LEAD ARTICLE

Acta Cryst. (1995). D**51**, 249–270

NMR – This Other Method for Protein and Nucleic Acid Structure Determination

BY KURT WÜTHRICH

Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland

(*Received* 29 *July* 1994; *accepted* 5 *September* 1994)

Abstract

For a quarter of a century X-ray diffraction in single crystals was unique in its ability to solve three-dimensional structures of proteins and nucleic acids at atomic resolution. The situation changed in 1984 with the completion of a protein structure determination by nuclear magnetic resonance (NMR) spectroscopy in solution, and today NMR is a second widely used method for biomacromolecular structure determination. This review describes the method of NMR structure determination of biological macromolecules, and attempts to place NMR structure determination in perspective with X-ray crystallography. NMR is most powerful for studies of relatively small systems with molecular weights up to about 30000, but these structures can be obtained in near-physiological milieus. The two techniques have widely different time scales which afford different insights into internal molecular mobility as well as different views of protein or nucleic acid molecular surfaces and hydration. Generally, in addition to information on the average three-dimensional structure, NMR provides information on a wide array of short-lived transient conformational states. Combining information from the two methods can yield a more detailed insight into the structural basis of protein and nucleic acid functions, and thus provide a more reliable platform for rational drug design and the engineering of novel protein functions.

Kurt Wüthrich was born in Switzerland on October 4, 1938. *He studied chemistry and physics at the University of Bern from* 1957 *to* 1962, *and obtained the Eidgenössisches Turn- und Sportlehrerdiplom and a PhD in inorganic chemistry at the University of Basel in* 1964. *He was a postdoctoral fellow at the University of California in Berkeley until* 1967, *and then a member of the technical staff at Bell Telephone Laboratories in Murray Hill, New Jersey. In* 1969 *he joined the ETH in Zurich, where he is now Professor of Biophysics. Wüthrich's scientific interests are in structural biology of proteins and nucleic acids, and in protein design in biological and biomedical research. His own work is focussed on the development of experimental techniques for structure determination of biological macromolecules in solution and for studies of molecular dynamics. In the late* 1960*'s his research was focused on NMR studies of the active sites of hemoproteins, and in the mid-*1970*'s he contributed key experimental observations on protein dynamics. From* 1977 *to* 1984 *his research group developed the now widely used NMR method for determination of three-dimensional protein structures in solution, and from* 1986 *to* 1991 *techniques were introduced that enable detailed studies of the hydration of proteins and nucleic acids in solution, and of the conformation of receptor-bound drug molecules.*

1. Introduction

In biological and biomedical research the area of structural biology, in particular atomic resolution studies of the three-dimensional structure of biological macromolecules and their intermolecular interactions, has never before had as central a role and enjoyed as much popularity as it does today. With the use of DNA recombination and chemical synthesis an unlimited array of different polypeptide sequences can be generated, but the relationship of the resulting primary structures with corresponding biological functions can only be rationalized through knowledge of the three-dimensional structure. In addition to its key role in academic institutions, structure determination of proteins, nucleic acids and other classes of biomolecules has, therefore, become a discipline that is also actively pursued in profit-oriented organizations, especially in the major pharmaceutical companies. Although X-ray diffraction with single crystals has for two decades, until 1984, been unique in its potential for efficient determination of three-dimensional molecular structures, it presently shares this role with nuclear magnetic resonance (NMR) spectroscopy in solution (Table 1).

Somewhat in contrast to the rather slow development of protein crystallography during the first decade after the initial structure determinations of myoglobin (Kendrew, 1963) and hemoglobin (Perutz, 1963) in the late 1950's (Dickerson & Geis, 1969), the first protein structure determination by NMR in 1984 (Williamson, Havel & Wüthrich, 1985) (Fig. 1) was followed by a rapid evolution of NMR structure determination from about 1987 onward. Starting in 1990, the publication of an annual survey of new experimental macromolecular

Table 1. *New three-dimensional structures of biological macromolecules published during the period* 1990–1993

Structures determined by X-ray crystallography, NMR or by any other method are listed separately. The data are from Hendrickson & Wüthrich (1991, 1992, 1993, 1994).

Year	X-rays (single crystals)	NMR (solution)	Other methods
1990	109	23	2
1991	123	36	—
1992	168	61	—
1993	207	59	—

structure determinations (Hendrickson & Wüthrich, 1991, 1992, 1993, 1994) provides a guide to the contributions by the different techniques available (Table 1). The few entries in the 'Other methods' column of Table 1 refer to structures solved by X-ray fibre diffraction or electron diffraction in crystals. These techniques can provide the desired information, but applications have mainly been focussed on highly complex systems so that the actual number of structures solved is still very small. Other experimental techniques that are traditionally associated with the structural biology of proteins and nucleic acids and have important roles in the empirical identification of different conformational states, *e.g.* circular dichroism and other spectroscopic methods, or measurements based on studies of thermodynamic or hydrodynamic properties of macromolecules in solution, do not appear in Table 1 because they cannot provide sufficient data for *de novo* three-dimensional structure determinations. Table 1 thus leads to the conclusion that X-ray diffraction in crystals and NMR in solution are the only efficient methods presently available for biomacromolecular structure determination at atomic resolution. Considering that the number of structure determinations listed in Table 1 exceeds by far the total number of macromolecular structures solved up to 1989, we further learn that knowledge of three-dimensional structures is still very limited when compared with the available sequence information. Scarcity of three-dimensional structures is indeed a major bottleneck in protein engineering and rational drug design, but as is evident from Table 1, the situation is improving rapidly.

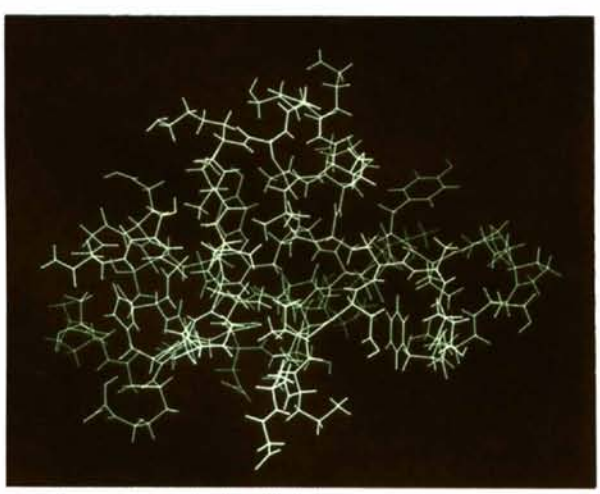

Fig. 1. NMR solution structure of the protein BUSI IIA (proteinase inhibitor IIA from bull seminal plasma). All heavy atoms of the 57-residue polypeptide chain are drawn. For the structure determination the protein had been dissolved in water. (Drawing prepared with the atomic coordinates from Williamson, Havel & Wüthrich, 1985.)

The present article presents a survey of the method of macromolecular structure determination by NMR (Wüthrich, 1986). It further attempts to place X-ray crystallography and NMR spectroscopy in solution into perspective with each other, with particular focus on the use of the two methods for protein structure determination. Foremost, I would like to entertain the question whether the numbers in Table 1 tell us everything. Does the introduction of NMR as a second method simply increase the total number of structures solved, for example, from 168 to 229 in 1992, or are there also fundamentally new insights indicating that refined or even entirely novel views of the structural basis of protein functions could result from combining the results obtained with the two techniques?

2. Some reminiscences: growing up besides X-ray crystallography

Over the years the fact that X-ray diffraction with protein or nucleic acid single crystals (Blundell & Johnson, 1976) and NMR spectroscopy with protein or nucleic acid solutions (Bax, 1989; Clore & Gronenborn, 1991; Kaptein, Boelens, Scheek & van Gunsteren, 1988; Wüthrich, 1976, 1986, 1989*a*,*b*) are both focussed on investigations of the three-dimensional molecular structures, has quite often resulted in exchanges determined by a sense of competition rather than by due emphasis on the complementary nature of the information obtained with the two techniques (see section 9 below). My personal experience was that of frequent contacts with leading colleagues in the field of protein crystallography, which greatly added to my postgraduate education in structural biology (my formal university education was in chemistry, physics and physical education). On many occasions I was invited to join protein crystallography meetings, and more than once I was the only 'outsider' in a crystallography school or similar. The flattering invitation to contribute this lead article to *Acta Crystallographica* is a continuation of these interdisciplinary contacts, which I accept as a very special privilege.

The physical phenomenon of nuclear magnetic resonance was first described in 1946 (Bloch, Hansen & Packard, 1946; Purcell, Pound & Bloembergen, 1946), and the first NMR spectrum of a protein was reported in 1957 (Saunders, Wishnia & Kirkwood, 1957). During the following decade there were a small number of

additional reports on NMR experiments with direct observation of the protein spectrum, but for technical reasons of sensitivity and spectral resolution most NMR studies focussed on measurements of the influence of proteins and nucleic acids on the solvent, using 'relaxation enhancement experiments' (for reviews of this early work see, for example, Ehrenberg, Malmström & Vänngard, 1967; Jardetzky, 1967; Mildvan & Cohn, 1970). Although some qualitative indications of internal flexibility of biopolymers in solution were obtained, and the usefulness of the method for certain analytical studies, such as titrations of individual histidyl side chains in proteins (*e.g.* Meadows, Jardetzky, Epand, Rüterjans & Scheraga, 1968) was convincingly demonstrated, these early NMR applications in structural biology could not, for technical reasons, bear directly on the three-dimensional structure. It was also too early for assessment of the NMR data in light of corresponding X-ray crystal structures, since these were yet to be solved.

During the second decade of NMR with biological macromolecules from 1967 to 1976, superconducting magnets providing higher polarizing magnetic fields, and the availability of commercial Fourier transform NMR spectrometers provided improved sensitivity and spectral resolution, enabling measurements relating directly to features of the three-dimensional molecular structures. In the course of this ten-year period, X-ray crystal structures were solved for several small proteins which were also amenable to NMR investigations, so that results obtained with the two methods could be compared directly. In this environment many fundamental features of protein ^1H NMR spectra could be rationalized, as is described in some technical detail in the following section 3, and it was soon clear that NMR investigations could provide important complementary information to that contained in X-ray crystal structures: (i) the NMR spectral features were usually in qualitative agreement with the assumption that at least part of the crystal structure was preserved in solution. (ii) Based on the concept of ring-current shifts (see section 3), conformational transitions could be characterized, for example, the reversible denaturation of hen egg-white lysozyme (McDonald & Phillips, 1967). In particularly favorable situations, mainly in hemoproteins where the immediate environment of the heme group is unique in its NMR spectral properties (Wüthrich, 1970), conformational changes observed by NMR could, by reference to the crystal structure, be attributed to unique locations on the molecular structure. An illustrative example is the observation of tertiary-structure changes upon ligand binding in myoglobin and hemoglobin (Shulman, Ogawa, Wüthrich, Yamane, Peisach & Blumberg, 1969). (iii) In the crystal structure of cytochrome *c* the second axial ligand of the heme Fe atom could initially not be identified from the X-ray data (Dickerson & Geis, 1969). Within the framework of the global polypeptide fold available from the crystallographic study, this missing structural detail was unambiguously determined by NMR (McDonald, Phillips & Vinograde, 1969; Wüthrich, 1969). (iv) ^{13}C NMR relaxation studies provided a qualitative picture of variable segmental mobility on the nanosecond time scale in different areas of globular protein molecules (Allerhand, Doddrell, Glushko, Cochran, Wenkert, Lawson & Gurd, 1971). (v) The proton-exchange reaction (1),

$$NH + {}^2H_2O \rightarrow N^2H + H^2HO, \qquad (1)$$

could be studied for individual backbone amide protons after dissolving the protein in ^2H_2O, providing information on hydrogen bonding and solvent accessibility of core regions of the protein. The basic pancreatic trypsin inhibitor (BPTI), which was one of the first proteins for which a high-resolution X-ray crystal structure became available (Deisenhofer & Steigemann, 1975), had a key role in the development of these studies (Karplus, Snyder & Sykes, 1973; Masson & Wüthrich, 1973). It is remarkable that corresponding NMR information on hydrogen bonding in tRNA was obtained in the early 1970's (Kearns, Patel & Shulman, 1971), years before an X-ray crystal structure determination of a tRNA was completed. (vi) The NMR observation of ring flips of phenylalanine and tyrosine provided precise data on concerted large-amplitude internal mobility in the protein core (Fig. 2). The observation of these ring-flipping motions on the millisecond to microsecond time scale (Wüthrich & Wagner, 1975) was a genuine surprise for the following reasons. In the refined X-ray crystal structure of BPTI (Deisenhofer & Steigemann, 1975) the aromatic rings of phenylalanine and tyrosine are among the side chains with the smallest temperature factors. For each ring the relative values of the *B* factors increase toward the periphery, so that the largest positional uncertainty is indicated for the C atom 4 on the symmetry axis through the $C^\beta—C^1$ bond rather than

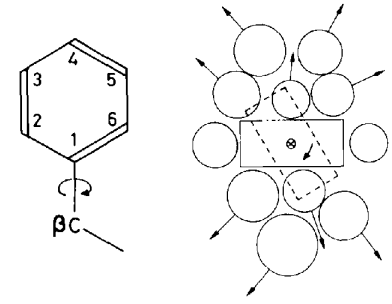

Fig. 2. 180° ring flips of tyrosine and phenylalanine about the $C^\beta—C^1$ bond. On the left the atom numeration is given and the χ_2 rotation axis is identified with an arrow. The drawing on the right represents a view along the $C^\beta—C^1$ bond of a flipping ring in the interior of a protein, where the broken lines indicate a transient orientation of the ring plane during the flip. The circles represent atom groups near the ring and arrows indicate movements of atom groups during the ring flip. (From Wüthrich, 1986.)

for the C atoms 2, 3, 5 and 6 (Fig. 2), which undergo extensive movements during the ring flips. Theoretical studies performed in a collaboration of my group with Robert Huber's laboratory (Hetzel, Wüthrich, Deisenhofer & Huber, 1976) and independently by Gelin & Karplus (1975) then showed that the crystallographic *B* factors sample multiple-rotation states about the $C^{\alpha}—C^{\beta}$ bond, whereas the ring flips about the $C^{\beta}—C^{1}$ bond seen by NMR are very rapid 180° rotations connecting two indistinguishable equilibrium orientations of the ring. The *B* factors do not manifest these rotational motions because the populations of all non-equilibrium rotational states about the $C^{\beta}—C^{1}$ bond are vanishingly small. Indications of aromatic ring mobility were also observed in other proteins (Campbell, Dobson & Williams, 1975; Hull & Sykes, 1975). Eventually the ring-flip phenomenon turned out to be a general feature of globular proteins, manifesting ubiquitous low-frequency internal motions in protein molecules, which have activation energies of 60–100 kJ mol^{-1}, amplitudes of $\geq$1.0 Å and activation volumes of about 50 Å^3 (Wagner, 1980), and involve concerted displacement of numerous groups of atoms (Fig. 2). In summary, during the period 1967–1976 the NMR spectroscopists used X-ray crystal data as a source of information on three-dimensional protein structures, which were, on the one hand, needed to rationalize the NMR spectral data and could, on the other hand, be complemented with genuinely novel information on transient conformational states present within the framework of the average molecular structure. In hindsight it is also clear that in spite of numerous premature claims made at the time, the NMR approaches used during these years were not of the kind that could lead to *de novo* three-dimensional protein structure determination, not to speak of the fact that the instrumentation available then would not have been suitable for this purpose.

A third decade of subjecting protein solutions to magnetic fields was needed from 1977 to 1986, to achieve complete *de novo* NMR structure determinations of biological macromolecules. In a first phase until 1981, the key two-dimensional (2D) NMR experiments were adapted for use with macromolecules (Nagayama, Wüthrich, Bachmann & Ernst, 1977; Nagayama, Wüthrich & Ernst, 1979; Nagayama, Anil-Kumar, Wüthrich & Ernst, 1980; Anil-Kumar, Ernst & Wüthrich, 1980; Anil-Kumar, Wagner, Ernst &Wüthrich, 1981), and in parallel the method of sequential resonance assignments was developed (Billeter, Braun & Wüthrich,

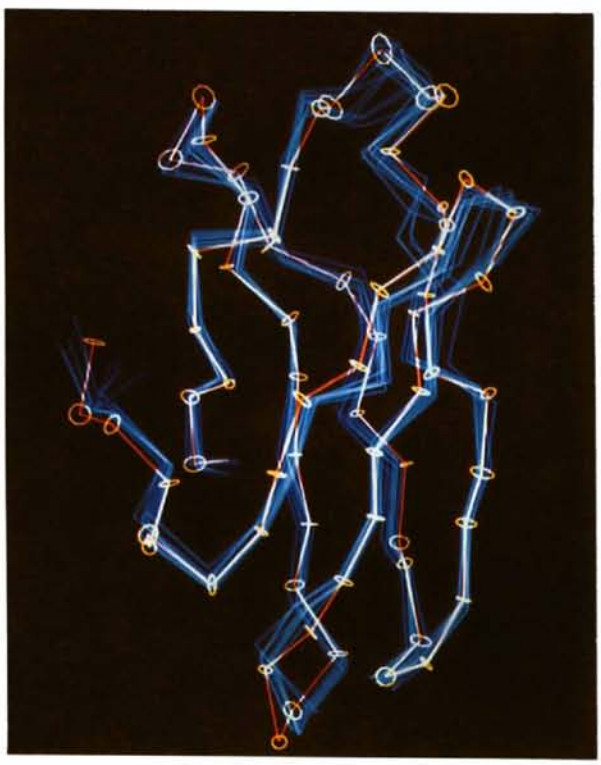

Fig. 3. Comparison of the structure of the protein Tendamistat derived by NMR in solution with that derived by X-ray crystallography. The NMR structure is represented by a bundle of nine conformers (blue), and the crystal structure by a red line (white in those regions where it overlaps the blue lines) representing the average atom coordinates and by yellow circles indicating the size of the *B* factors for the C^{α} atoms. Only the backbone is shown, that is, the blue and red lines represent virtual bonds connecting the C^{α} positions. (Drawing prepared with the atomic coordinates from Kline, Braun & Wüthrich, 1986; Pflugrath, Wiegand, Huber & Vértesy, 1986.)

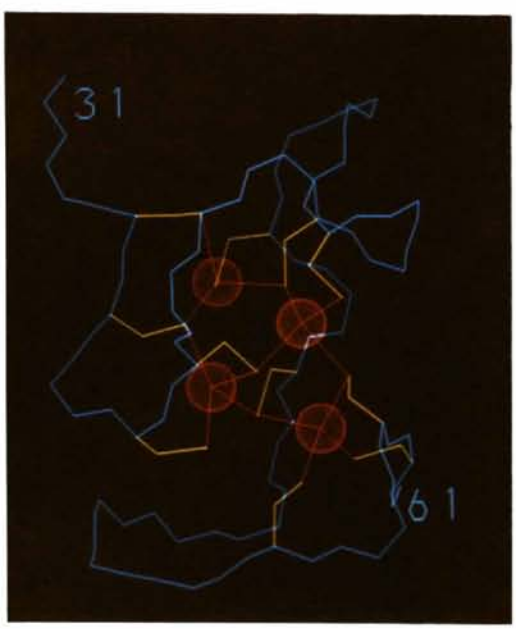

Fig. 4. Three-dimensional structure of the α-domain from human metallothionein-2 determined by NMR in solution. The blue line represents the polypeptide backbone from residues 31 to 61, the dotted spheres the four metal ions, yellow lines the 11 cysteinyl side chains, and red lines the coordinative bonds between the Cys S atoms and the metal ions. (Drawing prepared with the atom coordinates of Messerle, Schäffer, Vašák, Kägi & Wüthrich, 1990.)

1982; Dubs, Wagner & Wüthrich, 1979; Wagner & Wüthrich, 1982; Wider, Lee & Wüthrich, 1982). The aforementioned dynamic phenomena could now be attributed to unique locations along the polypeptide chain of BPTI, and with the assumption that the crystal structure was preserved in solution this also afforded a three-dimensional map of internal mobility. Obtaining sequence-specific resonance assignments also opened the door for three-dimensional structure determination by NMR (Wüthrich, Wider, Wagner & Braun, 1982) (see section 5 below). By 1984 the computational tools for protein structure calculation from NMR data had been developed (Braun, Bösch, Brown, Gō & Wüthrich, 1981; Havel & Wüthrich, 1984, 1985) and the structure determination of a globular protein, the bull seminal protease inhibitor IIA (BUSI IIA), was completed (Williamson, Havel & Wüthrich, 1985). When I presented the NMR structure of BUSI (Fig. 1) in some lectures in the spring of 1984, the reaction was one of disbelief and suggestions that our structure must have been modelled after the crystal structure of a homologous protein. Apparently the structural biology community had thoroughly adjusted to the role of NMR as a provider of some exotic supplementary data, although not being a suitable approach for *de novo* structure determination at atomic resolution. In the discussion following a seminar in Munich on May 14, 1984, Robert Huber proposed to settle the matter by independently solving a new protein structure in his laboratory by X-ray crystallography and in my laboratory by NMR. Each of us received 100 g (!) of the pure α-amylase inhibitor Tendamistat from Hoechst AG and, as is well documented in the literature (Billeter, Kline, Braun, Huber & Wüthrich, 1989; Braun, Epp, Wüthrich & Huber, 1989; Kline & Wüthrich, 1985; Kline, Braun & Wüthrich, 1986, 1988; Pflugrath, Wiegand, Huber & Vértesy, 1986), this project was started. Complete NMR assignments and the secondary-structure determination in solution were obtained by the end of the summer of 1984 (Kline & Wüthrich, 1985), and eventually virtually identical three-dimensional structures were obtained in solution (Kline, Braun & Wüthrich, 1986) and in crystals (Pflugrath, Wiegand, Huber & Vértesy, 1986) (Fig. 3). Some editorials then caused many to believe that Tendamistat rather than the earlier work by Tim Havel and Michael Williamson with BUSI (Williamson, Havel & Wüthrich, 1985) was the first complete *de novo* NMR structure determination of a globular protein. The fact that a genuinely novel type of polypeptide fold had been determined in the NMR structure of Tendamistat apparently convinced many of the critics that NMR could actually do the job. The Tendamistat experience was comforting when we came out with the next structure determination, rabbit metallothionein, in the spring of 1985 (Braun, Wagner, Wörgötter, Vašák, Kägi & Wüthrich, 1986). In response to the first public presentations of this structure at Yale University on June 10 and at the University of Pittsburgh on June 12, I was confronted with a crystal structure of rat metallothionein that was very different from our NMR structure (Furey, Robinson, Clancy, Winge, Wang & Stout, 1986). We subsequently found that the NMR structures of three homologous metallothioneins from rabbit, rat and man (Fig. 4) are virtually identical, and several years later the crystal structure of rat metallothionein was redetermined and found to actually coincide very closely with our NMR structure (Braun, Vašák, Robbins, Stout, Wagner, Kägi & Wüthrich, 1992). Clearly, establishing credibility for NMR structures of proteins did involve close encounters with the crystallography community!

Today, nearly 40 years after the initial recording of a protein ^{1}H NMR spectrum, it has become more customary than ever before to subject solutions of proteins and nucleic acids to strong magnetic fields. New advances during the fourth decade of this activity include the widespread use of ^{13}C and ^{15}N isotope labelling, the introduction of 3D and 4D NMR experiments for studies of isotope-labelled compounds (for a recent review see Bax & Grzesiek, 1993), and the development of heteronuclear filters for investigation of partially labelled macromolecules or multimolecular complexes (for a recent review see Otting & Wüthrich, 1990). The results of this evolution of the NMR methodology for biomolecular structure determination are in part also reflected in the numbers of new structures listed in Table 1.

3. NMR spectra of proteins and nucleic acids

The following isotopes in natural biological macromolecules give NMR signals of practical interest: ^{1}H, ^{13}C, ^{15}N and ^{31}P, and ^{2}H and ^{19}F may also be used after introduction of these isotopes by suitable chemical modification. The resonance positions of the different isotopes are widely different; they are not observed within the same NMR spectrum, and one actually needs different equipment for observation of the resonances from different nuclei. Thus, for example, the exchange reaction (1) can be followed by the disappearance from the ^{1}H NMR spectrum of the resonance lines corresponding to protons that were replaced with deuterons. Within the spectrum of a given isotope, individual nuclei in a complex molecular structure can generally be observed as separate lines. A one-dimensional (1D) ^{1}H NMR spectrum of the protein BPTI is shown in Fig. 5(*a*), where the resonance positions of different H atoms are given by the *chemical shift*, δ, in parts per million (p.p.m.) relative to a reference substance [usually a water-soluble derivative of $Si(CH_3)_4$ at 0 p.p.m.]. The chemical shift is determined primarily by the chemical structure. Thus, in a ^{1}H NMR spectrum of a protein (Fig. 5*a*) the resonance lines of all methyl groups appear on the extreme right at around 1 p.p.m., those of the potentially exchangeable amide protons [(1)] on the

extreme left from about 7 to 11 p.p.m., and in-between one observes methylene groups at 2–3 p.p.m., the α-protons at 4–5 p.p.m., and the protons of the aromatic rings at 6–7.5 p.p.m. The ^1H chemical shifts are further affected by the solvent milieu, so that a given chemical structure is observed in slightly different resonance positions when in aqueous solution or in a less polar environment. As to the resonance line widths, these may be as narrow as 0.1 Hz for small molecules, of the order 10 Hz in small proteins with molecular weights around 10 000, and broader in larger molecules. These few basic facts suffice to rationalize the salient features of ^1H NMR spectra of proteins and nucleic acids (Jardetzky & Roberts, 1981; Wüthrich, 1976, 1986).

In an extended, flexible polypeptide chain (Fig. 6*a*) all amino-acid side chains are exposed to the same solvent environment, so that multiple copies of a specified amino acid in the sequence have nearly identical ^1H chemical shifts. Therefore, the ^1H NMR lines of denatured 'random-coil' polypeptide chains correspond closely to the sum of the resonances in the constituent amino-acid residues, provided the increased macromolecular line widths are taken into account. In Fig. 5(*b*) a random-coil spectrum of BPTI was computed as the sum of the resonances of the amino-acid residues in the sequence of this protein. The increased complexity of the NMR spectrum of globular BPTI (Fig. 5*a*) results primarily from conformation-dependent ^1H chemical-shift dispersion, which is a consequence of a generalized solvent effect: interior peptide segments in globular proteins are shielded from the solvent and are surrounded by other peptide segments (Fig. 6*b*). Considering that the interior of a globular protein is highly aperiodic, each amino-acid residue is thus subjected to a unique microenvironment. In Fig. 5 these effects are shown in detail for the threonyl and tyrosyl residues in BPTI. For example, in the unfolded polypeptide chain the methyl groups of the three threonyl residues give rise to a single line corresponding in intensity to nine protons, whereas in the folded BPTI the chemical shifts of the three methyl groups are dispersed so that three separate lines of intensity three protons each, are observed. The NMR lines between 7.5 and 11 p.p.m. in the globular protein correspond to slowly exchanging amide protons. They have no counterpart in the unfolded protein because the spectrum of Fig. 5(*b*) was calculated with the assumption that these protons had been completely exchanged with ^2H [(1)].

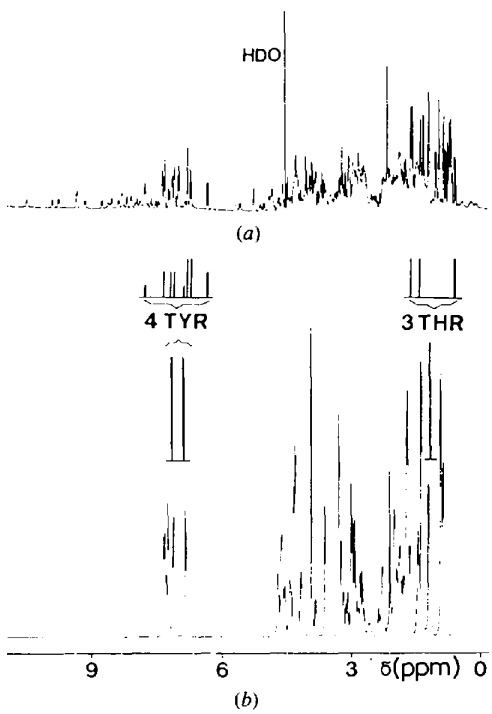

Fig. 5. (*a*) One-dimensional ^1H NMR spectrum of a freshly prepared ^2H_2O solution of the protein BPTI (protein concentration = 5 m*M*, p^2H = 4.5, T = 318 K; ^1H NMR frequency = 360 MHz). (*b*) Computed random-coil ^1H NMR spectrum of BPTI in ^2H_2O solution, where all the labile protons in N—H and O—H groups are replaced by ^2H [(1)]. All resonance lines in these spectra have been assigned to distinct H atoms of BPTI. Stick diagrams indicate the positions and intensities of the γCH_3 resonances of the threonyl residues in the sequence positions 11, 32 and 54, and the aromatic protons of the tyrosyl residues 10, 21, 23 and 35. In (*a*), HDO identifies the solvent water resonance. (From Wüthrich, 1986.)

At the origin of the ^1H chemical-shift dispersion in globular proteins are anisotropic diamagnetic susceptibilities of molecular fragments in the polypeptide chain. Dominant effects are due to the local *ring-current fields* near aromatic amino-acid residues (Fig. 7) and to the formation of intramolecular hydrogen bonds. *Ring-current shifts* arise because spins above or below the ring plane experience a local field that opposes the

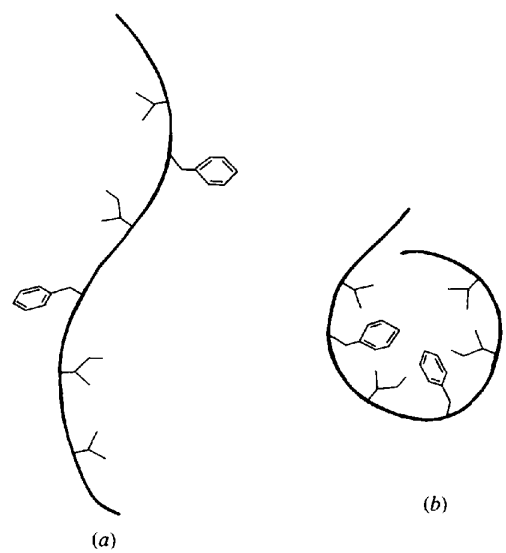

Fig. 6. Scheme depicting the surroundings of amino-acid side chains in an extended, 'random-coil' polypeptide chain (*a*) and in a globular protein (*b*). (From Wüthrich, 1986.)

external polarizing magnetic field $\mathbf{B}_o$ (Fig. 7), and therefore the corresponding resonance lines are displaced to higher field (*i.e.* to the right in the spectrum). Conversely, spins located in the ring plane outside of the confines of the ring are shifted to lower field. In BPTI the conformation-dependent shifts of the three threonyl γCH_3 lines at −0.64, 0.16 and 0.38 p.p.m. (Fig. 5) can be explained nearly quantitatively by the ring-current effects predicted from the crystal structure, and the same holds for methyl groups in general. Similar but usually much smaller anisotropic local magnetic fields are induced by other atom groups, for example, carbonyl groups. Very large shifts may also result near prosthetic groups, in particular the heme groups in hemoproteins (Wüthrich, 1970).

Fig. 5 clearly illustrates the importance of conformation-dependent chemical-shift dispersion for obtaining the spectral resolution needed for structural studies by NMR. In the unfolded polypeptide chain, multiple copies of the same amino-acid type could not be distinguished (Fig. 5*b*), whereas in the folded protein each residue can be individually observed (Fig. 5*a*). These observations dictate an essential requirement for NMR structure determinations: NMR techniques will be required that enable work with the complex spectrum of the intact folded protein (Fig. 5*a*), since the desired data cannot be obtained by adding up information from studies of small, constituent oligopeptide fragments.

Some pertinent conclusions on the use of ^{1}H chemical shifts for structural studies can also be drawn from these qualitative considerations on conformation-dependent chemical-shift dispersion in globular proteins: (i) when concentrating on resonance lines with sizeable ring-current shifts, in particular side-chain methyl groups, one can check on the compatibility of the observed ^{1}H shifts with the crystal structure of the protein. (ii) Ring-current-shifted lines can be used as highly sensitive indicators of conformation changes. (iii) Because of the high symmetry of the local magnetic fields, conformation-dependent ^{1}H shifts are not a useful parameter for *de novo* three-dimensional structure determination. This is readily apparent from Fig. 7, since atoms located in corresponding positions above or below the plane of the aromatic ring experience identical ring-current shifts, and the ring-current field is further circularly symmetrical with respect to the z axis. (iv) The ^{1}H NMR spectrum cannot be predicted from the crystal structure of the protein, both because the effects of weaker local fields on the chemical shifts cannot be quantitatively accounted for, and in regions near aromatic rings the crystallographic atom coordinates are not sufficiently precise to enable quantitative shift calculations.

4. NMR studies of dynamic processes on the chemical-shift time scale

The aforementioned ring-flip phenomenon is manifested in the ^{1}H NMR spectra because chemical-shift dispersion in the spatial protein structure can also arise for protons within the same residue. This is schematically illustrated in Fig. 8 for a phenylalanine ring. In the covalent structure the two proton pairs 2,6 and 3,5 (Fig. 2) are in symmetry-related equivalent positions. In an isotropic solvent this chemical-shift equivalence is preserved. A 'symmetrical' three-line spectrum is then observed, with identical shifts for positions 2 and 6, and 3 and 5, respectively. In the interior of a protein the twofold symmetry of the covalent structure is masked by different non-bonding shielding effects on the individual ring protons, so that an 'asymmetric' five-line spectrum is obtained. These effects from the protein environment can be averaged out by rotational motions of the aromatic ring. With increasing frequency of the ring flips, which

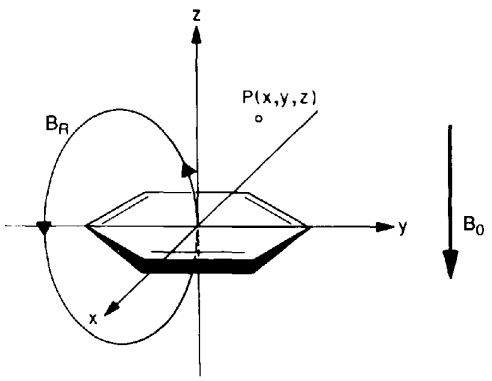

Fig. 7. Local ring-current field, $\mathbf{B}_R$, of a phenylalanine ring. $\mathbf{B}_R$ is induced by the external polarizing field $\mathbf{B}_o$. For simplicity only a single field line through the ring centre to the left along the y axis is drawn, but $\mathbf{B}_R$ is of course circularly symmetrical relative to the Z axis. In a three-dimensional molecular structure, protons located at the position $P(x, y, z)$ relative to the ring centre experience $\mathbf{B}_R$ in addition to the external field $\mathbf{B}_o$, and the resulting 'ring-current shifts' are dependent on the relative coordinates of the ring and of $P(x, y, z)$.

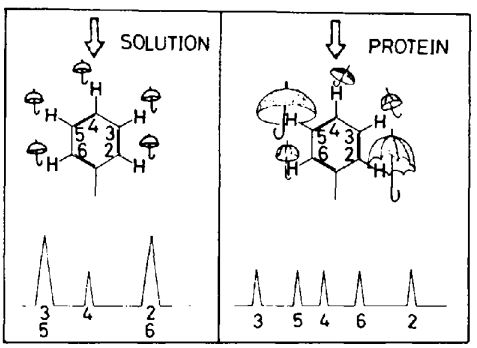

Fig. 8. 'Solvent effects' on the ^{1}H NMR spectrum of a phenylalanine ring in solution and in a globular protein. The size of the umbrellas reflects the shielding of the individual aromatic H atoms from the external polarizing magnetic field $\mathbf{B}_o$ (arrow). The resonance assignments at the bottom refer to the atom numbering indicated in the rings (From Wüthrich, 1986.)

can be achieved by raising the temperature within the stability range of the globular protein conformation (Fig. 9), the asymmetric two-line spectra of the 2,6- or 3,5-protons first broaden, then merge into a very broad two-proton line, and eventually give rise to a sharp two-proton resonance. Thereby the extent of the conformation-dependent dispersion of the chemical shifts determines the NMR time scale for observation of these exchange processes, *i.e.* the larger the chemical-shift difference between the two lines representing the asymmetric state, the higher the ring-flip frequencies that are accessible for quantitative measurement. These correlations are readily apparent from the spectral simulations in Fig. 9, which also show that the frequencies covered by these experiments extend from about $1\,s^{-1}$ to $2 \times 10^4\,s^{-1}$. Since chemical-shift differences are proportional to the applied magnetic field $\mathbf{B}_o$, the accessible frequency range is also dependent on the spectrometer frequency. Thus, higher ring-flipping frequencies could be measured at a ^1H resonance frequency of 600 MHz than in the 360 MHz spectrum of Fig. 9.

A similar situation to the one where two chemically equivalent ring protons exchange positions between two different environments (size of the umbrellas in Fig. 8) prevails when a single group of protons undergoes exchange between two different environments that are characterized by different chemical shifts. Examples of dynamic processes in proteins that would thus be accessible to NMR observation are motions of amino-acid side chains in and out of the local ring-current field of an aromatic side chain (Fig. 7), or exchange of an amide proton between two or multiple hydrogen-bonding partners.

5. The nuclear Overhauser effect (NOE) and the pivotal role of sequence-specific NMR assignments

Nuclear Overhauser effects are due to dipolar (through-space) interactions between different nuclei and are correlated with the inverse sixth power of the internuclear distance. The possibility of measuring interatomic distances between pairs of H atoms (or, in practice, of establishing upper boundaries on $^1H—^1H$ distances over the range of about 2.0–5.0 Å) *via* observation of NOE's represents the physical basis for macromolecular structure determination by NMR. NOE's can be observed in double-irradiation 1D NMR experiments as the fractional change in intensity of one NMR line when another resonance is irradiated, or as *cross peaks* in 2D NOE spectroscopy (2D [^1H,^1H]-NOESY) Fig. 10) or higher dimensional spectra (see below). The physical processes leading to NOE's are quite slow so that the *NOE build-up* can be followed over a time range of milliseconds to seconds. $^1H—^1H$ distance measurements in macromolecules are based on the analysis of NOE build-up curves recorded typically over a period of up to about 300 ms, which is needed to prevent possible pitfalls due to 'spin diffusion' (Gordon & Wüthrich, 1978; Wagner & Wüthrich, 1979; Anil-Kumar, Wagner, Ernst & Wüthrich, 1981).

The 2D [^1H, ^1H]-NOESY spectrum of Fig. 10 contains resonance lines on the diagonal from the upper right to the lower left which correspond to the 1D ^1H NMR

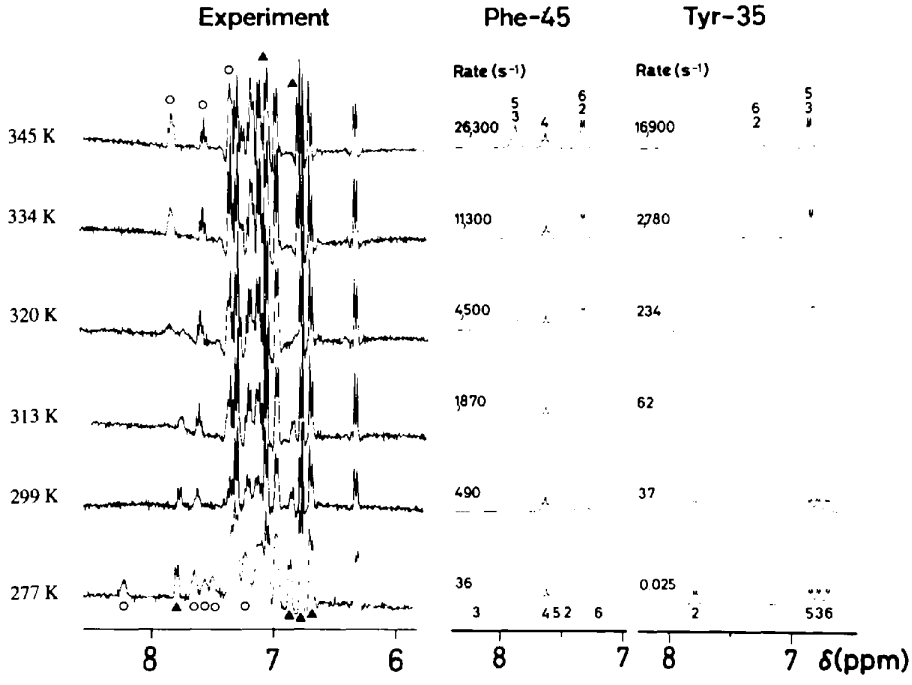

Fig. 9. NMR observation of aromatic ring flips. The experiment consists of a series of one-dimensional ^1H NMR spectra of BPTI in ^2H_2O (protein concentration = 10 m*M*, p^{2}H = 7.8, ^1H NMR frequency = 360 MHz) recorded at the different temperatures indicated on the left of each trace. The spectral region from 6.0 to 8.5 p.p.m. is shown. The resonances of the aromatic rings of Phe45 (○) and Tyr35 (▲) are identified in the spectra at 277 and 345 K and were simulated in the centre and on the right, respectively, using the ring-flip frequencies indicated with the individual simulations. The resonance assignments in the simulated spectra at 277 and 345 K refer to the atom numeration in Fig. 2. (From Wüthrich, 1986.)

spectrum. The cross peaks in the plane outside of the diagonal manifest NOE's between pairs of diagonal peaks, as indicated for one cross peak by a horizontal and a vertical line in the lower left of the figure. In a qualitative approach which is sufficient for obtaining a low-quality three-dimensional protein structure (Havel & Wüthrich, 1985), each NOESY cross peak of observable intensity (Fig. 10) can be attributed the same upper boundary on the distance between a pair of H atoms, i and j, of, for example, $U_{ij} < 4.0$ Å. A large number of NOE's are typically observed in globular proteins (Fig. 10) or in double-helical nucleic acids, indicating that these molecular structures contain numerous pairs of closely spaced H atoms. The structural information thus obtained from a NOESY spectrum is schematically represented by the upper part of Fig. 11. Clearly, on this level of detail the NOE distance constraints are not a viable basis for a three-dimensional structure determination, although they may be the result of a beautiful, well resolved NOESY experiment. The situation is analogous to that of a crystallographer who has obtained beautiful X-ray diffractions of a native protein crystal.

The crucial link between the NOE distance constraints and the molecular structure is provided by *sequence-specific resonance assignments* (Wüthrich, Wider, Wagner & Braun, 1982). Based on the assignments the distance constraints can be attributed to specific sites along the polymer chain, and hence the NOE experiments define the formation of loops by chain segments of variable lengths (lower part of Fig. 11). Considering that the overall length of the extended polypeptide chain of a protein with 100 residues is about 350 Å, the NOE's with $U_{ij} < 4.0$ Å impose stringent constraints on the polypeptide conformation. With the use of suitable mathematical procedures, notably distance geometry, the conformation space can then be searched for spatial arrangements of the polymer chain that are compatible with the uniquely assigned experimental distance constraints. Alhough the problems to be solved in using the two methods are technically very different, obtaining sequence-specific resonance assignments in protein structure determinations by NMR thus has a similar impact to that of solving the phase problem in protein crystallography.

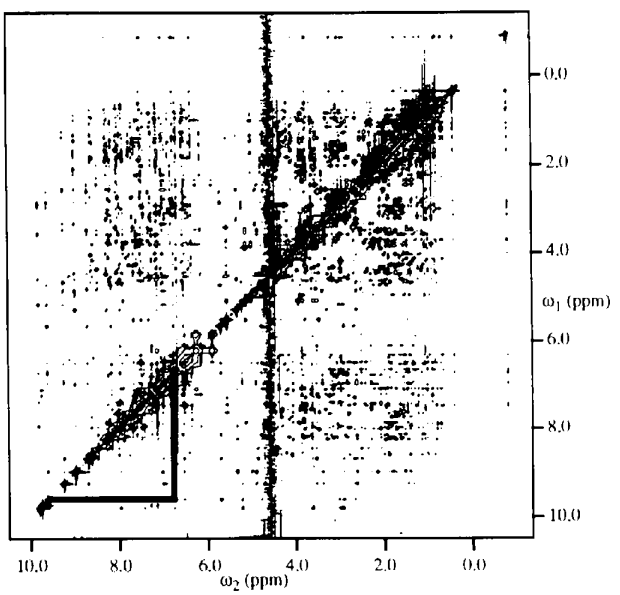

Fig. 10. Two-dimensional [^{1}H,^{1}H]-NOESY (nuclear Overhauser enhancement spectroscopy) of the protein toxin K (protein concentration = 10 mM, solvent = 90% H_2O/10% ^2H_2O, pH = 4.6, T = 309 K, ^{1}H frequency = 600 MHz, mixing time = 40 ms). The horizontal line and the vertical line through a cross peak in the lower left indicate how the cross peaks correlate pairs of H atoms (represented by the chemical shift on the diagonal) which are related by an NOE.

6. The practice of macromolecular structure determination by NMR

Fig. 12 presents an overview of the course of a NMR structure determination. Individual steps are the sample preparation, the NMR measurements, the assignments of the NMR lines to individual atoms in the polymer chain, the collection of conformational constraints, the structure calculation and structure refinements. Although the entire structure determination might conceivably at some future stage be handled as a single computer-supported process, in present practice the sequence of steps does correspond largely to the flow diagram of Fig. 12. As is also indicated in Fig. 12, it is a special feature of protein structure determination by NMR that the secondary polypeptide structure, including the connections between individual segments of regular secondary structure, may be known early-on from the data used for obtaining the resonance assignments, actually before the structure calculation is even started.

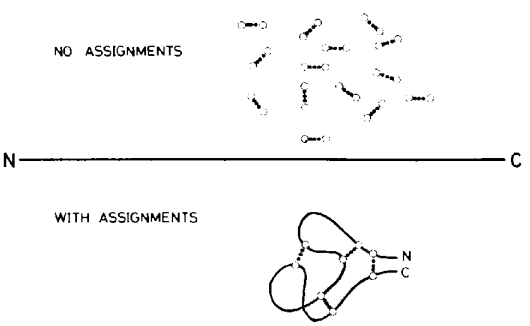

Fig. 11. Information content of ^{1}H–^{1}H NOE's in a polypeptide chain without (upper part) and with (lower part) sequence-specific resonance assignments. Open circles represent H atoms of the polypeptide. Dotted lines between pairs of circles indicate NOE distance constraints. The polypeptide chain is represented by the horizontal line in the centre, and the solid line in the lower drawing. (From Wüthrich, 1986.)

6.1. *Sample preparation*

As even minor macromolecular impurities may be detrimental to the outcome of a structure determination, the samples used must be highly homogeneous. The material is dissolved in 0.5 ml of water, and the ionic strength, pH and temperature may be adjusted so as to ensure near-physiological conditions (for proteins it is advantageous to work in the slightly acidic pH range from 3 to 5). The concentration should be about 1 m*M*, ideally 2–5 m*M*, so that, for example, 5–25 mg of a protein with molecular weight 10 000 should be available for a structure determination. Although this concentration is high relative to that of most proteins in their physiological milieu, it is not far from the total protein concentration in typical body fluids. For molecular sizes above approximately 10 000, quite often also for smaller sizes, the NMR study will have to include the preparation of compounds enriched with ^{15}N and/or ^{13}C. Uniformly isotope-labelled recombinant proteins are obtained with commonly available techniques; nonetheless, preparation of a suitable protein sample is often the bottleneck in a project. For RNA and DNA, selective or uniform labelling has been introduced more recently; it is still used only by a few groups and labelling techniques are still in the process of development.

I could imagine that most readers of this journal would readily agree that being able to work with solutions is a great asset of the NMR method. However, working with solutions also has its inherent potential difficulties. Most of all, many biological macromolecules tend to self-aggregate at the concentrations needed for NMR spectroscopy, which has slowed down many projects. Furthermore, in the course of an investigation it may be non-trivial to achieve identical solution conditions in different NMR samples of the same compound, which typically results in small chemical-shift differences that slow down the combined analysis of different NMR spectra.

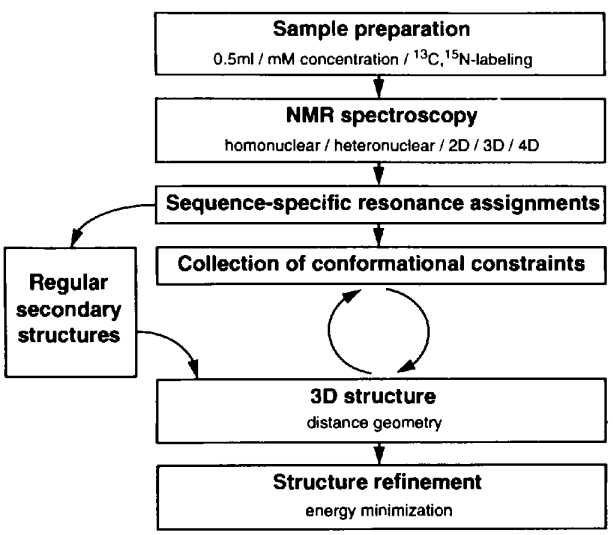

Fig. 12. Diagram outlining the course of a macromolecular structure determination by NMR in solution (see text).

6.2. *NMR spectroscopy with biological macromolecules*

The selection of NMR experiments for macromolecular structure determination is dictated mainly by two requirements: (i) hundreds or even thousands of resonance lines in the ^1H NMR spectrum of a biological macromolecule must be observed as separate peaks, and selective correlations between pairs of spins (or groups of equivalent spins, for example, the three protons in a methyl group) need to be established. (ii) The maximum possible number of ^1H–^1H NOE's must be identified, since the information needed for three-dimensional structure determination comes primarily from NOE distance constraints. In present practice these demands are met by multidimensional NMR (Ernst, Bodenhausen & Wokaun, 1987; Wüthrich, 1986).

Figs. 5*a* and 10 may serve to illustrate the problems to be solved. In a 1D ^1H NMR spectrum of a protein (Fig. 5*a*) only a small percentage of the resonance lines are well separated, and there is extensive overlap of multiple resonances in most spectral regions. In the 2D ^1H NMR spectrum the corresponding peaks on the diagonal from the upper right to the lower left are similarly overlapped, but the additional cross peaks in the 2D frequency plane outside of the diagonal, which manifest correlations between pairs of diagonal peaks, are much better resolved than the resonance lines in the 1D NMR spectrum. The 2D ^1H NMR experiment can thus yield selective ^1H–^1H correlations also in crowded spectral regions. Depending on the type of experiment performed, different kinds of spin–spin correlations are manifested by the cross peaks (see below).

With increasing size of the molecule studied and concomitant increase of the number of NMR peaks, it becomes more and more difficult to resolve and assign individual resonances in homonuclear 2D ^1H NMR spectra. In heteronuclear 3D or 4D spectra recorded with recombinant proteins that are uniformly labelled with ^{15}N and/or ^{13}C, the peaks are spread out in a third and possibly fourth dimension along the ^{15}N and ^{13}C chemical-shift axes (for a recent review, see Bax & Grzesiek, 1993). As an illustration, Fig. 13 shows a 3D ^{15}N-correlated [^1H,^1H]-NOESY spectrum, where the NMR peaks are spread out along a third axis which corresponds to the NMR frequencies of the ^{15}N spins in the ^{15}N-labelled protein. As a result, the same number of NMR peaks that would be observed in a 2D [^1H,^1H]-NOESY spectrum (Fig. 10) are distributed among multiple ^1H–^1H planes, typically 64 or 128 (Fig. 14). The ensuing further improved separation of the peaks in the individual planes is indispensable for work with larger proteins, and the preparation of isotope-labelled proteins as well as the use of higher-dimensional heteronuclear NMR experiments are now general

practice. In addition to the ^{1}H NMR data, this approach provides ^{13}C and ^{15}N NMR information, which may be used to support the structure determination (*e.g.* Wishart, Sykes & Richards, 1991) and to provide supplementary information on dynamic features of the molecule studied. Nonetheless, the key purpose of the heteronuclear NMR experiments is to obtain the maximum possible number of individually assigned ^{1}H–^{1}H NOE's in the system studied.

6.3. *Resonance assignments*

Obtaining resonance assignments in biopolymers is non-trivial for reasons that may best be rationalized by comparison with NMR applications in other fields. For example, in studies of organic chemical compounds, resonance assignments are typically based on the identification of unique *spin systems*, which are characteristic for the covalent molecular structure. In homonuclear ^{1}H NMR, each spin system includes all those H atoms that are related by scalar spin–spin couplings (*e.g.* Wüthrich, 1986). In macromolecules this encompasses those groups of H atoms that are separated in the chemical structure by steps of, at most, three covalent bonds. For most amino-acid residues the ensemble of all H atoms thus forms one spin system. For example, as is indicated by the dotted lines in Fig. 15, the three equivalent methyl protons of alanine are separated by three bonds from the α-proton, which in turn relates to the amide proton *via* three bonds, so that a continuous network of scalar couplings can be established. Exceptions are the four aromatic amino-acid side chains and some peripheral labile side-chain protons, where the connection to the aliphatic H atoms of the same side chain must be established *via* ^{1}H—^{1}H NOE's (Wüthrich, 1986). Fig. 15 also shows that alanine and valine can readily be distinguished, since their ^{1}H spin systems consist of three lines, and five lines, respectively. In a typical NMR application in organic chemistry, the assignment problem would be solved at this stage. In contrast, since proteins invariably contain more than one residue of most of the 20 common amino acids, the identification of the groups of scalar-coupled spins belonging to individual amino-acid residues is in general not sufficient to define a unique sequence location.

The aforementioned situation is illustrated in Fig. 16 with the *lac* repressor DNA-binding domain, which is one of the early proteins that were assigned in my laboratory in Zürich (Zuiderweg, Kaptein & Wüthrich, 1983*a*). This polypeptide contains eight valyl residues in the positions 4, 9, 15, 20, 23, 24, 30 and 38. Although the eight valyl spin systems could be identified, there was no way to assign individual sequence positions from this information alone. This intrinsic problem was solved with the sequential assignment strategy (Dubs, Wagner & Wüthrich, 1979; Billeter, Braun & Wüthrich, 1982; Wagner & Wüthrich, 1982): Unique segments of two or several sequentially adjoining amino-acid residues are identified by NMR experiments, and then by comparison with the chemically determined amino-acid sequence

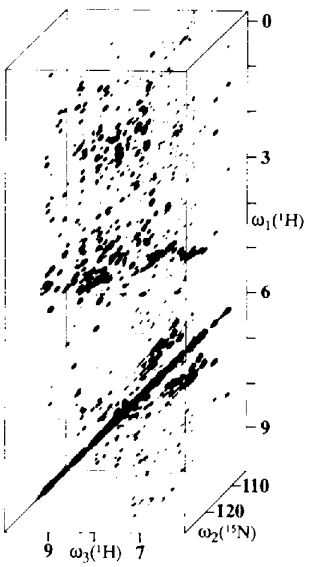

Fig. 13. 3D ^{15}N-correlated [^{1}H,^{1}H]-NOESY spectrum of oxidized *E. coli* glutaredoxin (protein concentration = 1.5 m*M*, solvent = 90% H_2O/10% $^{2}H_2O$, pH = 7.0, T = 301 K, ^{1}H frequency = 500 mHz, mixing time = 100 ms). The protein was uniformly labelled with ^{15}N to the extent of $\geq$ 95%. (From Sodano, Chary, Björnberg, Holmgren, Kren, Fuchs & Wüthrich, 1991.)

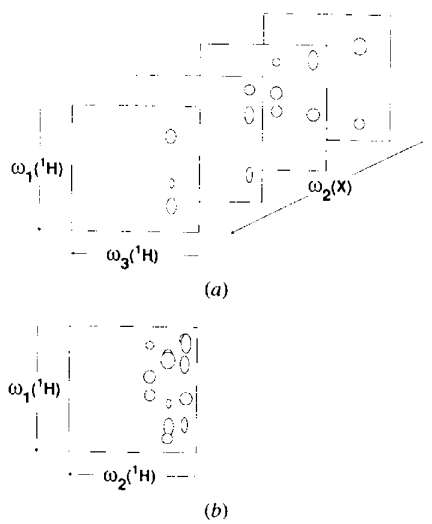

Fig. 14. Scheme illustrating the improved peak separation in a 3D heteronuclear-resolved [^{1}H,^{1}H]-NMR experiment (*a*) when compared with the corresponding 2D [^{1}H,^{1}H]-NMR experiment (*b*). The two spectra contain the same number of peaks. In the 3D spectrum (Fig. 13) these are distributed among multiple $\omega_1(^1H)$–$\omega_3(^1H)$ planes that are separated along the heteronuclear chemical-shift axis, $\omega_2(X)$. The 2D spectrum corresponds to a projection of all the peaks in (*a*) onto a single plane, causing overlap among some of the peaks [$\omega_2(^1H)$ in the 2D spectrum corresponds to $\omega_3(^1H)$ in the 3D spectrum]. (From Otting & Wüthrich, 1990.)

attributed to discrete positions in the polypeptide chain. The desired relations between protons in sequentially neighbouring amino-acid residues i and $i+1$ can be established by NOE's manifesting close approach between αCH_i and NH_{i+1} ($d_{\alpha N}$), NH_i and NH_{i+1} (d_{NN}) (Fig. 15), and possibly βCH_i and NH_{i+1} ($d_{\beta N}$) (Wüthrich, 1986). For *Xxx*-Pro dipeptide segments, corresponding connectivities are observed to δCH_2 of Pro in the place of the amide proton. In Fig. 15, for example, observation of a NOESY cross peak corresponding to $d_{\alpha N}$ or d_{NN} tells us that the protein studied contains a dipeptide segment Val-Ala. This dipeptide is then matched against the independently known amino-acid sequence. If the latter contains Val-Ala only once, the assignment problem is solved. In the *lac* repressor DNA-binding domain, the Val-Ala resonances could thus unambiguously be assigned to the residues in positions 9 and 10 (Fig. 16). Otherwise, to distinguish between multiple Val-Ala sites, tri- or tetrapeptide segments including Val-Ala must be identified by NMR and matched against the amino-acid sequence. Since in globular proteins with up to 200 amino-acid residues the probability is small that identical tri- or tetrapeptide segments occur more than once, the procedure is generally applicable (Wüthrich, 1986). For small proteins with up to about 100 amino-acid residues, NOE-based sequential assignments can rely entirely on homonuclear ^1H NMR experiments, and for somewhat bigger proteins with up to about 200 residues they can be established based on the improved resolution of 3D NMR experiments with ^{15}N-labelled proteins (Fig. 13). No prior knowledge of the polypeptide conformation is needed, since at least one of the two distances $d_{\alpha N}$ or d_{NN} (Fig. 15) is always sufficiently short to be observed by NOE's (Billeter, Braun & Wüthrich, 1982). A further attractive feature of this approach is that the identification of the sequential NOE's forms an integral part of the data collection for the protein structure determination (see below).

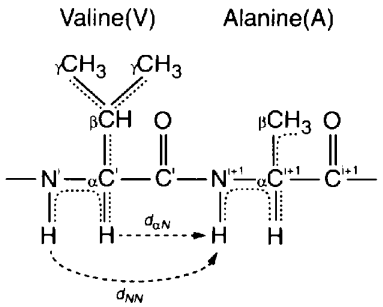

Fig. 15. Sequential resonance assignment based on sequential ^1H–^1H NOE's. In the dipeptide segment -Val-Ala- the dotted lines indicate ^1H–^1H relations which can be established by scalar through-bond spin–spin couplings. The broken arrows indicate relations between protons in sequentially neighbouring residues, which can be established by NOESY cross peaks manifesting short sequential distances $d_{\alpha N}$ (between αCH and the amide proton of the following residue) and d_{NN} (between the amide protons of neighbouring residues).

Sequential assignments can alternatively be obtained entirely *via* heteronuclear scalar couplings. Although this intuitively straightforward approach was applied already in 1979 to obtain assignments for the polypeptide toxin apamin, using the unlabelled peptide and multiple-irradiation 1D NMR experiments (Okhanov, Afanas'ev & Bystrov, 1980), it became a practical technique only a decade later with the availability of recombinant isotope-labelled proteins. For a polypeptide chain that is uniformly labelled with ^{15}N and ^{13}C, the entire macromolecular structure can be represented as a single heteronuclear spin system, since there are observably large scalar couplings also between atoms in neighbouring amino-acid residues (Fig. 17). Using 3D and 4D heteronuclear triple-resonance experiments (Montelione & Wagner, 1989, 1990; Ikura, Kay & Bax, 1990; Bax & Grzesiek, 1993; Szyperski, Wider, Bushweller & Wüthrich, 1993), the resonance lines of sufficiently large fragments of the polypeptide structure are grouped together to enable sequence-specific resonance assignments (*e.g.* Clore & Gronenborn, 1991). The largest proteins for which the polypeptide backbone has been assigned using $^{13}C/^{15}N$-double-labelling and triple-resonance experiments to correlate groups of atoms *via* the spin–spin couplings of Fig. 17 are in the range 250–300 residues.

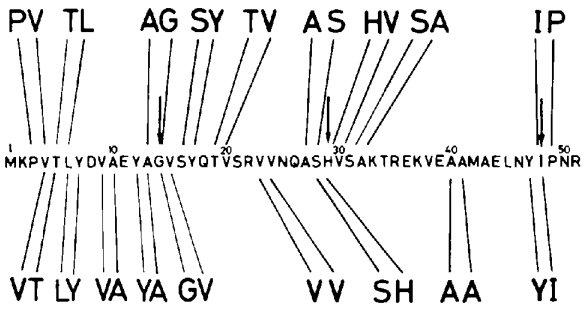

Fig. 16. Amino-acid sequence of the DNA-binding domain of residues 1–51 of the *E. coli lac* repressor with identification of unique residues (arrows), and indication of selected unique dipeptide segments that are useful for obtaining sequence-specific resonance assignments. (From Wüthrich, 1986.)

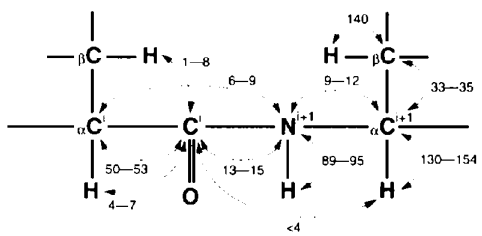

Fig. 17. Segment of a polypeptide chain with indication (by double-headed arrows, the numbers are the spin–spin coupling constants in Hz) of the scalar spin–spin couplings that provide the basis for obtaining sequential assignments by triple-resonance experiments with uniformly $^{13}C/^{15}N$ doubly labelled proteins. i and $i+1$ identify backbone atoms of two sequentially neighbouring residues.

For nucleic acids, assignment procedures were largely patterned after those for proteins. In contrast to the situation with polypeptide chains, assignments using intranucleotide, sequential and interstrand ^{1}H–^{1}H NOE's are necessarily based on the assumption of a particular conformation type for the compound studied (Wüthrich, 1986). For B-type DNA, this approach is highly efficient for duplexes with up to about 20 base pairs. A conformation-independent approach is also available, which uses heteronuclear scalar ^{1}H–^{31}P couplings to establish sequential assignments. In unlabelled compounds these assignments cannot readily be extended to include the resonances of the bases, which limits their usefulness (Wüthrich, 1986). With suitably ^{15}N- and ^{13}C-labelled compounds, base–backbone connections can be established *via* heteronuclear scalar couplings, which may in the future enable complete assignments for oligonucleotide segments with non-regular conformations.

6.4. *Collection of conformational constraints and structure calculation from NMR data*

As was explained in section 5 above, NOE upper distance constraints contain the crucial information needed for macromolecular structure determination (Fig. 11), although supplementary constraints may be derived from other NMR measurements (Wüthrich, 1986; 1989*a*). For a high-quality structure determination, the maximum possible number of NOE conformational constraints must be collected as input for the calculation of the complete three-dimensional protein structure. This is accomplished by using the chemical-shift lists obtained as a result of the sequence-specific resonance assignments to attribute the cross peaks in 2D (Fig. 10), 3D (Fig. 13) or 4D |^{1}H,^{1}H]-NOESY spectra to distinct pairs of H atoms. It is clear from section 3 that one has to anticipate some accidental chemical-shift degeneracies even in the NMR spectra of folded proteins. Therefore, as indicated in Fig. 12, the data collection is achieved in several cycles, where ambiguities in the NOESY cross-peak assignments can usually be resolved by reference to preliminary structures calculated from incomplete input data sets (Güntert, Berndt & Wüthrich, 1993).

In present practice the NOE input data have the format of allowed distance ranges, which circumvents intrinsic difficulties that might arise from attempts at quantitative distance measurements. The lower limit is usually taken to correspond to the sum of two H-atom radii, *i.e.* 2.0 Å. For short-range connectivities between protons separated by not more than three single bonds in the chemical structure, that is, intraresidual and sequential distances (Wüthrich, 1986), the NOE intensities are translated into corresponding upper boundaries, typically in steps of 2.5, 3.0 and 4.0 Å. For longer-range connectivities, a predetermined upper limit, usually 4.0 or 5.0 Å depending on the protein and the experimental conditions used, is applied. Unless stereospecific, individual assignments for pairs of diastereotopic substituents were obtained, in particular for the β-methylene groups in amino-acid side chains and the isopropyl groups of valine and leucine, correction factors for the use of *pseudoatoms* must be added to the observed upper limits on distances to H atoms of prochiral centres (Wüthrich, Billeter & Braun, 1983). The resulting allowed distance ranges may then be as large as from 2.0 to about 10 Å. Supplementary conformational constraints, for example, from spin–spin coupling constants, are represented in the input by similar allowed ranges, which account for the limited accuracy of the measurements, and implicitly for uncertainties in the definition of the measured parameters because of internal mobility. In computer-supported collection of the input data, the stepwise classification is preferably replaced by a smoothed fit of upper distance boundaries *versus* NOE intensity, and the calibration may be updated in successive cycles of the structure calculation.

Distance geometry algorithms can be used to obtain the Cartesian coordinates of spatial molecular structures that are consistent with a predetermined set of intramolecular interatomic distances (Crippen & Havel, 1988). To this end, such algorithms identify the conformations

$$\{r_i = (x_i, y_i, z_i);\ i = 1, 2, \ldots N\}, \tag{2}$$

which are consistent with the inequality

$$L_{ij} \leq |r_i - r_j| \leq U_{ij}, \tag{3}$$

where L_{ij} and U_{ij} are lower and upper boundaries, respectively, on the Euclidian distance $|r_i - r_j|$ between the points r_i and r_j, and N is the number of points (or atoms) in the system studied. The problem described by the inequality (3) may alternatively be expressed in the form of an error function, which can then be minimized. In the process of this conversion of distance information into Cartesian coordinates, one inevitably encounters the problem of local minima (Némethy & Scheraga, 1977). Metric matrix methods (Havel & Wüthrich, 1984) and variable target functions in dihedral angle space (Braun & Gō, 1985; Güntert, Braun & Wüthrich, 1991) have been implemented for the search for protein conformations obeying inequality (3), and the energy functions in molecular dynamics programs have been supplemented with terms representing the constraints expressed by inequality (3) (Brünger, Clore, Gronenborn & Karplus, 1986; Kaptein, Zuiderweg, Scheek, Boelens & van Gunsteren, 1985). Overall, corresponding results are obtained with these different computational approaches, although details may be subject to further analysis.

Model calculations performed in conjunction with the first protein structure determination (Fig. 1) had shown that the quality of an NMR structure determination depends critically on the density of NOE distance constraints, while the procedure is remarkably robust

with regard to low precision of the individual distance constraints (Havel & Wüthrich, 1985). As mentioned earlier, multiple rounds of data collection and structure calculation (Fig. 12) are usually needed to obtain the desired high-density input.

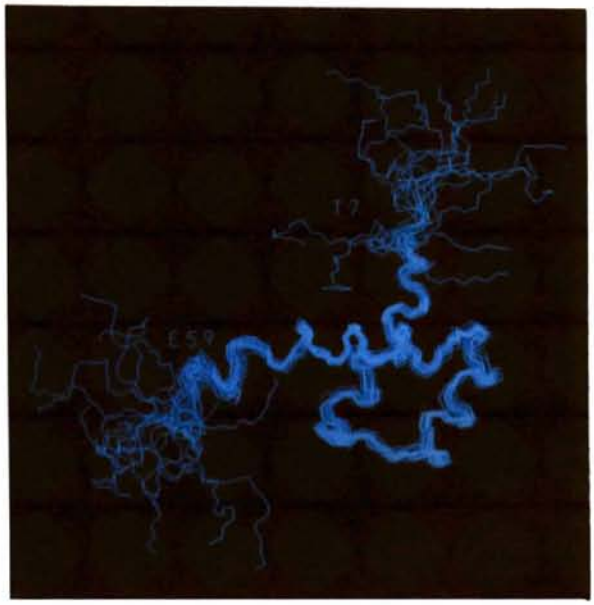

Fig. 18. Polypeptide backbone of 19 energy-refined conformers selected to represent the NMR solution structure of the *Antennapedia* homeodomain. Conformers 2–19 were superimposed for pairwise minimum r.m.s.d. of the backbone atoms N, C^{α} and C′ of residues 7–59 with conformer 1. From residues 7–59 (identified in the figure) the structure is well defined by the NMR data, but the chain-terminal segments 0–6 and 60–67 are (flexibly) disordered. (Drawing prepared with the atomic coordinates from Qian, Billeter, Otting, Müller, Gehring & Wüthrich, 1989.)

The result of a single distance geometry calculation (Fig. 1) represents one molecular geometry that is compatible with the NMR data, but one cannot tell whether this solution is unique. Thus, the calculation is repeated with different initial conditions. For each calculation, convergence is judged by the residual constraint violations, and all satisfactory solutions are included in a group of conformers used to represent the NMR structure of the protein. The final result of a structure detrmination is then commonly presented as a superposition of a group of conformers for pairwise minimum root-mean-square deviation (r.m.s.d.) relative to a predetermined conformer, or relative to the mean atom coordinates. Fig. 18 shows a superposition of the polypeptide backbone in 19 conformers of the *Antennapedia* homeodomain that were selected to represent the solution structure (Qian, Billeter, Otting, Müller, Gehring & Wüthrich, 1989). The precision of the structure determination, as reflected by the dispersion among this group of conformers, varies along the polypeptide chain. From residues 7 to 59, all conformers are similar, have nearly the same overall dimensions and contain the same secondary structure, showing that this part of the structure is well defined by the NMR data. In contrast, both chain ends are (dynamically) disordered, as is manifested by the large dispersion among the 19 conformers. Large variations are typically observed also in the precision of the structure determination of the amino-acid side chains, with well defined interior side chains representing the protein core and flexibly disordered side chains located predominantly near the protein surface (Fig. 19).

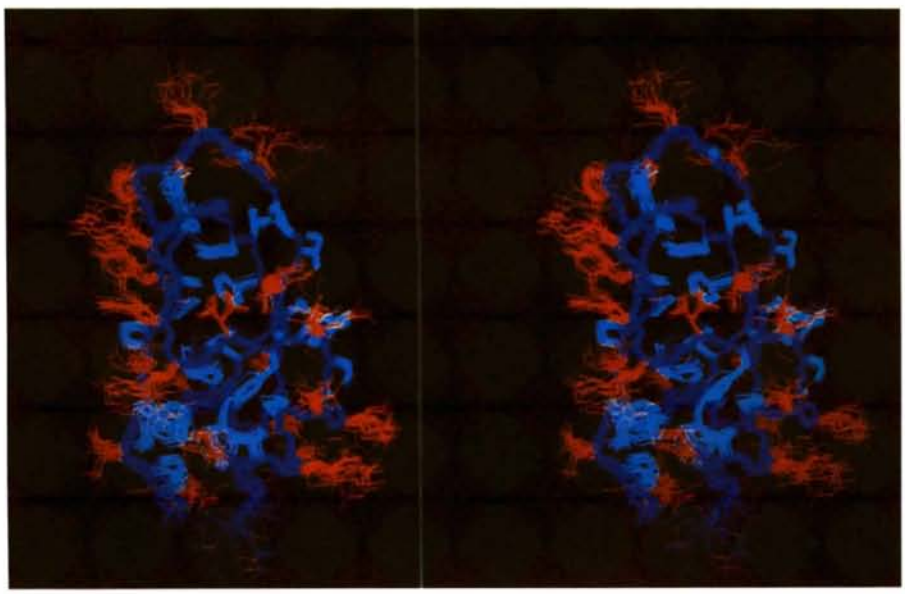

Fig. 19. Stereoview of the NMR solution structure of the protein BPTI represented as a superposition of 20 conformers. All heavy atoms are displayed. Colour code: green, polypeptide backbone; blue, core side chains; red, surface side chains. A close fit among the 20 conformers indicates regions of the molecule where the structure is well defined, whereas a large dispersion implicates (dynamic) disorder. (Drawing prepared using the atom coordinates from Berndt, Güntert, Orbons & Wüthrich, 1992.)

Structure determination with nucleic acids is pursued along similar lines as described here for proteins. As a general guide to an evaluation of the results obtained, one can assume that 'global folds', for example, formation of duplexes, triplexes, quadruplexes or loops can be well defined by NMR data. On the other hand, because of the short range and the low precision of individual NOE distance measurements, information on details in a given

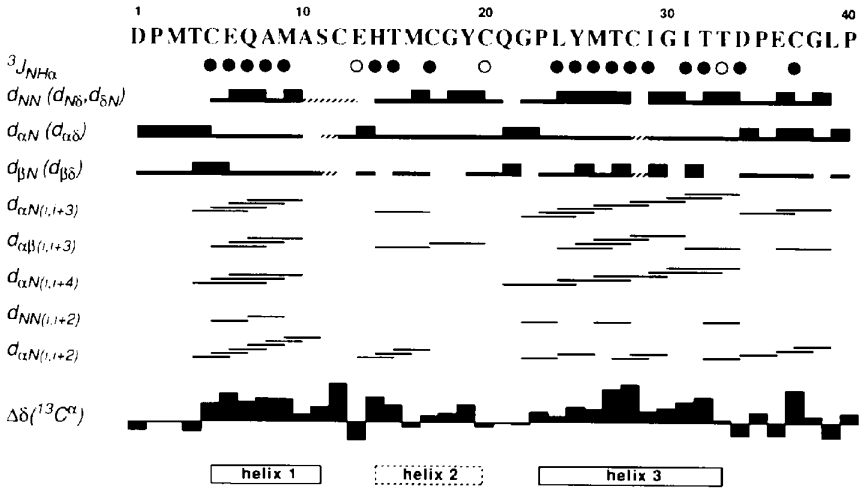

Fig. 20. Standard presentation of the data for regular secondary-structure identification in proteins. The 40-residue protein pheromone Er-2 is used as an illustration. Below the amino-acid sequence, open and filled circles identify residues with $^3J_{HN\alpha} > 8.0$ Hz, and $^3J_{HN\alpha} < 6.0$ Hz, respectively. For the sequential NOE connectivities $d_{\alpha N}$, d_{NN} and $d_{\beta N}$ ($d_{N\delta}$, $d_{\alpha\delta}$ and $d_{\beta\delta}$ for *Xxx*-Pro dipeptides, $d_{\delta N}$ for Pro-*Xxx*), thick and thin bars indicate strong and weak NOE intensities (hatched bars indicate connectivities that were observed only at 303 K, whereas all the other data were measured at 283 K). The observed medium-range NOE's $d_{\alpha N}(i, i+3)$, $d_{\alpha\beta}(i, i+3)$, $d_{\alpha N}(i, i+4)$, $d_{NN}(i, i+2)$, and $d_{\alpha N}(i, i+2)$, are indicated by lines connecting the two residues that are related by the NOE. $^{13}C^{\alpha}$ chemical shifts relative to the random coil values, $\Delta\delta(^{13}C^{\alpha})$, are plotted at the bottom of the figure, where positive values are shifts to lower field and the largest shift is about 5.8 p.p.m. Downfield shifts of the size observed here are indicative of helical secondary structure. The identification of the sequence of three helices indicated at the bottom was based on the patterns of continuous strong d_{NN} connectivities, small coupling constants $^3J_{HN\alpha}$ and $\Delta\delta(^{13}C^{\alpha})$-values > 2 p.p.m., the presence of medium-range NOE's, and slowed exchange for the hydrogen-bonded amide protons [(1); not shown in the figure]. The identification of helix 2 from these data is uncertain; the subsequent determination of the complete three-dimensional structure showed that this polypeptide segment forms a distorted 3_{10}-helix. (From Ottiger, Szyperski, Luginbühl, Ortenzi, Luporini, Bradshaw & Wüthrich, 1994.)

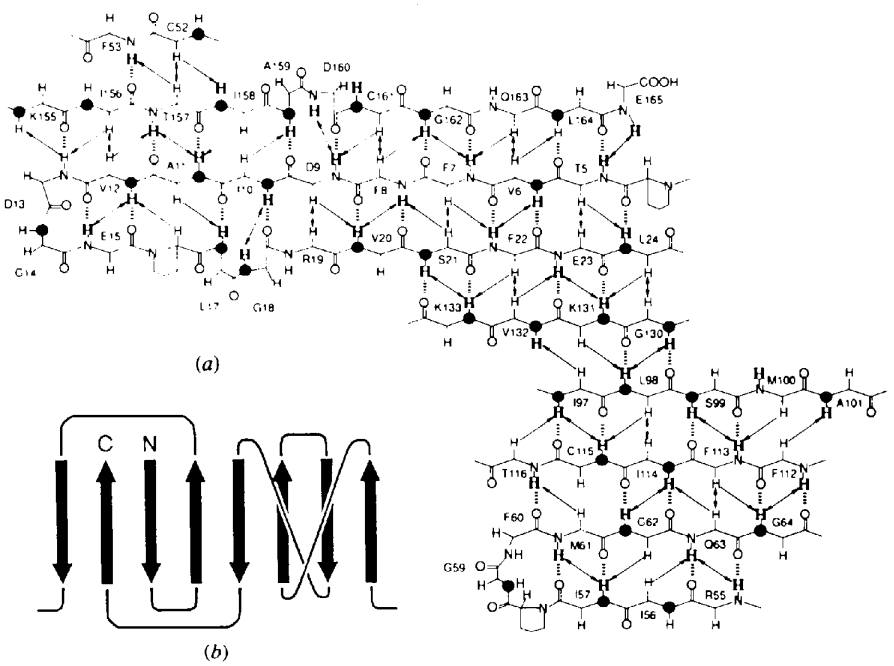

Fig. 21. (*a*) β-sheet indentified in recombinant human cyclophilin. All backbone atoms are shown, and except for Pro the amino-acid side chains have been omitted. The individual residues are identified near the C^{α} position by the one-letter amino-acid symbol and the sequence number. (Black circles at the amide N-atom position identify those residues which were labelled with ^{15}N in residue-specific labelling experiments.) Arrows represent the experimentally observed interstrand ^1H–^1H NOE's. Slowly exchanging amide protons are printed in bold. Hydrogen bonds that were unambiguously identified on the basis of interstrand NOE's and slowed exchange of the amide proton are identified with dashed lines. (*b*) Scheme of the β-sheet, which has a global (+1, −3, −1, −2, +1, −2, −3) topology, (Adapted from Wüthrich, Spitzfaden, Memmert, Widmer & Wider, 1991.)

nucleic acid structure type, for example bending of DNA duplexes or non-regular base stacking, may be more difficult to characterize.

6.5. *Stucture refinement*

The individual conformers obtained as the solutions of distance geometry calculations generally contain some close interatomic contacts and correspondingly high conformational energies. These are routinely removed in a separate final refinement step (Fig. 12). Alternatively, in typical molecular dynamics calculations the energy minimization may be performed in concert with the fitting to the distance constraints. Methods for direct refinement against the NOESY spectra by complete relaxation matrix calculations are in principle available (*e.g.* Boelens, Koning, van der Marel, van Boom & Kaptein, 1989; Borgias, Gochin, Kerwood & James, 1990; Mertz, Güntert, Wüthrich & Braun, 1991; Yip & Case, 1989), but because of a variety of fundamental and technical difficulties they have only rarely been used in actual structure determinations so far.

7. Identification of regular polypeptide secondary structures in proteins

The fact that regular secondary structures can be identified at an early stage of a protein structure determination by NMR (Fig. 12) is based on the following. The sequential NOE's (Fig. 15) are dependent on the local backbone conformation and thus indicative of regular secondary-structure elements (Billeter, Braun & Wüthrich, 1982). Corresponding information on the local backbone conformation comes from $^3J_{HN\alpha}$ coupling constants (Pardi, Bileter & Wüthrich, 1984). Furthermore, when identifying sequential NOE's, all other NOE's between amide protons, H^α and H^β must also be analyzed, which includes medium-range NOE's that are unique for helices and tight turns (Wüthrich, Billeter & Braun, 1984). This information is usually collected in a presentation of the type of Fig. 20, where patterns of successive small $^3J_{HN\alpha}$ constants and strong d_{NN} NOE's combined with the presence of medium-range NOE's provide reliable identification of helical secondary structure. Using additional long-range $d_{\alpha\alpha}(i,j)$ NOE connectivities and amide proton-exchange data, β-sheet structures can similarily be identified (Fig. 21). $^{13}C^\alpha$ chemical shifts have recently been added as yet another indicator of regular secondary structure (Spera & Bax, 1991; Wishart, Sykes & Richards, 1991), as is illustrated at the bottom of Fig. 20 and explained in the figure caption.

Early identification of regular secondary structure is not a necessary step in a protein structure determination (Fig. 12), since the conformational constraints of Fig. 20 are in any case part of the input for the structure calculation. However, it may be a valuable result in itself in situations where a structure determination cannot immediately be completed. This situation prevailed in all early NMR structure determinations (*e.g.* Williamson, Marion & Wüthrich, 1984; Kline & Wüthrich, 1985; Zuiderweg, Kaptein & Wüthrich, 1983*b*), and today it is typical for projects where uniform or residue-selective ^{15}N-labelling enables complete sequence-specific polypeptide backbone assignments *via* sequential NOE's (Fig. 15) but the extent of the side-chain assignments in the absence of ^{13}C-labelling is insufficient as a basis for a complete structure determination. In many instances, reports on backbone assignments and secondary-structure identification are, therefore, published long before a complete three-dimensional NMR structure is available.

8. Combined use of X-ray diffraction in single crystals and NMR in protein solutions

As was discussed in some detail in section 3 above, the chemical shifts in proteins or nucleic acids cannot be calculated with sufficiently high precision from the X-ray crystal structures to reliably predict the NMR spectra. Therefore, knowledge of the X-ray structure is of very limited use in a NMR structure determination. For the resonance assignments with sequential NOE's one may at some stage be able to derive information on what type of connectivity to look for. More useful support from knowledge of the X-ray structure may be obtained for the collection of conformational constraints. If there is evidence that the crystal and solution structures have the same architecture, the X-ray structure may be used as a starting reference to resolve ambiguities in the assignment of NOE distance constraints (Güntert, Berndt & Wüthrich, 1993), which may help to reduce the number of cycles needed to collect the input data for a high-quality structure determination (Fig. 12).

An attractive possibility for combined use of X-ray and NMR data arises when both a low-resolution electron-density map and a secondary-structure determination by NMR are available. In NMR secondary-structure determinations, the covalent connections between regular secondary-structure elements can unambiguously be determined (Figs. 20 and 21), so that this NMR information may be helpful in tracing the electron-density map for the determination of the corresponding X-ray crystal structure. In our experience this approach can be more efficient than either unassisted chain tracing or complete NMR structure determination (Kallen, Spilzfaden, Zurini, Wider, Widmer, Wüthrich & Walkinshaw, 1991). Furthermore, since identical global architectures are usually found for corresponding crystal and solution structures of globular proteins (see below), a high-quality NMR solution structure can, in principle, be used to solve the corresponding crystal structure by molecular replacement. This process was carried through, for example, with the α-amylase inhibitor Tendamistat mentioned previously (Braun, Epp, Wüthrich & Huber, 1989).

9. Complementary information from X-ray diffraction in crystals and NMR in solution

Since X-ray diffraction in crystals and NMR in solution can both be used independently to determine the complete three-dimensional structure of proteins, application of the two methods with the same proteins provides a basis for meaningful comparisons of corresponding structures in single crystals and in non-crystalline states. This is highly relevant, since the solution conditions for NMR studies may concide closely with the natural physiological environment of the protein, or they may be varied over a wide range for studies of structural transitions with pH, temperature or ionic strength (Wüthrich, 1986). Further complementarity of the two methods results from the fact that the time scales of the two measurements are widely different. Extensive similarities between corresponding crystal and solution in structures as well as major differences in conformational features of the two states have been documented, and are briefly surveyed in this section.

9.1. *Molecular architecture and core structure of globular proteins*

By now a significant number of globular proteins are available for which structures have independently been determined in crystals and in solution (for a recent review, see Billeter, 1992). For Tendamistat (Fig. 3) and BPTI (Fig. 19) particularly detailed comparisons of the two states have been made, which also includes solving the crystal structure by molecular replacement with the NMR structure (Braun, Epp, Wüthrich & Huber, 1989). Some general observations emerged from the analyses of these two proteins that have been found to be true generally for globular proteins (Billeter, 1992). The dominant impression is one of near-identity of the molecular architecture in solution and in single crystals (Fig. 3). In addition to the polypeptide backbone, the core side chains, which have small solvent accessibility, are well defined in both structures, and coincide very closely. This is also true for aromatic rings, which are seen in the same equilibrium orientations in the crystal structure and the NMR structure. The flipping of the aromatic rings is not explicitly manifested in NMR structures determined following the protocol of Fig. 12, although one has additional direct information on these dynamic processes (section 4 above). The same holds for the amide-proton exchange data [(1)]. High-frequency internal mobility of protein molecules in solution is allowed for in the calibration for the conformational constraints used as input for the structure calculation (Braun, Bösch, Brown, Gō & Wüthrich, 1981; Wüthrich, 1986). Largely as a consequence of this conservative calibration, the result of a structure determination by NMR spectroscopy consists of an ensemble of similar but not identical conformers that characterize the solution structure, and the spread between these conformers (usually expressed by the r.m.s.d.) gives an indication of the precision of the structure determination. The local r.m.s.d.'s along the polypeptide chain parallel qualitatively the variations of the crystallographic B factors (Fig. 3). [This comparison has been made with NMR conformers obtained with distance geometry calculations which, besides the experimental NMR constraints, take only the van der Waals atomic volumes into consideration, but corresponding observations result when the NMR conformers are further refined by energy minimization (Billeter, 1992).] Regions of the molecule showing increased disorder in the crystal structure are usually also poorly defined by the NMR data and *vice versa*. (There is nonetheless an important difference between the two methods: whereas disordered polypeptide segments may be difficult or impossible to see in electron-density maps, they will usually give outstandingly prominent lines in NMR spectra, enabling additional NMR measurements to be made that relate directly to the local mobility.)

Over the years there have been a small number of reports describing major rearrangements between the molecular architecture of globular proteins in crystals and in solution. Although such conformational rearrangements would, in principle, not be unexpected, in most of these instances there have been reasons to believe that the apparent structural differences might be attributable to shortcomings of one or both of the structure determinations (see metallothionein in section 2 above). Overall it appears that there is no well documented case as yet of a globular protein showing a global change of the molecular architecture between single crystals and non-denaturing solution milieus.

Although the presently available results thus show that there is usually very close coincidence between the molecular architecture and the detailed structure of the molecular core in corresponding X-ray and NMR structures of globular proteins, there is also extensive complementarity in the information on the protein core that is accessible with the two methods: X-ray diffraction can provide the desired infomation for big molecules and multimolecular assemblies such as intact viruses, whereas NMR applications are limited to smaller systems up to a molecular weight of about 30 000–40 000. On the other hand, NMR measurements provide additional, quantitative information on concerted internal mobility of proteins. Representative rate processes with activation energies of 50–100 kJ mol^{-1} and frequencies at ambient temperature in the millisecond to microsecond time range include the ring flips of phenylalanine and tyrosine (see sections 2 and 4 above), exchange of interior hyration water molecules with the bulk of solvent (see section 9.3 below), and interconversion of disulfide bonds between the *pro-R* and *pro-S* chiral forms (Otting, Liepinsh & Wüthrich, 1993). Furthermore, NMR studies of amide-proton exchange [(1)] and *cis–trans* isomerization of

Xxx-Pro peptide bonds (Wüthrich, 1976) afford insight into additional conformational equilibria in the protein core.

9.2. *Surface structure of globular proteins*

The increase of structural disorder near the surface of a protein tends to be more pronounced in the NMR structure than in the crystal structure (Fig. 19). Furthermore, significant differences between the average atomic coordinates in the crystal and in solution have in some instances been observed near the protein surface. Such differences may include individual amino-acid side chains, or clusters of solvent-exposed side chains. In some cases these differences could be attributed unambiguously to protein–protein contacts in the protein crystals (*e.g.* Billeter, 1992).

Surface disorder in structures determined with the protocol of Fig. 12 is usually directly related to a scarcity of NOE distance constraints near the protein surface. [When compared with the protein core there is also a scarcity of packing constraints, and surface side chains may be more intimately coupled *via* non-bonding contacts to the rapidly fluctuating solvent lattice (see section 9.3 below) than to the main body of the protein molecule (Fig. 19).] In turn, however, there are clear-cut implications that scarcity of NOE constraints is a direct consequence of dynamic disorder, so that the surface features seen in Fig. 19 are indeed meaningful and not simply an artifact of the NMR method. A nice example is hirudin. Using different NMR techniques, in particular pH titration of amide-proton chemical shifts (Bundi & Wüthrich, 1979), the formation of transient hydrogen bonds on the surface of this protein could be characterized in detail (Szyperski, Antuch, Schick, Betz, Stone & Wüthrich, 1994). Although these local surface structures are compatible with the protein structure determined with the approach of Fig. 12, the corresponding NOESY cross peaks have not been observed, presumably because the resonance lines are either broadened beyond detection or the NOE's are quenched by local conformational mobility (Wüthrich, 1986). A particularly clear-cut illustration of scarcity of NOE distance constraints due to line broadening by intramolecular mobility is afforded by the aforementioned interconversion of the disulfide bond Cys14—Cys38 in BPTI. In the otherwise precisely defined protein core, the *pro-R* and *pro-S* forms of this disulfide bond are both observed among the 20 conformers used to represent the NMR structure, whereas only the *pro-R* form is present in the corresponding crystal structure (Fig. 22) (Berndt, Güntert, Orbons & Wüthrich, 1992). At the time of the BPTI structure determination this ambiguity in the NMR structure was clearly related to scarcity of data near residues 14 and 38 in the input for the structure calculation. Subsequently, additional NMR studies revealed that there is an equilibrium between these two local conformational isomers, which are indeed both present and interchange at a frequency corresponding to lifetimes in the two states in the millisecond range. Because of the ensuing line broadening due to chemical-shift averaging (see Fig. 9, section 4), most of the NOE's involving H atoms of Cys14 and Cys38 could not be observed (Otting, Liepinsh & Wüthrich, 1993). In this way the NMR structure manifests a minor conformation with a population of about 2%, which could not be detected by diffraction techniques.

Overall, X-ray crystal structures and NMR solution structures provide qualitatively different information on the protein surface. In the crystals a sizeable proportion of surface amino-acid side chains are subject to similar packing constraints in protein–protein interfaces as the interior side chains in the protein core, and therefore they are rather precisely defined by the X-ray diffraction data. In NMR solution structures determined according to the protocol of Fig. 12, the surface is largely disordered (Fig. 19), where in addition to amino-acid side chains the backbone of the chain ends and extended surface loops may also be poorly defined (see section 9.4 below). Additional NMR experiments can provide information needed for more detailed characterization of the molecular surface (see section 9.3 below), but care must be exercised in the data analysis because of the presence of a multitude of equilibria between two or multiple transient local conformational states, of which the relative populations are usually not independently known.

9.3. *Protein hydration*

A recent major addition to the characterization of protein structures in solution is based on the ability to observe individual molecules of hydration water by

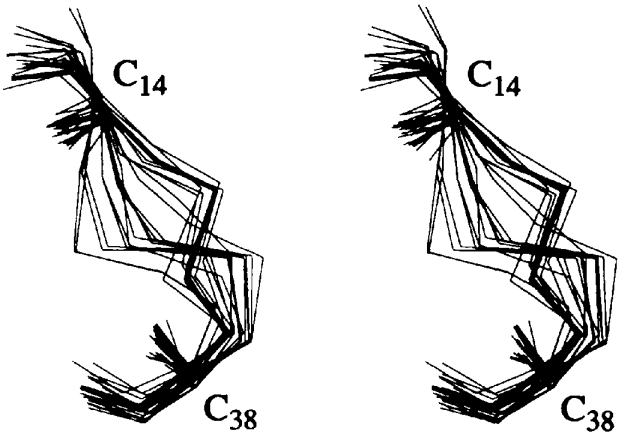

Fig. 22. Stereoview of the disulfide bond Cys14–Cys38 in BPTI observed in the X-ray crystal structure (bold line) and the NMR solution structure represented by 20 conformers (thin lines) after superposition of the backbone atoms N, C^{α} and C′ for minimal r.m.s.d. (from Berndt, Güntert, Orbous & Wüthrich, 1992).

NMR. Thereby the location of the hydration waters is determined by the observation of NOE's between water protons and H atoms of the polypeptide chain (Otting & Wüthrich, 1989). Because of the dependence of the NOE on the inverse sixth power of the ^1H–^1H distance, only water molecules in a first hydration layer are observed. For the hydration studies it is of further crucial importance that the NOE intensity is also related to a correlation function describing the stochastic modulation of the dipole–dipole coupling between the interacting protons. This correlation function may be governed either by the Brownian rotational tumbling of the hydrated protein molecule, or by interruption of the dipolar interaction through translational diffusion of the interacting spins, whichever is faster (Otting, Liepinsh & Wüthrich, 1991*a*). When comparing hydration data from X-ray or neutron diffraction in protein crystals with NMR data in solution, two different situations are encountered.

Firstly, interior hydration waters which are an integral part of the protein architecture have by now been observed in identical locations in the crystal and solution structures of many different proteins, implicating that their conservation between the two states is a general feature of globular proteins. However, as was directly demonstrated for BPTI and is clearly indicated for all other interior waters observed so far by NMR, additional information on molecular mobility is obtained from the NMR studies. BPTI contains four interior hydration water molecules, which are inaccessible to the solvent in a rigid model of the three-dimensional structure in crystals as well as in solution. Nonetheless, these waters exchange with the bulk solvent at rates corresponding to residence times in the protein hydration sites which are shorter than 20 ms (Otting, Liepinsh & Wüthrich, 1991*b*). (Using a different NMR approach it was recently shown that for at least one of the four interior hydration waters the residence time is actually shorter than 4 μs; Denisov & Halle, 1995*a*,*b*.) The effective correlation time for NOE's with these water molecules, as determined by the rotational tumbling of the protein molecule, is of the order of several nanoseconds. Long-lived 'interior-type' water has also been observed in the protein–DNA interface of the solution structure of an *Antennapedia* homeodomain–DNA complex (Qian, Otting & Wüthrich, 1993). Although hydration of the protein–DNA interface has previously been reported for crystal structures of DNA complexes with repressor proteins, most notably a *trp* repressor–DNA complex (Otwinowski, Schevitz, Zhang, Lawson, Joachimiak, Marmorstein, Luisi & Sigler, 1988), the NMR observation again provides novel additional information on dynamic aspects of the intermolecular recognition. Similar to the aforementioned interior waters in BPTI, the interfacial water molecules in the homeodomain–DNA complex exchange with the bulk water with lifetimes somewhere in the range from milliseconds to nanoseconds. Molecular dynamics simulations indicate that at any given moment the intermolecular hydration cavity contains from one to five water molecules, and that in concert with several flexibly disordered amino-acid side chains these waters move around the cavity (Fig. 23). With these functionally essential amino-acid side chains as well as water molecules moving between multiple contact sites on the DNA, the implications are of specific intermolecular recognition by a continuously ongoing scanning process on the milisecond to microsecond time scale.

Secondly, if one includes surface hydration water into the description of the protein structure, one finds that the

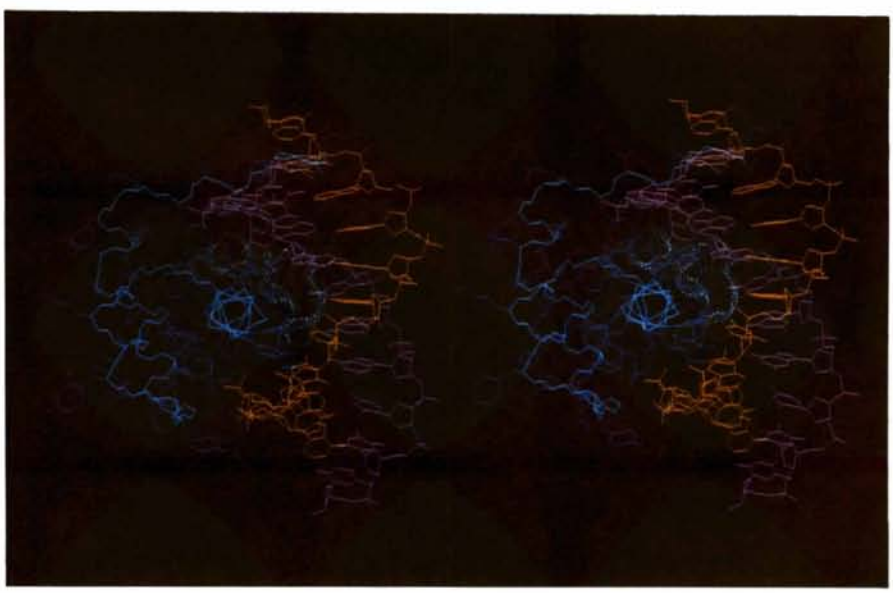

Fig. 23. Stereoview of an *Antennapedia* homeodomain complex with a 14-base pair DNA duplex. An all-heavy presentation of residues 3–57 of the homeodomain and the base pairs 3–13 of the DNA duplex is shown. Colour code: DNA strands, brown and magenta; polypeptide backbone, light blue; amino-acid side chains, dark blue. Blue dots signify the surface of a cavity with possible locations of hydration water molecules in the protein–DNA interface. (Drawing prepared using the atomic coordinates from Billeter, Qian, Otting, Müller, Gehring & Wüthrich, 1993.)

two methods provide genuinely different, complementary information. Individual hydration water molecules continuously exchange in and out of particular hydration sites on the protein, and NMR measures the average duration of these 'visits'. Diffraction experiments, on the other hand, probe the total fraction of time that a water molecule spends in a particular hydration site, but they are insensitive to the residence time at that site on any particular visit (Saenger, 1987). Typically, hydration water molecules in protein crystals are observed in discrete sites covering 30 to 60% of the molecular surface, depending on the protein and the crystal form, and although they usually give rise to somewhat bigger B factors the diffraction data on surface waters are not qualitatively different from those on the aforementioned interior water molecules. Considering the high water concentration in aqueous protein solutions used for NMR studies, one anticipates intuitively that the entire surface of the solute is covered with water molecules. In experiments with the nonapeptide oxytocin, all polypeptide H atoms did indeed show NOE's with the water, and the residence times for all hydration water molecules were found to be very short, in the approximate range 20–200 ps at 283 K (Otting, Liepinsh & Wüthrich, 1991*a*). Similar, although less complete, data were obtained with BPTI (Otting, Liepinsh, & Wüthrich, 1991*a*). Again, all surface hydration water molecules were found to have lifetimes in the hydration sites shorter than 300 ps, and these data could be rationalized with a molecular dynamics simulation extending over 2 ns (Brunne, Liepinsh, Otting, Wüthrich & van Gunsteren, 1993). These studies showed that there is no apparent correlation between the fact that particular surface hydration sites are or are not occupied in the crystal stucture and the length of the residence times of the hydration waters in these sites in aqueous solution.

The following comments may help to more fully appreciate the nature of the NMR information on protein hydration. In contrast to the situation in X-ray crystallography, the information on hydration water is not included in the input for the structure determination (Fig. 12), so that studies of hydration require additional, particularly demanding experiments (*e.g.* Otting, Liepinsh, Farmer & Wüthrich, 1991). Interior hydration water molecules with lifetimes exceeding 1 ns are characterized by strong negative NOE's, surface waters with lifetimes shorter than about 300 ps are observed by very weak positive NOE's [the NOE's with water molecules having residence times between about 800 and 300 ps are expected to be too weak to be detected in laboratory-frame experiments, but they can be observed as 'ROE's' in rotating-frame ROESY experiments (Otting, Liepinsh & Wüthrich, 1991*a*)]. NMR thus provides a qualitative criterion for distinguishing between short-lived, 'non-specific' surface hydration on the one hand, and on the other hand longer lived hydration in the protein core, in intermolecular interfaces or in particular solvent-accessible surface-hydration sites as observed in A·T tracts of DNA duplexes (Kubinec & Wemmer, 1992; Liepinsh, Otting & Wüthrich, 1992). By now, long-lived water has been reported in many different systems. In contrast, because of the low sensitivity for observation of NOE's associated with very short correlation times, surface hydration waters have so far been identified only in systems where the peptide, protein or nucleic acid concentration was higher than 5 mM. For future studies it will be of particular interest to extend solvation studies of the molecular surface to non-aqueous solvent components, in particular chemicals that promote denaturation of proteins and nucleic acids (Wüthrich, 1994).

9.4. *Biopolymers that have not been crystallized*

In all instances where a biological macromolecule cannot be crystallized, NMR is currently the only method capable of providing a three-dimensional structure. Examples are already too numerous to warrant individual citations; they include non-globular polypeptides, proteins containing extensive disordered loops or chain ends, proteins in strongly denaturing milieus, RNA's and DNA's, nucleic acid-binding proteins *etc.* In many instances, only low precision can be achieved for the NMR stucture determination of such compounds, which may be unsatisfactory and frustrating for those involved but highly relevant as it reflects the pronounced structural mobility that presumably prevented crystallization. From the technical point of view the reasons opposing structure determination at high precision are the same as discussed previously for the surface of globular proteins. For future applications, NMR studies of non-crystallizable intermediate states in protein and nucleic acid folding (*e.g.* Dyson, Merutka, Waltho, Lerner & Wright, 1992) will be particularly attractive.

Experience has shown that small proteins that cannot be crystallized are often accessible to diffraction studies in complexes with other proteins or with nucleic acids, where the flexible chain segments are stabilized in a unique conformation by the intermolecular contacts in the complex. This opens a wide array of possibilities for obtaining complementary information by X-ray diffraction and by NMR, since the complexes are often too big for complete NMR structure determination. Examples from our laboratory include the hirudin–thrombin system (Szyperski, Güntert, Stone, Tulinksky, Bode, Huber & Wüthrich, 1992) and systems involving procarboxypeptidases, their activation segments and the corresponding carboxypeptides (Billeter, Vendrell, Wider, Avilés, Coll, Guasch, Huber & Wüthrich, 1992).

10. Concluding remarks

A high-quality NMR solution structure is characterized by the facts that solvent-inaccessible interior regions of macromolecules, in particular the core of globular

proteins, are determined with high precision comparable to that achieved by X-ray crystallography at about 2.0 Å resolution, while solvent-accessible surface regions appear flexibly disordered. Although technically feasible, it is, therefore, not meaningful to characterize the quality of a NMR structure determination with a single parameter. Furthermore, a detailed comparison of corresponding crystal and solution structures is warranted only for the solvent-inaccessible interior molecular regions.

X-ray crystallography and solution NMR provide complementary information on biomacromolecular structures. For future practice it is advisable that the two techniques emphasize applications where they can provide unique information. Diffraction techniques with single crystals have the decisive advantages of being applicable for big systems and of providing high-precision mean structures. NMR in solution can provide high-precision data primarily on solvent-inaccessible molecular regions. Although the limitations on the precision of NMR structure determinations for molecular surfaces, or for the entire structure of non-globular molecules, are often frustrating, the resulting data are highly relevant and represent meaningful reference states for rationalizing the structural foundations and ensuing mechanistic aspects of specific intermolecular recognition. With the widespread availability of both techniques it seems attractive to use NMR data on the surface of small proteins as a basis for rational interpretation of larger crystal structures with regard to their physical–chemical and functional behaviour in solution. The aforementioned studies of surface hydration represent a step in this direction. While the crystallographic solvation data appear not to be directly relevant for rationalizing the solution state, molecular-dynamics simulations with the presently available potentials for aqueous solutions and using the atomic coordinates of the non-hydrated crystal structure seem to reproduce the experimental NMR observations in solution rather faithfully (Brunne, Liepinsh, Otting, Wüthrich & van Gunsteren, 1993).

Our research on NMR structure determination has been supported by the ETH Zürich, the Schweizerischer Nationalfonds, and the Kommission für wissenschaftliche Forschung (KWF) in joint projects with Spectrospin AG. I thank Mrs R. Hug and Mr R. Marani for the careful processing of the manuscript.

References

ALLERHAND, A., DODDRELL, D., GLUSHKO, V., COCHRAN, D. W., WENKERT, E., LAWSON, P. J. & GURD, F. R. N. (1971). *J. Am. Chem. Soc.* **93**, 544–546.

ANIL-KUMAR, ERNST, R. R. & WÜTHRICH, K. (1980). *Biochim. Biophys. Res. Commun.* **95**, 1–6.

ANIL-KUMAR, WAGNER, G., ERNST, R. R. & WÜTHRICH, K. (1981). *J. Am. Chem. Soc.* **103**, 3654–3658.

BAX, A. (1989). *Annu. Rev. Biochem.* **58**, 223–256.

BAX, A. & GRZESIEK, S. (1993). *Acc. Chem. Res.* **26**, 131–138.

BERNDT, K. D., GÜNTERT, P., ORBONS, L. P. M. & WÜTHRICH, K. (1992). *J. Mol. Biol.* **227**, 757–775.

BILLETER, M. (1992). *Q. Rev. Biophys.* **25**, 325–377.

BILLETER, M., BRAUN, W. & WÜTHRICH, K. (1982). *J. Mol. Biol.* **155**, 321–346.

BILLETER, M., KLINE, A. D., BRAUN, W., HUBER, R. & WÜTHRICH, K. (1989). *J. Mol. Biol.* **206**, 677–687.

BILLETER, M., QIAN, Y. Q., OTTING, G., MÜLLER, M., GEHRING, W. J. & WÜTHRICH, K. (1993). *J. Mol. Biol.* **234**, 1084–1093.

BILLETER, M., VENDRELL, J., WIDER, G., AVILÉS, F. X., COLL, M., GUASCH, A., HUBER, R. & WÜTHRICH, K. (1992). *J. Biomol NMR*, **2**, 1–10.

BLOCH, F., HANSEN, W. W. & PACKARD, M. E. (1946). *Phys. Rev.* **69**, 127.

BLUNDELL, T. L. & JOHNSON, L. N. (1976). *Protein Crystallography*. New York: Academic Press.

BOELENS, R., KONING, T. M. G., VAN DER MAREL, G. A., VAN BOOM, J. H. & KAPTEIN, R. (1989). *J. Magn. Reson.* **82**, 290–308.

BORGIAS, B. A., GOCHIN, M., KERWOOD, O. S. & JAMES, T. L. (1990). *Progr. NMR Spectrosc.* **22**, 83–100.

BRAUN, W., BÖSCH, C., BROWN, L. R., GŌ, N. & WÜTHRICH, K. (1981). *Biochim. Biophys. Acta*, **667**, 377–396.

BRAUN, W., EPP, O., WÜTHRICH, K. & HUBER, R. (1989). *J. Mol. Biol.* **206**, 669–676.

BRAUN, W. & GŌ, N. (1985). *J. Mol. Biol.* **186**, 611–626.

BRAUN, W., VAŠÁK, M., ROBBINS, A. H., STOUT, C. D., WAGNER, G., KÄGI, J. H. R. & WÜTHRICH, K. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 10124–10128.

BRAUN, W., WAGNER, G., WÖRGÖTTER, E., VAŠÁK, M., KÄGI, J. H. R. & WÜTHRICH, K. (1986). *J. Mol. Biol.* **187**, 125–129.

BRÜNGER, A. T., CLORE, G. M., GRONENBORN, A. M. & KARPLUS, M. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 3801–3805.

BRUNNE, R. M., LIEPINSH, E., OTTING, G., WÜTHRICH, K. & VAN GUNSTEREN, W. F. (1993). *J. Mol. Biol.* **231**, 1040–1048.

BUNDI, A. & WÜTHRICH, K. (1979). *Biopolymers*, **18**, 299–311.

CAMPBELL, I. D., DOBSON, C. M. & WILLIAMS, R. J. P. (1975). *Proc. R. Soc. London Ser. B*, **189**, 503–509.

CLORE, G. M. & GRONENBORN, A. M. (1991). *Science*, **252**, 1390–1399.

CRIPPEN, G. M. & HAVEL, T. F. (1988). *Distance Geometry and Molecular Conformation*. New York: John Wiley.

DEISENHOFER, J. & STEIGEMANN, W. (1975). *Acta Cryst.* **B31**, 238–250.

DENISOV, V. P. & HALLE, B. (1995*a*). *J. Mol. Biol.* **245**, 682–697.

DENISOV, V. P. & HALLE, B. (1995*b*). *J. Mol. Biol.* **245**, 698–709.

DICKERSON, R. E. & GEIS, I. (1969). *The Structure and Action of Proteins*. New York: Harper & Row.

DUBS, A., WAGER, G. & WÜTHRICH, K. (1979). *Biochim. Biophys. Acta*, **577**, 177–194.

DYSON, H. J., MERUTKA, G., WALTHO, J. P., LERNER, R. A. & WRIGHT, P. E. (1992). *J. Mol. Biol.* **226**, 795–817.

EHRENBERG, A., MALMSTRÖM, B. G. & VÄNNGARD, T. (1967). Editors. *Magnetic Resonance in Biological Systems*. New York: Pergamon Press.

ERNST, R. R., BODENHAUSEN, G. & WOKAUN, A. (1987). *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford: Clarendon.

FUREY, W. F., ROBINSON, A. H., CLANCY, L. L., WINGE, D. R., WANG, B. C. & STOUT, C. D. (1986). *Science*, **231**, 701–710.

GELIN, B. R. & KARPLUS, M. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 2002–2006.

GORDON, S. L. & WÜTHRICH, K. (1978). *J. Am. Chem. Soc.* **100**, 7094–7096.

GÜNTERT, P., BERNDT, K. D. & WÜTHRICH, K. (1993). *J. Biomol. NMR*, **3**, 601–606.

GÜNTERT, P., BRAUN, W. & WÜTHRICH, K. (1991). *J. Mol. Biol.* **217**, 517–530.

HAVEL, T. F. & WÜTHRICH, K. (1984). *Bull. Math. Biol.* **46**, 673–698.

HAVEL, T. F. & WÜTHRICH, K. (1985). *J. Mol. Biol.* **182**, 281–294.

HENDRICKSON, W. A. & WÜTHRICH, K. (1991). Editors. *Macromolecular Structures* 1991. London: Current Biology.

HENDRICKSON, W. A. & WÜTHRICH, K. (1992). Editors. *Macromolecular Structures* 1992. London: Current Biology.
HENDRICKSON, W. A. & WÜTHRICH, K. (1993). Editors. *Macromolecular Structures* 1993. London: Current Biology.
HENDRICKSON, W. A. & WÜTHRICH, K. (1994). Editors. *Macromolecular Structures* 1994. London: Current Biology.
HETZEL, R., WÜTHRICH, K., DEISENHOFER, J. & HUBER, R. (1976). *Biophys. Struct. Mech.* **2**, 159–180.
HULL, W. E. & SYKES, B. D. (1975). *J. Mol. Biol.* **98**, 121–153.
IKURA, M., KAY, L. E. & BAX, A. (1990). *Biochemistry*, **29**, 4659–4667.
JARDETZKY, O. (1967). *Naturwissenschaften*, **54**, 149–155.
JARDETZKY, O. & ROBERTS, G. C. K. (1981). *NMR in Molecular Biology*. New York: Academic Press.
KALLEN, J., SPITZFADEN, C., ZURINI, M. G. M., WIDER, G., WIDMER, H., WÜTHRICH, K. & WALKINSHAW, M. D. (1991). *Nature (London)*, **353**, 276–279.
KAPTEIN, R., BOELENS, R., SCHEEK, R. M. & VAN GUNSTEREN, W. F. (1988). *Biochemistry*, **27**, 5389–5395.
KAPTEIN, R., ZUIDERWEG, E. R. P., SCHEEK, R. M., BOELENS, S. R. & VAN GUNSTEREN, W. F. (1985). *J. Mol. Biol.* **182**, 179–182.
KARPLUS, S., SNYDER, G. H. & SYKES, B. D. (1973). *Biochemistry*, **12**, 1323–1329.
KEARNS, D. R., PATEL, D. J. & SHULMAN, R. G. (1971). *Nature (London)*, **229**, 338–339.
KENDREW, J. (1963). *Science*, **139**, 1259–1266.
KLINE, A. D., BRAUN, W. & WÜTHRICH, K. (1986). *J. Mol. Biol.* **189**, 377–382.
KLINE, A. D., BRAUN, W. & WÜTHRICH, K. (1988). *J. Mol. Biol.* **204**, 675–724.
KLINE, A. D. & WÜTHRICH, K. (1985). *J. Mol. Biol.* **183**, 503–507.
KUBINEC, M. G. & WEMMER, D. E. (1992). *J. Am. Chem. Soc.* **114**, 8739–8740.
LIEPINSH, E., OTTING, G. & WÜTHRICH, K. (1992). *Nucleic Acids Res.* **20**, 6549–6553.
MCDONALD, C. C. & PHILLIPS, W. D. (1967). *J. Am. Chem. Soc.* **89**, 6332–6341.
MCDONALD, C. C., PHILLIPS, W. D. & VINOGRADE, S. N. (1969). *Biochem. Biophys. Res. Commun.* **36**, 442–449.
MASSON, A. & WÜTHRICH, K. (1973). *FEBS Lett.* **31**, 114–118.
MEADOWS, D. H., JARDETZKY, O., EPAND, R. M., RÜTERJANS, H. H. & SCHERAGA, H. A. (1968). *Proc. Natl Acad. Sci. USA*, **60**, 766–772.
MERTZ, J. E., GÜNTERT, P., WÜTHRICH, K. & BRAUN, W. (1991). *J. Biomol. NMR*, **1**, 257–269.
MESSERLE, B. A., SCHÄFFER, A., VAŠÁK, M., KÄGI, J. H. R. & WÜTHRICH, K. (1990). *J. Mol. Biol.* **214**, 765–779.
MILDVAN, A. S. & COHN, M. (1970). *Adv. Enzymol.* **33**, 1–70.
MONTELIONE, G. T. & WAGNER, G. (1989). *J. Am. Chem. Soc.* **111**, 5474–5475.
MONTELIONE, G. T. & WAGNER, G. (1990). *J. Magn. Reson.* **87**, 183–188.
NAGAYAMA, K., ANIL-KUMAR, WÜTHRICH, K. & ERNST, R. R. (1980). *J. Magn. Reson.* **40**, 321–334.
NAGAYAMA, K., WÜTHRICH, K., BACHMANN, P. & ERNST, R. R. (1977). *Biochem. Biophys. Res. Commun.* **78**, 99–105.
NAGAYAMA, K., WÜTHRICH, K. & ERNST, R. R. (1979). *Biochem. Biophys. Res. Commun.* **90**, 305–311.
NÉMETHY, G. & SCHERAGA, H. A. (1977). *Q. Rev. Biophys.* **10**, 239–352.
OKHANOV, V. V., AFANAS'EV, V. A. & BYSTROV, V. F. (1980). *J. Magn. Reson.* **40**, 191–195.
OTTIGER, M., SZYPERSKI, T., LUGINBÜHL, P., ORTENZI, C., LUPORINI, P., BRADSHAW, R. A. & WÜTHRICH, K. (1994). *Protein Sci.* **3**, 1515–1526.
OTTING, G., LIEPINSH, E., FARMER, B. T. II & WÜTHRICH, K. (1991). *J. Biomol. NMR*, **1**, 209–215.
OTTING, G., LIEPINSH, E. & WÜTHRICH, K. (1991*a*). *Science*, **254**, 974–980.
OTTING, G., LIEPINSH, E. & WÜTHRICH, K. (1991*b*). *J. Am. Chem. Soc.* **113**, 4363–4364.
OTTING, G., LIEPINSH, E. & WÜTHRICH, K. (1993). *Biochemistry*, **32**, 3571–3582.
OTTING, G. & WÜTHRICH, K. (1989). *J. Am. Chem. Soc.* **111**, 1871–1875.
OTTING, G. & WÜTHRICH, K. (1990). *Q. Rev. Biophys.* **23**, 39–96.
OTWINOWSKI, Z., SCHEVITZ, R. W., ZHANG, R. G., LAWSON, C. L., JOACHIMIAK, A., MARMORSTEIN, R. Q., LUISI, B. F. & SIGLER, P. B. (1988). *Nature (London)*, **335**, 321–329.
PARDI, A., BILLETER, M. & WÜTHRICH, K. (1984). *J. Mol. Biol.* **180**, 741–751.
PERUTZ, M. F. (1963). *Science*, **140**, 863–869.
PFLUGRATH, J., WIEGAND, E., HUBER, R. & VÉRTESY, L. (1986). *J. Mol. Biol.* **189**, 383–386.
PURCELL, E. M., POUND, R. V. & BLOEMBERGEN, N. (1946). *Phys. Rev.* **70**, 986–987.
QIAN, Y. Q., BILLETER, M., OTTING, G., MÜLLER, M., GEHRING, W. J. & WÜTHRICH, K. (1989). *Cell*, **59**, 573–580.
QIAN, Y. Q., OTTING, G. & WÜTHRICH, K. (1993). *J. Am. Chem. Soc.* **115**, 1189–1190.
SAENGER, W. (1987). *Annu. Rev. Biophys. Chem.* **16**, 93–114.
SAUNDERS, M., WISHNIA, A. & KIRKWOOD, J. G. (1957). *J. Am. Chem. Soc.* **79**, 3289–3290.
SHULMAN, R. G., OGAWA, S., WÜTHRICH, K., YAMANE, T., PEISACH, J. & BLUMBERG, W. E. (1969). *Science*, **165**, 251–257.
SODANO, P., CHARY, K. V. R., BJÖRNBERG, O., HOLMGREN, A., KREN, B., FUCHS, J. A. & WÜTHRICH, K. (1991). *Eur. J. Biochem.* **200**, 369–377.
SPERA, S. & BAX, A. (1991). *J. Am. Chem. Soc.* **113**, 5490–5492.
SZYPERSKI, T., ANTUCH, W., SCHICK, M., BETZ, A., STONE, S. R. & WÜTHRICH, K. (1994). *Biochemistry*. **33**, 9303–9310.
SZYPERSKI, T., GÜNTERT, P., STONE, S. R., TULINSKY, A., BODE, W., HUBER, R. & WÜTHRICH, K. (1992). *J. Mol. Biol.* **228**, 1206–1211.
SZYPERSKI, T., WIDER, G., BUSHWELLER, J. H. & WÜTHRICH, K. (1993). *J. Am. Chem. Soc.* **115**, 9307–9308.
WAGNER, G. (1980). *FEBS Lett.* **112**, 280–284.
WAGNER, G. & WÜTHRICH, K. (1979). *J. Magn. Reson.* **33**, 675–680.
WAGNER, G. & WÜTHRICH, K. (1982). *J. Mol. Biol.* **155**, 347–366.
WIDER, G., LEE, K. H. & WÜTHRICH, K. (1982). *J. Mol. Biol.* **155**, 367–388.
WILLIAMSON, M. P., HAVEL, T. F. & WÜTHRICH, K. (1985). *J. Mol. Biol.* **182**, 295–315.
WILLIAMSON, M. P., MARION, D. & WÜTHRICH, K. (1984). *J. Mol. Biol.* **173**, 341–359.
WISHART, D. S., SYKES, B. D. & RICHARDS, F. M. (1991). *J. Mol. Biol.* **222**, 311–333.
WÜTHRICH, K. (1969). *Proc. Natl Acad. Sci. USA*, **63**, 1071–1078.
WÜTHRICH, K. (1970). *Struct. Bonding*, **8**, 53–121.
WÜTHRICH, K. (1976). *NMR in Biological Research: Peptides and Proteins*. Amsterdam: North Holland.
WÜTHRICH, K. (1986). *NMR of Proteins and Nucleic Acids*. New York: John Wiley.
WÜTHRICH, K. (1989*a*). *Science*, **243**, 45–50.
WÜTHRICH, K. (1989*b*). *Acc. Chem. Res.* **22**, 36–44.
WÜTHRICH, K. (1994). *Toward a Molecular Basis of Alcohol Use and Abuse*, edited by B. JANSSON, H. JÖRNVALL, U. RYDBERG, L. TERENIUS & B. L. VALLEE, pp. 261–268. Basel: Birkhäuser Verlag.
WÜTHRICH, K., BILLETER, M. & BRAUN, W. (1983). *J. Mol. Biol.* **169**, 949–961.
WÜTHRICH, K., BILLETER, M. & BRAUN, W. (1984). *J. Mol. Biol.* **180**, 715–740.
WÜTHRICH, K., SPITZFADEN, C., MEMMERT, K., WIDMER, H. & WIDER, G. (1991). *FEBS Lett.* **285**, 237–247.
WÜTHRICH, K. & WAGNER, G. (1975). *FEBS Lett.* **50**, 265–268.
WÜTHRICH, K., WIDER, G., WAGNER, G. & BRAUN, W. (1982). *J. Mol. Biol.* **155**, 311–319.
YIP, P. & CASE, D. A. (1989). *J. Magn. Reson.* **83**, 643–648.
ZUIDERWEG, E. R. P., KAPTEIN, R. & WÜTHRICH, K. (1983*a*). *Eur. J. Biochem.* **137**, 279–292.
ZUIDERWEG, E. R. P., KAPTEIN, R. & WÜTHRICH, K. (1983*b*). *Proc. Natl Acad. Sci. USA*, **80**, 5837–5841.