

Nuclear magnetic resonance studies and molecular dynamics simulations of the solution conformation of a 'designed', α -helical peptide

Werner Klaus¹ and Rudolf Moser²

Gesellschaft für Biotechnologische Forschung (GBF), Abt. Molekulare Strukturforchung, Mascheroder Weg 1, D-3300 Braunschweig, Germany and ²Biochemisches Institut der Universität Zürich, Winterthurerstr. 190, CH-8052 Zürich, Switzerland

¹Present address: F.Hoffmann-LaRoche, Ltd, Dept. Central Research Unit, Building 65, CH-4002 Basel, Switzerland

The complete three-dimensional structure in methanol of an amphipathic α -helical peptide, that has been designed by taking into account the three-dimensional structures of small haemolytic peptides, secondary structure prediction algorithms and the well documented literature on α -helix stabilizing factors, has been elucidated by two-dimensional NMR spectroscopy. Initially various two-dimensional spectra (COSY, TOCSY, and NOESY) allowed the complete sequence specific assignment of all signals in the ¹H spectrum. Consequently trial structures were generated which were then subjected to molecular dynamics simulations using 121 NOE-derived distances and 25 vicinal coupling constant values as structural restraints to give a final set of calculated structures. These structures are in complete agreement with the results of a circular dichroism study and reveal that the peptide adopted a highly ordered α -helical conformation. Details of the structure which throw light on future peptide/protein design are discussed.

Key words: amphipathic α -helix/haemolytic peptide/NMR structure/secondary structure design

Introduction

Currently one of the unresolved key problems in molecular biology is the understanding of the relationship between the primary amino acid sequence of a protein or peptide and its three-dimensional structure. The contribution of stabilizing forces which determine the secondary and tertiary structure can best be studied by the synthesis of small polypeptides that are designed to adopt a specific desired conformation. Experimental structure elucidation, either spectroscopically or crystallographically, allows an assessment of the accuracy of the model construction and an analysis of the contribution of structure stabilizing factors.

Our limited understanding of how proteins or peptides fold into their specific conformation makes the use of short and easily accessible synthetic peptides a powerful tool (Wright *et al.*, 1988) and can reduce the complexity of such systems by permitting the selective examination of only one structural building block (e.g. an α -helix). The fact that short linear peptides in isolation are capable of adopting a stable α -helical conformation is of considerable importance not only for the initiation of the protein folding process but also because it validates the *de novo* design of peptides with tailored structures and novel properties (DeGrado, 1988; DeGrado *et al.*, 1989). Designing helical peptides is greatly facilitated by the existence of detailed information about helix stabilizing factors (Shoemaker *et al.*, 1987) and about the relative helix forming propensities of the

various amino acid residues (Chou and Fasman, 1974; Argos and Palau, 1982; Palau *et al.*, 1982).

A Helical Erythrocyte Lysing Peptide (HELP) with a length of 26 residues, having a stable, amphiphilic, α -helical structure, was designed using secondary structure prediction methods and molecular modelling techniques (see Moser, 1992). Well characterized lytically active peptides were taken as the main structural template and together with the results from modelling studies the amino acid sequence produced by the design strategy was synthesized by the Merrifield solid-phase peptide synthesis method. HELP induced pH-triggered cell disruption using human erythrocytes as targets, as would be expected from a surface-active amphiphilic helix.

The circular dichroism (CD) spectrum of HELP is typical of an all α -helical peptide and highly suggestive of the presence of a unique folded, helical structure (Greenfield and Fasman, 1969) with the value of $[\theta]_{222}$ corresponding to a helix content of ~100% (Provencher and Glöckner, 1981). In spite of the similarity of CD spectra recorded in both aqueous and methanolic solutions (see Moser, 1992) the NMR spectra of HELP in the latter solvent showed drastically enhanced spectral resolution. Thus methanol stabilizes the monomeric helical structure whereas the peptide tends to form an aggregated helical state in aqueous solution (see Moser, 1992). Hence the complete three-dimensional structure of HELP dissolved in methanolic solution was elucidated by high-resolution 2D ¹H-NMR-studies and molecular dynamics (MD) simulations.

Materials and methods

Peptide synthesis and sample preparation

HELP was synthesized by the Merrifield solid-phase peptide synthesis method using *tert*-butyloxycarbonyl-chemistry on a *p*-methylbenzhydrylamine resin in order to obtain a C-terminal amide. The synthetic peptide was purified by reverse-phase HPLC and its homogeneity was confirmed by amino acid composition analysis, complete sequence analysis and FAB mass spectral analysis. For the NMR-measurements one sample of 5.5 mg HELP was dissolved in 500 μ l CD₃OD (100% D, Janssen Chimica, Nettetal, Germany) to which 1.5 μ l DCI/D₂O (2% DCI) was added. A second sample consisted of 7 mg peptide dissolved in 500 μ l CD₃OH (Merck, Sharp and Dohme Isotopes) with the same amount of HCl/H₂O added.

Light scattering experiments

Light scattering experiments were performed on 1 mM and 4.5 mM methanolic solutions of the peptide using the system ALV-3000 (ALV Lasertriebgesellschaft mbH, Langen, Germany). The correlation functions were analysed with the help of the program CONTIN (Provencher, 1982). The observed size distributions were bimodal. A first peak corresponding to the solubilized peptide co-existed with a second peak indicative of dust or particle aggregates at least an order of magnitude larger. As the scattered intensity is proportional to the square of the molecular weight of the scatterers, the latter represent an insignificant contamination without any relevance. The

hydrodynamic radius was obtained from the third moment of the distribution of the first peak. The values of $1.4 \text{ nm} \pm 0.3 \text{ nm}$ for the first sample and $1.2 \text{ nm} \pm 0.2 \text{ nm}$ for the second sample are consistent with a non-aggregated state of the peptide under these conditions.

NMR experiments

NMR spectra were recorded on a BRUKER AM 600 spectrometer with time proportional phase increment (Redfield and Kunz, 1975) for quadrature detection in the t_1 -dimension and at temperatures between 277 K and 297 K. The temperature was adjusted using a stream of nitrogen gas which was precooled in a heat exchanger immersed in a thermostated liquid bath (Haake GmbH, Karlsruhe, Germany). Spectra were referenced internally with the residual signal of the methyl group of methanol set to 3.35 p.p.m. relative to 2,2-dimethyl-2-silapentanesulfonate (DSS). 1D NMR spectra recorded with peptide concentrations between 0.3 mM and 4.5 mM showed no concentration-dependent chemical shifts of the amide protons that would be indicative of peptide aggregation under these conditions.

A phase-sensitive COSY-spectrum (Marion and Wüthrich, 1983) was recorded in conjunction with a NOESY-spectrum (Jeener *et al.*, 1979; Macura *et al.*, 1981) on the CD_3OD -sample as part of the COCONOSY-experiment of 277 K (Haasnoot *et al.*, 1984; Gurevich *et al.*, 1984). A double-quantum filtered COSY-spectrum (Rance *et al.*, 1983) was acquired at the same temperature with the peptide dissolved in CD_3OH . With the same sample a 2D TOCSY was recorded in the 'CLEAN' version as described (Griesinger *et al.*, 1988) with the MLEV-17 mixing sequence (Bax and Davis, 1985; Davis and Bax, 1985) but without the trim pulses. The total mixing time was 100.2 ms. NOESY-spectra were acquired with the usual pulse sequence ($90-t_1-90^\circ-\tau-90^\circ-\text{acq}$) with mixing times ranging from 60 to 240 ms. In all measurements the solvent resonance line was irradiated during the relaxation delay (1 s) and additionally during the mixing time of the NOESY-experiments. All spectra were recorded using O1/O2-phase coherence (Zuiderweg *et al.*, 1986).

To evaluate $^3J_{\text{HN-H}\alpha}$ -values a 1D-spectrum with a high digital resolution was recorded, zero-filled to 64 k real data points, multiplied by a Gaussian window and Fourier-transformed. For overlapping resonance lines the coupling constants were extracted from the 2D-COSY-DQF-spectrum. For all 2D-spectra typically 512 t_1 -increments were used with 64 scans sampled per increment with a spectral width of 8064 Hz for both dimensions. The data were transferred to an ASPECT X32 computer for 2D-Fourier transformation and phase correction yielding data matrices of $1\text{k} * 1\text{k}$ real data points for COSY-DQF and NOESY and $2\text{k} * 1\text{k}$ for CLEAN TOCSY. The assignment of the cross-peaks was greatly facilitated by the program AURELIA (Neidig and Kalbitzer, 1990), which was also used for the determination of the volumes of the NOE cross-peaks. Their intensities were classified into five groups and translated into distances by an empirical calibration of cross-peaks arising from known distances within regular secondary structures (Billeter *et al.*, 1982). Distance restraints were used as upper bounds with appropriate corrections for methyl groups or methylene protons which were treated as pseudo-atoms (Wüthrich *et al.*, 1983).

Restrained molecular dynamics simulations

Restrained molecular dynamics simulations and analysis of the resulting structures were carried out on a TRACE (Multiflow Corp.) mini-supercomputer using the GROMOS software package (van Gunsteren and Berendsen, 1987). Creation of the peptide in an extended configuration and display of structures

was accomplished using the GBF-program BRAGI (Schomburg and Reichelt, 1988) running under UNIX on a Hewlett Packard Graphics System 9000. The approach for the creation of the initial 'pseudo-random' structures was influenced by Nilges *et al.* (1988), whereas the protocol that was used for the main part of the structure calculations was adopted from Bazzo *et al.* (1988) and Klaus *et al.* (1991).

The initial structures for the molecular dynamics calculations were created as follows. The peptide, in an extended conformation with all backbone angles artificially set to 180° , was subjected to energy minimization to relieve the internal strain. Molecular dynamics runs without restraints and with velocities taken from a Maxwellian distribution corresponding to temperatures between 900 K and 1350 K were performed for time periods between 3 ps and 6 ps. These structures were energy minimized and used as input for the calculations with restraints. Each one of these starting structures was subjected to energy minimization using 500 steps of steepest descent with a small force constant of $200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ for the distance restraints and 10 kJ mol^{-1} for the dihedral restraints. This was followed by 2 ps of MD at 290 K using the same force constants. The calculations were continued at 1000 K with force constants set to $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ and 35 kJ mol^{-1} for 4 ps and with values of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ and 70 kJ mol^{-1} for another 4 ps. For the rest of the calculations the force constant for the distance restraints was set to $1500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. 3 ps of cooling to 600 K were followed by 3 ps of dynamics, 2 ps cooling to 290 K and 30 ps of dynamics at this temperature. During the 'high temperature stage' between 600 K and 1000 K the peptide bonds were additionally restrained to prevent *cis/trans*-isomerizations. For the last 20 ps of each run, which were used for analysing the resulting structure, a force constant of $3000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ for the distance restraints was applied. The cutoff radius for calculating non-bonded interactions was set to 0.8 nm for both energy minimization and molecular dynamics with a list of interactions updated every 10 cycles. To constrain bondlengths the SHAKE algorithm was used allowing a time step of 1 fs for the simulations.

Results

An expansion showing the amide region of the ^1H NMR spectrum of HELP is given in Figure 1. Most of the signals are clearly resolved and allow direct measurement of the $^3J_{\text{HN-H}\alpha}$ -coupling constants. These, as listed in Table I, are a strong indication of a helical secondary structure of the peptide.

Assignment of spin systems

The TOCSY-spectrum (Figure 2) with a 100 ms mixing time shows the correlation of ten amide protons with methyl-groups resonating around 1.0 p.p.m. which can only arise from leucine spin-systems. The only other methyl-bearing amino acids present in the peptide—Thr4 and Thr7—can be distinguished from the leucines as the Thr-amide proton shows characteristic cross-peaks in the 'fingerprint region' with both its α - and β -CH. The NH of Gly3 is readily assigned as it is the only proton visible at $T = 290 \text{ K}$ which shows correlations with just two others at 3.86 and 4.01 p.p.m. This is confirmed by inspection of the COSY spectrum where a cross-peak between both C_α -protons of Gly1 can be observed. The NH of Gly1 is observed only in spectra recorded at $T = 277 \text{ K}$ where it resonates at 8.28 p.p.m. Both Phe-residues are assigned based on their cross-peak pattern which is typical of AMX-spin systems, and are readily correlated with their amide protons in the TOCSY spectrum. This spectrum also gives the total correlation of the 'long side chain' protons of the two Lys and two Arg residues with their respective NHs. This

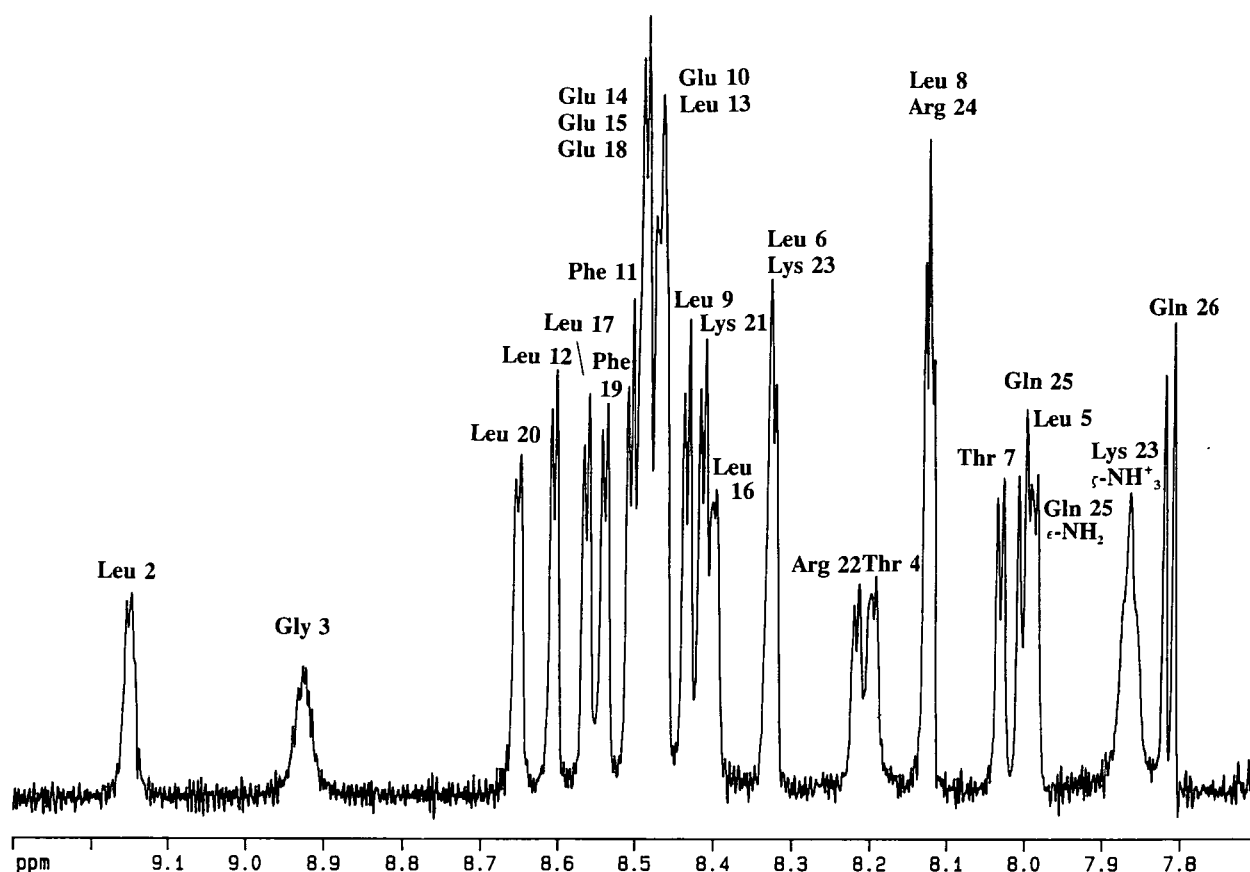


Fig. 1. ^1H NMR spectrum of 4.5 mM HELP in CD_3OH recorded at 600 MHz and at a temperature of 290 K. Prior to Fourier transformation the data were multiplied by a Gaussian window function.

leaves six amide protons which belong to the peptide's four Glus and two Glns. Four of them resonate between 8.45 p.p.m. and 8.48 p.p.m. whereas the others are located at 7.81 p.p.m. and 7.99 p.p.m. This implies that the first group comprises the four Glu-NHs and the second the NHs of the Glns. This finding is confirmed in the course of the sequence specific assignment of the amide protons.

Sequence specific assignment

Figure 3 shows NOE connectivities between the amide protons of adjacent residues. The assignment starts with the highest field shifted amide proton at 7.81 p.p.m. which has only one cross-peak with the NH at 7.99 p.p.m. As both of these are of the Gln-type they are assigned to the C-terminal Gln26 and Gln25, respectively. The chain of NH-NH-connectivities can be followed up the backbone to Phe19. Another string of NOEs starts with Leu2 (or Gly1 in the spectrum recorded at 277 K) on the N-terminal side of the peptide and extends to Leu9. Having assigned Phe19, the remaining Phe-residue must be number 11 in the chain, which shows NH-NH-NOE-cross-peaks with one Glu (which must be Glu10) and Leu12. The latter has a cross-peak with another Leu, Leu13. This leaves the 'core' of the peptide, which comprises Glu14 to Glu18, whose amide protons all resonate between 8.40 p.p.m. and 8.55 p.p.m. Cross-peaks between these resonances are best seen in the NOESY recorded in CD_3OD at 277 K. In this spectral region there are cross-peaks connecting a segment of Glu-Leu-Leu-Glu. Their sequence specific assignment remains ambiguous in the first instance, because of the inherent symmetry of this peptide. However, there is one weak cross-peak between Leu20 and a Glu-residue which is interpreted as a contact between the amide protons of Leu20

Table I. $^3J_{\text{HN-H}\alpha}$ and corresponding dihedral angles φ and Ψ for HELP in CD_3OH . $\langle\varphi\rangle_{\text{MD}}$ and $\langle\Psi\rangle_{\text{MD}}$ are the values of the dihedral angles as an average over the molecular dynamics analysis runs

Residue	$^3J_{\text{NH-C}\alpha\text{H}}$	φ_{exp}	$\langle\varphi\rangle_{\text{MD}}$	$\langle\Psi\rangle_{\text{MD}}$
Gly1	—	—	—	-141.3
Leu2	3.8	-57	-55.9	-37.9
Gly3	5.5	-70	-58.5	-51.9
	4.4	-62		
Thr4	3.9	-57	-62.5	-40.1
Leu5	4.3	-61	-59.8	-45.1
Leu6	4.0	-58	-59.0	-46.6
Thr7	4.6	-63	-64.2	-39.7
Leu8	4.4	-62	-61.9	-44.3
Leu9	4.1	-59	-60.4	-43.3
Glu10	4.1	-59	-59.6	-45.6
Phe11	4.3	-61	-60.7	-45.4
Leu12	4.0	-58	-60.4	-43.2
Leu13	4.3	-61	-60.4	-47.0
Glu14	3.9	-57	-59.8	-43.5
Glu15	4.3	-61	-60.3	-43.4
Leu16	4.2	-60	-59.6	-48.4
Leu17	4.2	-60	-59.6	-43.0
Glu18	3.9	-57	-58.0	-49.3
Phe19	4.0	-58	-61.4	-43.3
Leu20	4.0	-58	-59.9	-43.3
Lys21	3.9	-57	-60.5	-45.9
Arg22	3.9	-57	-58.2	-42.0
Lys23	4.0	-58	-60.5	-37.8
Arg24	4.3	-61	-62.8	-39.0
Gln25	6.0	-74	-74.1	-11.1
Gln26	7.1	-83	-84.0	61.4

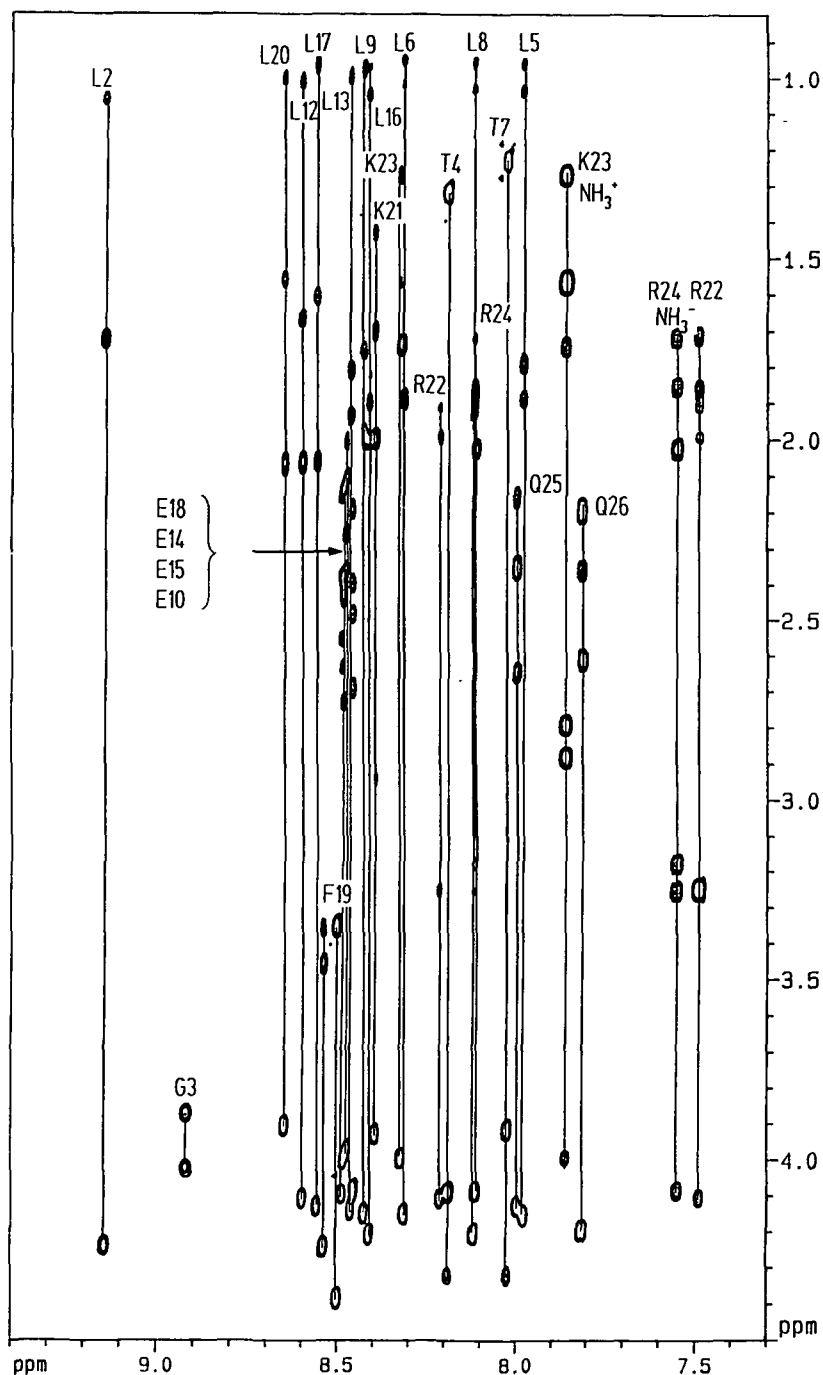


Fig. 2. NH (F_2 axis)–aliphatic (F_1 axis) region of a CLEAN TOCSY-spectrum of 4.5 mM HELP in CD_3OH recorded at 290 K with a mixing time of 100.2 ms. The spin systems are indicated by continuous lines.

and Glu18. This allows further the unambiguous assignment of the fragment mentioned above as Glu18–Leu17–Leu16–Glu15 leaving one NH unassigned which must be Glu14. This is consistent with the observation of NOEs between the C_α -proton of Phe11 and both the NH of Glu14 and Glu15 as well as with other NOEs in the fingerprint region and in other parts of the spectrum. Table II lists the assigned proton resonances of the peptide, and a summary of the observed NOEs is given in Figure 4. The predominance of leucine and glutamic acid residues in the sequence gives rise to some degeneracies in the proton chemical shifts resulting in overlap of cross-peaks. This effect is especially severe in the spectral regions where NOESY peaks between C_α -protons and C_β -protons respectively and NH-

protons of adjacent residues are observed. For the same reason a stereospecific assignment of methylene protons and methyl groups was not feasible.

Secondary structure of HELP

The NOESY spectra of HELP show all the patterns that are characteristic of an α -helical arrangement of the peptide backbone. Starting with Leu2 and extending up to Gln25, strong or medium strong NOEs corresponding to distances d_{NN} are observed between most of the adjacent amide protons, as shown in Figure 3. Some of the NOEs that are not seen in this figure, but are documented in Figure 4, are taken from NOESY spectra recorded at a lower temperature or from cross-sections through

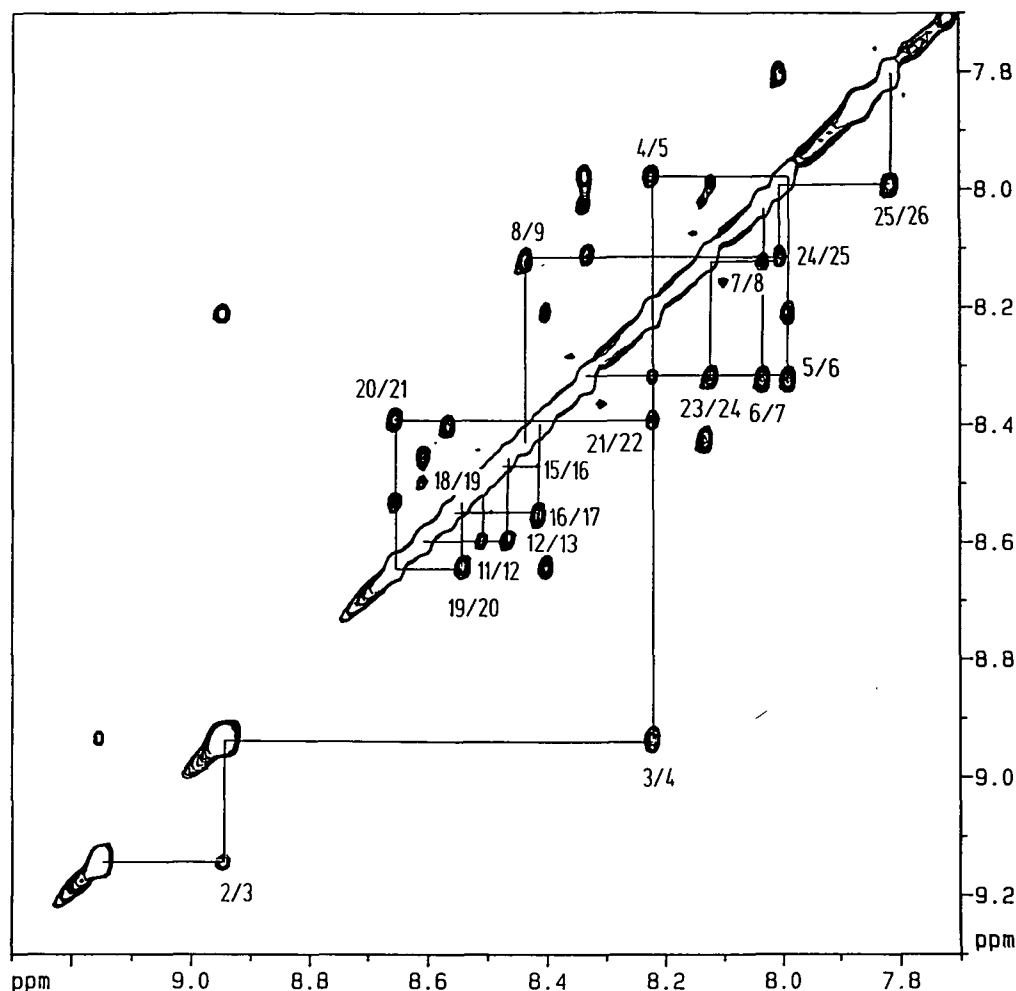


Fig. 3. NH-NH region of a NOESY-spectrum of HELP in CD₃OH at 290 K with a mixing time of 240 ms. Cross-peaks between the amide protons are labelled with the sequence positions of the residues. Sequential connectivities are indicated by continuous lines.

corresponding rows and columns in the case of cross-peaks close to the diagonal. The presence of both consecutive medium-range NOEs $d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$ throughout the sequence is another strong indication that the proposed secondary structure was accomplished. The aforementioned predominance of leucine and glutamic acid residues in the peptide explains also why the observation of medium-range NOEs $d_{\alpha\beta}(i, i+3)$ is not complete. The $^3J_{\text{HN-H}\alpha}$ coupling constants provide further evidence of a helical fold, as a series of three or more constants smaller than 6 Hz is diagnostic of that structure element (Wüthrich, 1986). With values smaller than 5 Hz, as are observed from Gly1 to Arg24, they are suggestive of a very large population of helical structures of HELP in methanolic solution. Based on similar considerations as given in Bradley *et al.* (1990) and with the same assumption of a two-state model with a helical and a random coil arrangement of the peptide backbone, we can estimate that the time-averaged helical population for residues 1–24 is more than 80%.

Molecular dynamics simulations

For the actual structure determination a total of 121 NOE-derived distances (70 of them being inter-residue ones) were incorporated in the GROMOS force field using the usual semi-harmonic potential for the distance information. In addition 25 dihedral angle restraints were included based on the calibration curve between $^3J_{\text{HN-H}\alpha}$ and the angle φ (Pardi *et al.*, 1984). Attempts

to start the calculations with an extended conformation of the peptide resulted only occasionally in reasonable final structures. A better approach using the 'random' structures created as described in Materials and methods, gave more satisfying results as *cis/trans* isomerizations of peptide bonds were less frequent.

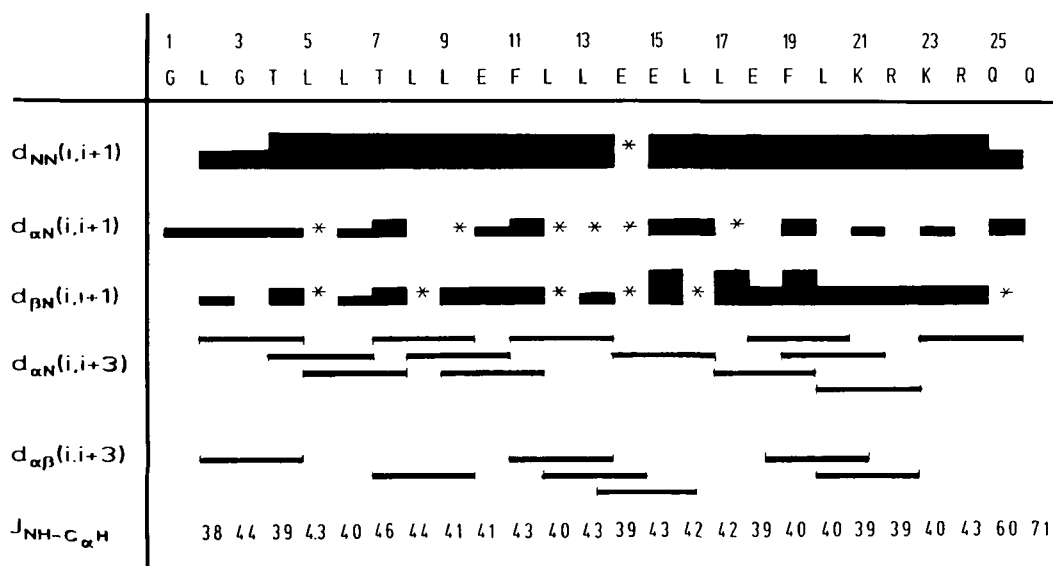
About 50 of these structures were subjected to molecular dynamics runs covering the time span between 0 ps and 23 ps. Structures which then showed more than three severe violations of the inter-residue distance restraints (>0.05 nm) were discarded as it was found that in the further course of the simulations these restraints could not be satisfied. Seven configurations with few, minor violations of the distance restraints and all showing helical structures were subjected to a final analysis. The numbers of distance violations together with the relevant energies after energy minimizing are listed in Table III. The energy values for the distance (E_{dr}) and dihedral (E_{dlr}) restraints are reasonably low, as well as the values for the total energy (E_{tot}). The only severely violated distance restraint in structures II and IV to VI is the intra-residue one between the NH of Thr7 and the β -proton of the same amino acid. A possible reason for this will be discussed below.

Hydrogen bonds within HELP

As a check of the helical folding of HELP, the formation and breaking of hydrogen bonds was monitored during the analysis of the molecular dynamics runs. Table IV lists all hydrogen bonds

Table II. ^1H Chemical shifts (in p.p.m.) of HELP in CD_3OH at $T = 290\text{ K}$ relative to DSS. The NH of Gly 1 is only observed at 277 K

Residue	NH	H_α	H_β	H_γ	Others
Gly1	8.28	4.02, 3.85			
Leu2	9.14	4.24	1.72, 1.72	1.73	$\delta\text{-CH}_3$ 1.04, 1.00
Gly3	8.91	4.01, 3.86			
Thr4	8.18	4.10	4.32	$\gamma\text{-CH}_3$ 1.31	
Leu5	7.97	4.13	1.88, 1.79	1.79	$\delta\text{-CH}_3$ 1.03, 0.97
Leu6	8.31	4.14	1.89, 1.74	1.74	$\delta\text{-CH}_3$ 1.01, 0.96
Thr7	8.02	3.91	4.32	$\gamma\text{-CH}_3$ 1.22	
Leu8	8.11	4.20	1.89, 1.84	2.00	$\delta\text{-CH}_3$ 1.02, 0.97
Leu9	8.42	4.14	1.75, 1.75	1.98	$\delta\text{-CH}_3$ 0.98, 0.98
Glu10	8.45	4.07	2.38, 2.18	2.68, 2.47	
Phe11	8.49	4.37	3.34, 3.34	Ring: 7.19 7.26 7.31	
Leu12	8.59	4.10	2.06, 2.06	1.66	$\delta\text{-CH}_3$ 1.00, 1.00
Leu13	8.45	4.12	1.92, 1.92	1.79	$\delta\text{-CH}_3$ 0.99, 0.99
Glu14	8.47	4.00	2.42, 2.11	2.71, 2.71	
Glu15	8.47	3.97	2.25, 2.00	2.09, 2.34	
Leu16	8.40	4.19	1.82, 1.89	2.00	$\delta\text{-CH}_3$ 1.04, 0.98
Leu17	8.55	4.11	1.60, 1.96	2.06	$\delta\text{-CH}_3$ 0.98, 0.96
Glu18	8.48	4.08	2.37, 2.13	2.61, 2.54	
Phe19	8.53	4.23	3.44, 3.35	Ring: 7.19 7.26 7.31	
Leu20	8.64	3.89	2.07, 1.54	2.10	$\delta\text{-CH}_3$ 1.00, 1.00
Lys	8.39	3.91	2.00, 2.00	1.42, 1.42	$\delta\text{-CH}_2$ 1.70, 1.70 $\epsilon\text{-CH}_2$ 2.94, 2.94
Arg22	8.21	4.10	2.00, 2.00	1.84, 1.71	$\delta\text{-CH}_2$ 3.25, 3.25 $\epsilon\text{-NH}$ 7.48
Lys23	8.32	3.98	1.74, 1.74	1.26, 1.26	$\delta\text{-CH}_2$ 1.55, 1.55 $\epsilon\text{-CH}_2$ 2.88, 2.79
Arg24	8.11	4.07	2.01, 2.01	1.84, 1.71	$\delta\text{-CH}_3$ 3.18, 3.24 $\epsilon\text{-NH}$ 7.54
Gln25	7.99	4.11	2.14, 2.14	2.63, 2.34	$\delta\text{-NH}_2$ 7.98, 7.04
Gln26	7.81	4.19	2.18, 2.18	2.60, 2.35	$\delta\text{-NH}_2$ 7.60, 7.66
CONH ₂	7.59, 6.90				

**Fig. 4.** Summary of the inter-residue NOEs of HELP. The heights of the bars correspond to the intensities of the NOEs. An asterisk indicates that no NOE could be observed due to overlap of the signals.

occurring in more than 10% of the total analysis time and averaged over all structures. The occupancy of more than 80% of H-bonds between the amide proton of residue i and the carbonyl oxygen of residue $i-4$ for Leu6 and from Leu8 up to Lys23 gives clear evidence for the presence of a helical secondary structure in this core region of HELP. The somewhat lower

populated hydrogen-bond at Leu5 is as expected as there will be some unravelling of the helix at the N-terminus. Thr7 deserves special attention as there are three H-bonds which are significantly populated. With an occurrence of 54.3%, the expected H-bond between the amide proton of this amino acid and the carbonyl oxygen of Gly3 is markedly underrepresented as compared to

Table III. Energies and distance restraints violations of the seven final structures (in kJ mol^{-1}) using a force constant $K_{\text{dr}} = 3000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$

	$E_{\text{tot}}^{\text{a}}$	E_{dr}^{b}	$E_{\text{dtr}}^{\text{c}}$	d > 0.02 nm
I	-2068.9	11.2	2.9	5 (0.034)
II	-2080.7	12.5	3.4	3 (0.055)
III	-1998.8	22.9	4.7	7 (0.054)
IV	-2091.8	10.5	2.7	3 (0.047)
V	-2045.6	15.8	9.2	5 (0.052)
VI	-2058.0	12.9	4.2	4 (0.056)
VII	-2031.6	7.3	2.8	2 (0.030)

^a E_{tot} , total energy.^b E_{dr} , energy associated with the distance restraints.^c E_{dtr} , energy associated with the dihedral restraints.**Table IV.** Hydrogen bonds within HELP averaged over all seven final structures as a result of the last 20 ps of the molecular dynamics runs with more than 10% occurrence

Donor	Acceptor	Occurrence (in %)
Thr4 NH	Gly1 O=C	11.4
Leu5 NH	Gly1 O=C	64.3
Leu6 NH	Leu2 O=C	85.7
Thr7 NH	Gly3 O=C	54.3
Thr7 O _γ H	Thr4 O=C	12.9
Thr7 O _γ H	Gly3 O=C	11.4
Leu8 NH	Thr4 O=C	72.9
Leu9 NH	Leu5 O=C	85.7
Glu10 NH	Leu6 O=C	84.3
Phe11 NH	Thr7 O=C	88.6
Leu12 NH	Leu8 O=C	84.3
Leu13 NH	Leu9 O=C	87.1
Glu14 NH	Glu10 O=C	92.9
Glu15 NH	Phe11 O=C	90.0
Leu16 NH	Leu13 O=C	85.7
Glu18 NH	Glu14 O=C	90.0
Glu18 O _{ε2} H	Glu14 O _{ε1}	17.1
Phe19 NH	Glu15 O=C	80.0
Leu20 NH	Leu16 O=C	87.1
Lys21 NH	Leu17 O=C	85.7
Arg22 NH	Glu18 O=C	87.1
Lys23 NH	Phe19 O=C	78.6
Lys23 NH	Leu20 O=C	12.9
Arg24 NH	Leu20 O=C	70.0
Arg24 NH	Lys21 O=C	20.0
Gln25 NH	Lys21 O=C	62.9
Gln25 NH	Arg22 O=C	17.1
Gln26 NH	Arg22 O=C	44.3
Gln26 NH	Lys23 O=C	20.0
Gln26 NH	Arg24 O=C	18.6
Gln26 N _{ε2} H	Lys23 O=C	18.6
Gln26 N _{ε2} H	Arg22 O=C	14.3
Gln26 CONH ₂	Gln25 O=C	24.3

the other (i, i-4) H-bonds. Instead one observes H-bonds between the side chain OH of Thr7 as donor and the backbone oxygen of Gly3 and Thr4, respectively, with a total occurrence of 24.3%. This coincides with the intra-residue distance restraint violation between the amide proton and the β -proton of Thr7. At the C-terminus, one observes besides the normal (i, i-4) H-bonds H-bonds between residues i and i-3 and i-2, respectively, a feature which occurs frequently at this end of a helix. There is no evidence for an unravelling of the peptide, but the C-terminus rather deviates from an ideal α -helix towards a 3_{10} -helix.

Table V. (A) RMS-deviations of the seven final structures versus an 'ideal helix' averaged over all backbone atoms (in nm).

	res. 1-26	res. 3-23	res. 5-21
I	0.094	0.074	0.058
II	0.125	0.084	0.069
III	0.144	0.085	0.076
IV	0.226	0.121	0.077
V	0.095	0.060	0.044
VI	0.137	0.101	0.070
VII	0.114	0.087	0.072
Total	0.134	0.087	0.066
	± 0.042	± 0.019	± 0.012

Table V. (B) average of the pairwise RMS-deviations among the seven final structures (in nm)

	res. 1-26	res. 3-23	res. 5-21
	0.185	0.127	0.101
	± 0.078	± 0.039	± 0.027

As a further criterion for the correct helical folding of the calculated structures, they were compared to an 'ideal' helix, which was model-built on the graphics display and then energy minimized. The root mean square (RMS)-deviations for them averaged over the atoms of the backbone are given in Table V. The mean value of 0.133 nm for all structures and summed over all residues is reasonably low, and it decreases further to 0.066 nm when averaged only over residues 5-21. This might be explained by some curvature which is common to all the NMR-derived structures of this peptide. A backbone superposition of the seven final structures is shown in Figure 5.

Discussion

The actual determination of the three-dimensional structure of the designed helical peptide HELP in methanolic solution gives us the possibility to check on the principles underlying the design, as described by Moser (1992).

Evidence for the proposed helical structure

Already the first one-dimensional ¹H-NMR-spectrum of HELP dissolved in CD₃OD showed evidence for the proposed secondary structure. The observation of J-coupling constants around 4 Hz and the fact that nearly all amide protons showed very slow exchange with the solvent protons indicated the presence of a very stable, hydrogen-bonded helical arrangement of the peptide in solution. The detailed analysis of the NMR-spectra and the results of the molecular dynamics calculations show indeed that at least residues 5-24 are arranged in an α -helical structure. So we conclude that the criteria for the selection of the amino acids in the sequence, i.e. the helical propensity and the hydrophathy, were correct.

Global fold for HELP

Inspection of the final structures shows clearly that the proposed amphiphilic arrangement of the side chains emerged as a consequence of the folding of the peptide. Indeed, the polar side chains of glutamic acid constitute the hydrophilic face, allowing at least to a certain extent the formation of a hydrogen bond between the side chain carboxylate groups of Glu15 and Glu18. As intended the apolar Leu-residues occupy the hydrophobic face. Additionally, one observes a slight bending of the helix axis towards this side suggesting some clustering of the hydrophobic

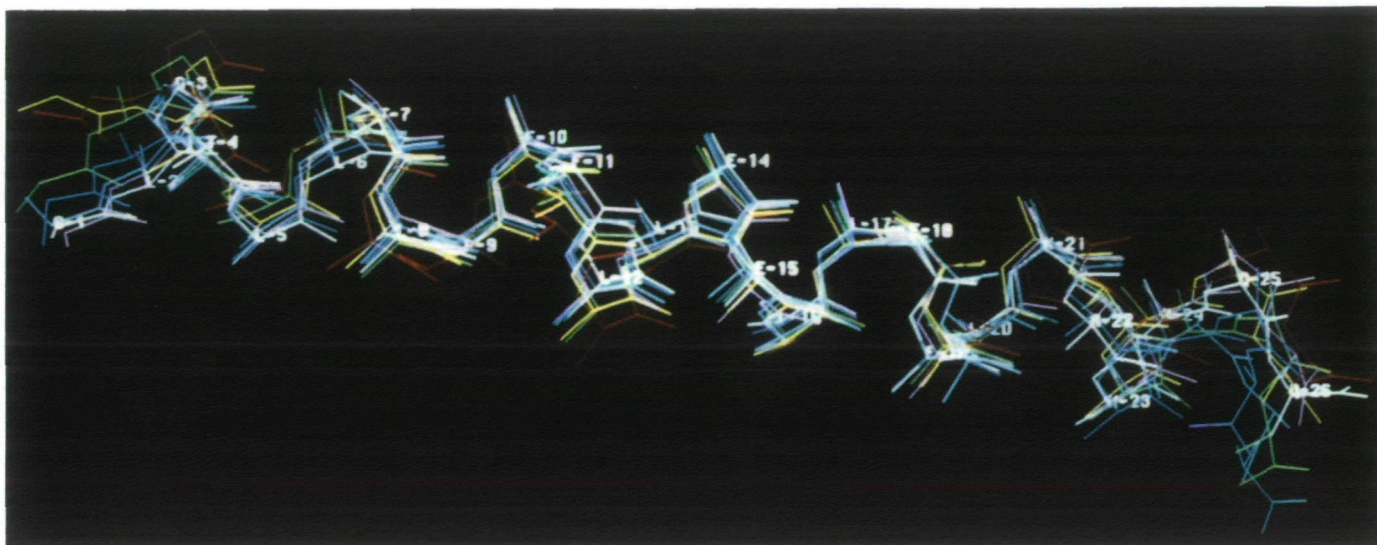


Fig. 5. Backbone representation of the seven final structures of HELP, superimposed for minimum RMS deviation between residues Gly3 and Lys23.

side chains. This may also be supported by a mutual residual repulsion of the Glu side chains.

It is clear from our results the chosen length of 26 residues is more than the minimum length for a peptide to adopt a stable secondary structure. End effects do not seem to play such an important role, especially in our design with the C-terminal α -carboxyl group blocked by an amide group. Unravelling of the helix is observed only at the N-terminus, where the first four residues are involved. Based on the pattern of hydrogen bonds the C-terminus shows a tendency towards formation of an even more compact 3_{10} -helix. This arrangement compares very well with the solution structure of the peptide toxin melittin (Bazzo *et al.*, 1988), whose sequence served as a template for HELP and from which the six C-terminal residues were taken over unchanged.

The severe distance restraint violation within Thr7, as mentioned above, can be explained in at least two ways. The first explanation would be a misinterpretation of the NOE-derived distance between the amide proton and the beta proton of this amino acid. However, this can be ruled out as three of the seven final structures do not show any violation of this restraint. The second reason could be that the force constant for the distance restraints was not high enough to counteract the energetically favourable formation of a hydrogen bond between the carbonyl oxygen of Gly3 and the side chain OH of Thr7. The latter locks these structures in a slightly distorted conformation of the side chain, for which there is no experimental evidence. This also explains the observed lowered occurrence of the H-bond between the amide proton of this residue and the carbonyl oxygen of Gly3, which one would expect for a regular α -helix, and the existence of an H-bond between the side chain oxygen of this threonine and Gly3.

The dynamic properties of HELP—as manifested in the formation and breaking of the hydrogen bonds observed during the molecular dynamics simulations—correlate very well with the estimate of the helical content as deduced from the population-weighted average of the $^3J_{\text{HN-H}\alpha}$ coupling constants. With this determination at an atomic resolution of the three-dimensional structure of a designed helical peptide, the peptide design circle is closed. Having these results at hand, one can even go further towards designing more sophisticated amphiphilic α -helices,

which could be used to tailor specific minienzymes with proposed catalytic activities. HELP with its unexpected, extraordinarily high stability when compared with other naturally occurring small peptides may also be used as a valuable α -helix model enabling the study of the initial folding of proteins and of the factors which stabilize the three-dimensional structure.

Acknowledgements

We thank Prof. Dietmar Schomburg for his support of this work and Dr Hans-Jürgen Hecht for valuable discussions. We gratefully acknowledge Dr Victor Wray for his stimulating interest in this project as well as for his careful and critical reading of the manuscript. We thank Dr Martin Zulauf for performing the light scattering experiments. This work was supported by a grant of the Swiss National Science Foundation to B.Gutte.

References

- Argos, P. and Palau, J. (1982) *Int. J. Peptide Protein Res.*, **19**, 380–393.
- Bax, A. and Davis, D.G. (1985) *J. Magn. Res.*, **65**, 355–360.
- Bazzo, R., Tappin, M.J., Pastore, A., Harvey, T.S., Carver, J.A. and Campbell, I.D. (1988) *Eur. J. Biochem.*, **173**, 139–146.
- Billeter, M., Braun, W. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 321–346.
- Bradley, E.K., Thomason, J.F., Cohen, F.E., Kosen, P.A. and Kuntz, I.D. (1990) *J. Mol. Biol.*, **215**, 607–622.
- Chou, P.Y. and Fasman, G.D. (1974) *Biochemistry*, **13**, 211–221.
- Davis, D.G. and Bax, A. (1985) *J. Am. Chem. Soc.*, **107**, 2820–2821.
- DeGrado, W.F. (1988) *Adv. Prot. Chem.*, **39**, 51–118.
- DeGrado, W.F., Wasserman, Z.R. and Lear, J.D. (1989) *Science*, **243**, 622–628.
- Greenfield, N. and Fasman, G.D. (1969) *Biochemistry*, **8**, 4108–4116.
- Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870–7872.
- Gurevich, A.Z., Barsukov, I.L., Arseniev, A.S. and Bystrov, V.F. (1984) *J. Magn. Res.*, **56**, 471–478.
- Haasnoot, C.A.G., van de Ven, F.J.M. and Hilbers, C.W. (1984) *J. Magn. Res.*, **56**, 343–349.
- Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.
- Klaus, W., Dieckmann, T., Wray, V., Schomburg, D., Wingender, E. and Mayer, H. (1991) *Biochemistry*, **30**, 6936–6942.
- Macura, S., Huang, Y., Sutter, D. and Ernst, R.R. (1981) *J. Magn. Res.*, **43**, 259–281.
- Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
- Moser, R. (1992) *Prot. Engng*, **5**, 323–331.
- Neidig, K.-P. and Kalbitzer, H.R. (1990) *J. Magn. Res.*, **88**, 155–160.
- Niiles, M., Clore, G.M. and Gronenborn, M. (1988) *FEBS Lett.*, **239**, 129–136.
- Palau, J., Argos, P. and Puigdomenech, P. (1982) *In J. Peptide Protein Res.*, **19**, 394–401.

- Pardi,A., Billeter,M. and Wüthrich,K. (1984) *J. Mol. Biol.*, **180**, 741–751.
- Provencher,S.W. and Glöckner,J. (1981) *Biochemistry*, **20**, 33–37.
- Provencher,S.W. (1982) *Comp. Phys. Commun.*, **27**, 213–242.
- Rance,M., Sorensen,O.W., Bodenhausen,G., Wagner,G., Ernst,R.R. and Wüthrich,K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 479–485.
- Redfield,A. and Kunz,S.D. (1975) *J. Magn. Res.*, **19**, 250–254.
- Schomburg,D. and Reichelt,J. (1988) *J. Mol. Graphics*, **6**, 161–165.
- Shoemaker,K.R., Kim,P.S., York,E.J., Stewart,J.M. and Baldwin,R.L. (1987) *Nature*, **326**, 563–567.
- van Gunsteren,W.F. and Berendsen,H.J.C. (1987) *GROMOS Manual*. Biomos, Groningen.
- Wright,P.E., Dyson,H.J. and Lerner,R.A. (1988) *Biochemistry*, **27**, 7167–7175.
- Wüthrich,K., Billeter,M. and Braun,W. (1983) *J. Mol. Biol.*, **169**, 949–961.
- Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York.
- Zuiderweg,E.R.P., Hallenga,K. and Olejniczak,E.T. (1986) *J. Magn. Res.*, **70**, 336–343.

Received on November 5, 1991; revised and accepted on February 25, 1992