, K. Wüthrich, *J. Biomol. NMR* **1**, 111–130 (1991).
28. G. J. Kleywegt, R. Boelens, R. Kaptein, *J. Magn. Reson.* **88**, 601–608 (1990).
29. C. Cieslar, T. A. Holak, H. Oschkinat, *J. Magn. Reson.* **89**, 184–190 (1990).
30. G. W. Vuister, R. Boelens, A. Padilla, R. Kaptein, *J. Biomol. NMR* **1**, 421–438 (1991).
31. R. Bernstein, A. Ross, C. Cieslar, T. A. Holak, *J. Magn. Reson., Ser. B* **101**, 185–188 (1993).
32. H. Oschkinat, T. A. Holak, C. Ciesler, *Biopolymers* **31**, 699–712 (1991).
33. M. C. Baran, Y. J. Huang, H. N. Moseley, G. T. Montelione, *Chem. Rev.* **104**, 3541–3556 (2004).
34. A. S. Altieri, R. A. Byrd, *Curr. Opin. Struct. Biol.* **14**, 547–553 (2004).
35. P. Güntert, *Prog. NMR Spectrosc.* **43**, 105–125 (2003).
36. W. Gronwald, H. R. Kalbitzer, *Prog. NMR Spectrosc.* **44**, 33–96 (2004).
37. W. J. Chazin, M. Rance, P. E. Wright, *J. Mol. Biol.* **202**, 603–622 (1988).
38. R. A. Williamson, G. Mortorell, M. D. Carr, G. Murphy, A. J. P. Docherty, R. B. Freedman, J. Feeney, *Biochemistry* **33**, 11745–11759 (1994).

39. P. L. Weber, S. C. Brown, L. Mueller, *Biochemistry* **26**, 7282–7290 (1987).
40. P. C. Driscoll, G. M. Clore, D. Marion, P. T. Wingfield, A.M. Gronenborn, *Biochemistry* **29**, 3542–3556 (1990).
41. G. M. Clore, A. Bax, P. C. Driscoll, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* **29**, 8172–8184 (1990).
42. M. D. Carr, B. Birdsall, T. A. Frenkiel, C. J. Bauer, J. Jimenez-Barbero, V. I. Polshakov, J. E. McCormick, G. C. K. Roberts, J. Feeney, *Biochemistry* **30**, 6330–6341 (1991).
43. R. T. Clubb, V. Thanabal, C. Osborne, G. Wagner, *Biochemistry* **30**, 7718–7730 (1991).
44. W. J. Fairbrother, J. Cavanagh, H. J. Dyson, A. G. Palmer, S. L. Sutrina, J. Reizer, M. H. Saier, P. E. Wright, *Biochemistry* **30**, 6896–6907 (1991).
45. W. J. Fairbrother, A. G. Palmer, M. Rance, J. Reizer, M. H. Saier, P. E. Wright, *Biochemistry* **31**, 4413–4425 (1992).
46. B. J. Stockman, N. R. Nirmala, G. Wagner, T. J. Delcamp, M. T. DeYarman, J. H. Freisheim, *Biochemistry* **31**, 218–229 (1992).
47. M. Ikura, L. E. Kay, A. Bax, *Biochemistry* **29**, 4659–4667 (1990).
48. G. T. Montelione, B. A. Lyons, S. D. Emerson, M. Tashiro, *J. Am. Chem. Soc.* **114**, 10974–10975 (1992).
49. S. Grzesiek, J. Anglister, A. Bax, *J. Magn. Reson., Ser. B* **101**, 114–119 (1993).
50. R. T. Clowes, W. Boucher, C. H. Hardman, P. J. Domaille, E. D. Laue, *J. Biomol. NMR* **3**, 349–354 (1993).
51. B. A. Lyons, M. Tashiro, L. Cedergren, B. Nilsson, G. T. Montelione, *Biochemistry* **32**, 7839–7845 (1993).
52. T. M. Logan, E. T. Olejniczak, R. X. Xu, S. W. Fesik, *J. Biomol. NMR* **3**, 225–231 (1993).
53. B. T. Farmer, R. A. Venters, *J. Am. Chem. Soc.* **117**, 4187–4188 (1995).
54. T. Torizawa, A. M. Ono, T. Terauchi, M. Kainosho, *J. Am. Chem. Soc.* **127**, 12620–12626 (2005).
55. J. J. Prompers, A. Groenewegen, C. W. Hilbers, H. A. M. Pepermans, *J. Magn. Reson.* **130**, 68–75 (1998).
56. B. T. Farmer, R. A. Venters, *J. Biomol. NMR* **7**, 59–71 (1996).
57. T. Yamazaki, S. M. Pascal, A. U. Singer, J. D. Forman-Kay, L. E. Kay, *J. Am. Chem. Soc.* **117**, 3556–3564 (1995).
58. F. Lohr, S. Pfeiffer, Y. J. Lin, J. Hartleib, O. Klimmek, H. Rüterjans, *J. Biomol. NMR* **18**, 337–346 (2000).
59. V. Kanelis, L. Donaldson, D. R. Muhandiram, D. Rotin, J. D. Forman-Kay, L. E. Kay, *J. Biomol. NMR* **16**, 253–259 (2000).
60. D. Yang, L. E. Kay, *J. Am. Chem. Soc.* **121**, 2571–2575 (1999).
61. V. Tugarinov, R. Muhandiram, A. Ayed, L. E. Kay, *J. Am. Chem. Soc.* **124**, 10025–10035 (2002).
62. V. Tugarinov, L. E. Kay, *J. Am. Chem. Soc.* **125**, 13868–13878 (2003).
63. R. Richarz, K. Wüthrich, *Biopolymers* **17**, 2133–2141 (1978).
64. D. S. Wishart, B. D. Sykes, *J. Biomol. NMR* **4**, 171–180 (1994).
65. S. Spera, A. Bax, *J. Am. Chem. Soc.* **113**, 5490–5492 (1991).
66. G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **13**, 289–302 (1999).
67. A. Grishaev, M. Llinás, *Meth. Enzymol.* **394**, 261–295 (2005).
68. P. Güntert, *Prog. NMR Spectrosc.* **31**, 145–237 (1998).
69. A. G. Palmer, *Chem. Rev.* **104**, 3623–3640 (2004).
70. T. Carlomagno, *Annu. Rev. Biophys. Biomol. Struct.* **34**, 245–266 (2005).
71. M. P. Williamson, T. F. Havel, K. Wüthrich, *J. Mol. Biol.* **182**, 295–315 (1985).
72. N. Tjandra, A. Bax, *Science* **278**, 1111–1114 (1997).
73. G. Wagner, *Q. Rev. Biophys.* **16**, 1–57 (1983).

74. Y.-X. Wang, J. Jacob, F. Cordier, P. Wingfield, S. J. Stahl, S. Lee-Huang, D. A. Torchia, S. Grzesiek, A. Bax, *J. Biomol. NMR* **14**, 181–184 (1999).
75. V. Gaponenko, S. P. Sarma, A. S. Altieri, D. A. Horita, J. Li, R. A. Byrd, *J. Biomol. NMR* **28**, 205–212 (2004).
76. I. Bertini, C. Luchinat, M. Piccioli, *Meth. Enzymol.* **339**, 314–340 (2001).
77. J. Iwahara, C. D. Schwieters, G. M. Clore, *J. Am. Chem. Soc.* **126**, 5879–5896 (2004).
78. J. W. Keepers, T. L. James, *J. Magn. Reson.* **57**, 404–426 (1984).
79. T. Havel, in "Proteins: Structure, Dynamics, Design" (V. Renugopalakrishnan, P. R. Carey, I. C. P. Smith, S.-G. Huans, A. L. Storer, eds.), pp. 110–115. ESCOM Science Publishers, Leiden, Holland, 1991.
80. P. Güntert, W. Braun, M. Billeter, K. Wüthrich, *J. Am. Chem. Soc.* **111**, 3997–4004 (1989).
81. J. F. Doreleijers, M. L. Raves, T. Rullmann, R. Kaptein, *J. Biomol. NMR* **14**, 123–132 (1999).
82. M. Karplus, *J. Phys. Chem.* **30**, 11–15 (1959).
83. V. F. Bystrov, *Prog. Nucl. Magn. Reson.* **10**, 41–82 (1976).
84. A. Pardi, M. Billeter, K. Wüthrich, *J. Mol. Biol.* **180**, 741–751 (1984).
85. G. Wagner, W. Braun, T. F. Havel, T. Schaumann, N. Go, K. Wüthrich, *J. Mol. Biol.* **196**, 611–639 (1987).
86. G. W. Vuister, M. Tessari, Y. Karimi-Nejad, B. Whitehead, in "Biological Magnetic Resonance, 16: Modern Techniques in Protein NMR" (N. R. Krisha, L. Berliner, eds.), pp. 195–257. Kluwer Academic, New York, 1999.
87. M. J. Minch, *Concepts Magn. Reson.* **6**, 41–56 (1994).
88. L. J. Smith, M. J. Sutcliffe, C. Redfield, C. M. Dobson, *Biochemistry* **30**, 986–996 (1991).
89. S. Ludvigsen, K. V. Andersen, F. M. Poulsen, *J. Mol. Biol.* **217**, 731–736 (1991).
90. G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **115**, 7772–7777 (1993).
91. C. M. Deber, D. A. Torchia, E. R. Blout, *J. Am. Chem. Soc.* **93**, 4893–4897 (1971).
92. A. DeMarco, M. Llinás, K. Wüthrich, *Biopolymers* **17**, 617–636 (1978).
93. J. M. Schmidt, M. Blümel, F. Löhr, H. Rüterjans, *J. Biomol. NMR* **14**, 1–12 (1999).
94. M. Habeck, W. Rieping, M. Nilges, *J. Magn. Reson.* **177**, 160–165 (2005).
95. C. Peréz, F. Löhr, H. Rüterjans, J. Schmidt, *J. Am. Chem. Soc.* **123**, 7081–7093 (2001).
96. G. S. Harbison, *J. Am. Chem. Soc.* **115**, 3026–3027 (1993).
97. R. Brüschweiler, D. Case, *J. Am. Chem. Soc.* **116**, 11199–11200 (1994).
98. S. G. Hyberts, W. Märki, G. Wagner, *Eur. J. Biochem.* **164**, 625–635 (1987).
99. B. Reif, A. Diener, M. Hennig, M. Maurer, C. Griesinger, *J. Magn. Reson.* **143**, 45–68 (2000).
100. D. Sitkoff, D. A. Case, *Prog. NMR Spectrosc.* **32**, 165–190 (1998).
101. G. M. Clore, A. M. Gronenborn, N. Tjandra, *J. Magn. Reson.* **131**, 159–162 (1998).
102. M. W. F. Fischer, J. A. Losonczi, J. L. Weaver, J. H. Prestegard, *Biochemistry* **38**, 9013–9022 (1999).
103. G. A. Mueller, W. Y. Choy, D. Yang, J. D. Forman-Kay, R. A. Venters, L. E. Kay, *J. Mol. Biol.* **300**, 197–212 (2000).
104. A. Bax, A. Grishaev, *Curr. Opin. Struct. Biol.* **15**, 563–570 (2005).
105. S. Forsén, R. A. Hoffman, *Acta Chem. Scand.* **17**, 1787 (1963).
106. S. Forsén, R. A. Hoffman, *J. Chem. Phys.* **39**, 2892 (1963).
107. A. J. Dingley, S. Grzesiek, *J. Am. Chem. Soc.* **120**, 8293–8297 (1998).
108. K. Pervushin, A. Ono, C. Fernández, T. Szyperski, M. Kainosho, K. Wüthrich, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14147–14151 (1998).
109. G. Cornilescu, B. E. Ramirez, M. K. Frank, G. M. Clore, A. M. Gronenborn, A. Bax, *J. Am. Chem. Soc.* **121**, 6275–6279 (1999).
110. S. Hyberts, M. S. Goldberg, T. F. Havel, G. Wagner, *Protein Sci.* **1**, 736–751 (1992).

111. W. Rieping, M. Habeck, M. Nilges, *Science* **309**, 303–306 (2005).
112. G. M. Crippen, T. F. Havel, "Distance Geometry and Molecular Conformation," pp. 1–58. Research Studies Press, Taunton, England, 1988.
113. T. F. Havel, *Prog. Biophys. Mol. Biol.* **56**, 43–78 (1991).
114. W. Braun, N. Go, *J. Mol. Biol.* **186**, 611–626 (1985).
115. P. Güntert, W. Braun, K. Wüthrich, *J. Mol. Biol.* **217**, 517–530 (1991).
116. G. M. Clore, A. T. Brünger, M. Karplus, A. M. Gronenborn, *J. Mol. Biol.* **191**, 523–551 (1986).
117. A. T. Brünger, G. M. Clore, A. M. Gronenborn, M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3801–3805 (1986).
118. M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* **239**, 129–136 (1988).
119. B. Xia, V. Tsui, D. A. Case, H. J. Dyson, P. E. Wright, *J. Biomol. NMR* **22**, 317–331 (2002).
120. J. P. Linge, M. A. Williams, C. A. Spronk, A. M. Bonvin, M. Nilges, *Proteins* **50**, 496–506 (2003).
121. G. Liu, Y. Shen, H. S. Atreya, D. Parish, Y. Shao, D. K. Sukumaran, R. Xiao, A. Yee, A. Lemak, A. Bhattacharya, T. A. Acton, C. H. Arrowsmith, G. T. Montelione, T. Szyperski, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10487–10492 (2005).
122. C. A. E. M. Spronk, S. B. Nabuurs, E. Krieger, G. Vriend, G. W. Vuister, *Prog. NMR Spectrosc.* **45**, 315–337 (2004).
123. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Cryst.* **26**, 283–291 (1993).
124. R. A. Laskowski, J. A. C. Rullmann, M. W. MacArthur, R. Kaptein, J. M. Thornton, *J. Biomol. NMR* **8**, 477–486 (1996).
125. G. Vriend, *J. Mol. Graph.* **8**, 52–56 (1990).
126. S. B. Nabuurs, E. Krieger, C. A. Spronk, A. J. Nederveen, G. Vriend, G. W. Vuister, *J. Biomol. NMR* **33**, 123–134 (2005).
127. Y. J. Huang, R. Powers, G. T. Montelione, *J. Am. Chem. Soc.* **127**, 1665–1674 (2005).
128. A. Bax, *Protein Sci.* **12**, 1–16 (2003).
129. G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **120**, 6836–6837 (1998).
130. C. Gonzalez, J. A. C. Rullmann, A. M. J. J. Bonvin, R. Boelens, R. Kaptein, *J. Magn. Reson.* **91**, 659–664 (1991).
131. P. Yip, D. A. Case, *J. Magn. Res.* **83**, 643–648 (1989).
132. A. M. J. J. Bonvin, R. Boelens, R. Kaptein, *J. Biomol. NMR* **1**, 305–309 (1991).
133. H. Liu, P. D. Thomas, T. J. James, *J. Magn. Reson.* **98**, 163–175 (1992).
134. M. J. Dellwo, D. M. Schneider, A. J. Wand, *J. Magn. Reson., Ser. B* **103**, 1–9 (1994).
135. S. Vijay-Kumar, C. E. Bugg, W. J. J. Cook, *J. Mol. Biol.* **194**, 531–544 (1987).
136. N. Tjandra, S. E. Feller, R. W. Pastor, A. Bax, *J. Am. Chem. Soc.* **117**, 12562–12566 (1995).
137. J.-S. Hu, A. Bax, *J. Am. Chem. Soc.* **119**, 6360–6368 (1997).
138. A. C. Wang, A. Bax, *J. Am. Chem. Soc.* **117**, 1810–1813 (1995).
139. F. Löhr, J. M. Schmidt, S. Maurer, H. Rüterjans, *J. Magn. Reson.* **153**, 75–82 (2001).
140. J. Schmidt, *J. Magn. Reson.* **124**, 310–322 (1997).
141. A. DeMarco, M. Llinás, K. Wüthrich, *Biopolymers* **17**, 2727–2742 (1978).
142. A. J. Fischman, D. H. Live, H. R. Wyssbrod, W. C. Agosta, D. Cowburn, *J. Am. Chem. Soc.* **102**, 2533–2539 (1980).
143. D. Braun, G. Wider, K. Wüthrich, *J. Am. Chem. Soc.* **116**, 8466–8469 (1994).