Epidermal growth factor: exposure and dynamics of the aromatic side chains as investigated by photo-CIDNP and variable temperature ^{1}H-NMR

Antonio De Marco, Enea Menegatti* and Mario Guarneri*

Istituto di Chimica delle Macromolecole del CNR Via E. Bassini 15/A, 20133 Milano and **Istituto di Chimica Farmaceutica dell'Università di Ferrara, Via Scandiana 21, 44100 Ferrara, Italy*

Received 9 May 1983; received version received 15 June 1983

EGF is a 53 residue polypeptide of multiple biological activities. Homonuclear decoupling, spin-echo multiplet selection and Photo-CIDNP experiments lead us to fully assign the resonances from the aromatic side chains (5 Tyr, 1 His and 2 Trp). The photochemical experiment gives specific attribution of doublets in tyrosines and multiplets in tryptophans. The resonances from two tyrosines are broadened and shifted at high fields, suggesting the existence of hydrophobic domains in EGF, consistent with the presence of ring current shifted methyl resonances. However, the amide exchange in D_2O solution is considerably faster than that observed for globular proteins of the same size, and most of the aromatic residues are accessible to the flavin dye. The combined evidence suggests that EGF possesses a folded structure, but the molecule is rather floppy. From spectra measured at variable temperature a $\Delta G°$ at 25°C of about 8 kcal/mol is calculated.

Epidermal growth factor *^{1}H-NMR* *Photo-CIDNP* *Aromatic residues* *Protein dynamics*

1. INTRODUCTION

Epidermal growth factor is synthesized in the submaxillary glands of adult male mice and then secreted into the blood stream. Of the growth factors purified to date, EGF is one of the most biologically potent, both in vivo and in organ or cell cultures. Among its in vivo activities is the accelerated proliferation of skin [1] and epithelial tissues [2,3]. The increase in the active transport of nutrients is noticeable in cell cultures [4,5], as also the activation of DNA [6,7], RNA [6–8] and protein synthesis [7,8]. Although EGF is one of the best characterized proteins as to its biological and physico-chemical properties, little is known about its conformation in solution.

EGF is a single polypeptide of 53 amino acid residues, devoid of Ala, Phe and Lys [9], and linked by 3 intramolecular disulfide bonds [10,11]:

H_2N-Asn–Ser–Tyr–Pro–Gly–Cys–Pro–Ser–Ser

–^{10}Tyr–Asp–Gly–Tyr–Cys–Leu–Asn–Gly–

–Gly–Val–^{20}Cys–Met–His–Ile–Glu–Ser–

Leu–Asp–Ser–Tyr–^{30}Thr–Cys–Asn–Cys–

Val–Ile–Gly–Tyr–Ser–Gly–^{40}Asp–Arg–

Cys–Gln–Thr–Arg–Asp–Leu–Arg–Trp–

^{50}Trp–Glu–Leu–Arg-COOH

A space-filling model of EGF [12] based on CD spectra and predictive rules for secondary structure suggests that the protein possesses a highly compact structure, due mainly to the S–S bridges and to the presence of several β-bends.

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

Here we report the complete identification of the eight aromatic spin systems of EGF (5 Tyr, 2 Trp, 1 His) in the ^{1}H-NMR spectra, achieved by double resonance experiments and spin-echo multiplet selection [13]. At least 2 tyrosines appear to have reduced mobility. Variable temperature spectra up to 75°C make it possible to evaluate thermodynamic parameters for unfolding ($\Delta G°$ at 25°C ~8 kcal.mol^{-1}). Photo-CIDNP experiments show that at least 6 aromatic rings are, in varying degrees, accessible to the flavin dye. It is interesting to note that the Tyr resonances most broadened and most displaced from their random coil positions give a strong CIDNP response, suggesting that, despite the highly tertiary structure, the EGF molecule is somewhat loose. The overall NMR evidence from the present experiments suggests, in contrast with previous conclusions, that EGF may not be one of the most energetically stable proteins described to date.

2. EXPERIMENTAL

EGF receptor grade was purchased from Seragen Inc. (Boston, MA) and used without further purification. EGF solution was ~ 0.2 mM in ^2H_2O (pH 7.9). Conventional and photo-CIDNP ^{1}H-NMR spectra were recorded on a HX-270 and HX-360 Bruker spectrometer, respectively. Chemical shifts are quoted in ppm from internal sodium 3-trimethylsilyl [2,2,3,3$^2H_4$]propionate (TSP).

Although EGF was dissolved in ^2H_2O at pH 4.5, after dissolving the sample the pH was ~ 8.0, probably due to residual salts in the commercially purified protein. At the latter pH all amide protons exchanged too fast to allow their detection in the NMR spectrum. Resolution enhancement was applied via the Lorentz–Gauss transformation [14]. Spectra in fig. 1C,D were simulated with the PANIC program (Aspect 2000 Data Package). The spin-echo experiment shown in fig. 1B was performed according to the pulse sequence: (90°–t_1–180°–t_1–90°–t_2–90° acquisition), where t_1 is ½J and t_2 allows for partial relaxation [13]. Photo-CIDNP difference spectra were obtained by taking 'light' and 'dark' FIDs and subtracting the spectra after Fourier transformation. A 0.6 s light pulse was used with a 0.05 s delay before the 90° pulse.

3. RESULTS

Fig. 1A shows the aromatic region of the 270 MHz ^{1}H-NMR spectrum of mouse EGF. Since all amide protons were exchanged against deuterons before the experiment, on the basis of the amino acid composition, this part of the spectrum should contain the resonances from 1 His, 5 Tyr and 2 Trp. A spectral integration and double resonance experiments were sufficient to identify as many multiplets and spin–spin connectivities, as expected from these aromatic residues. A simplified

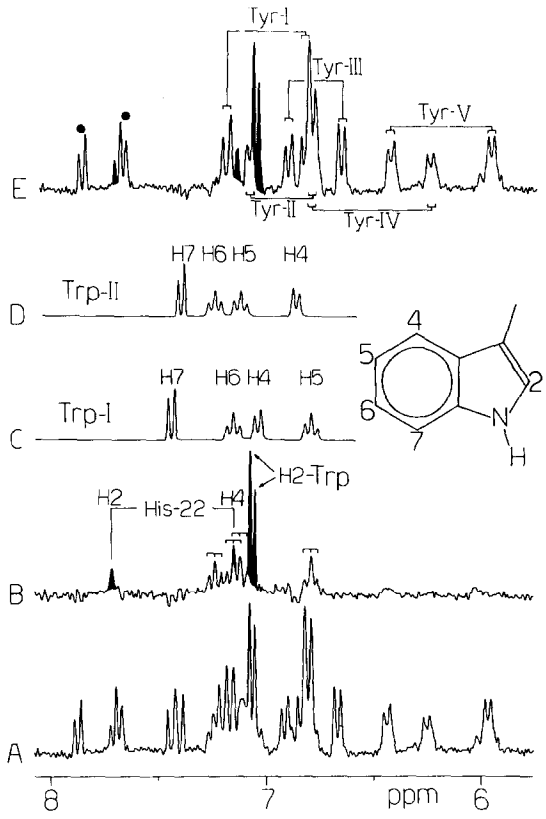

Fig. 1. (A) 270 MHz ^{1}H-NMR spectrum of 0.2 mM mouse EGF in ^2H_2O: aromatic region; (B) as (A), after the application of a pulse sequence, which essentially clears the spectrum from singlets and triplets (see section 2); (C,D), computer simulation of the separate resonances from the two Trp included in the protein sequence. (E), (A)–(C + D): besides the His and Trp singlets, this spectrum contains only the Tyr ring resonances, with spin–spin connectivities indicated. Singlet resonances have been shaded in spectra (B) and (E); (•) denotes the impurity discussed in the text; T = 25°C, pH = 7.9.

version of the crowded spectrum shown in Fig. 1A is presented in fig. 1B: a partially relaxed spin-echo sequence (see section 2) essentially eliminates the doublet resonances, leaving the singlet and triplet contribution. The 4 resonances blackened in fig. 1B were not perturbed in any of the double resonance experiments, and must hence arise from H2 and H4 of His-22 and from the two H2's of Trp-49 and Trp-50. By comparing the spectrum in fig. 1B with the double resonance experiments (not shown), the separate spin systems doublet (H4)–triplet(H5)–triplet(H6)–doublet(H7) for the two tryptophans (see diagram on the right-hand side of fig. 1C) were located in the spectrum. Their computer simulations (with the same intensities, linewidths and digital filtering as the experimental resonances) are shown in fig. 1C, D, where the two residues are indicated as Trp-I and Trp-II, not being specifically assigned to sites 49 and 50. The choice between the proposed and the reverse (H7–H6–H5–H4) multiplet sequence was made possible by the photo-CIDNP experiment discussed below. Fig. 1E was obtained by subtracting the calculated spectra C and D from spectrum A, hence it contains only the contribution from the singlets (blackened as in spectrum B) and from the tyrosyl spin systems. It is noticeable how, after subtraction of the triplets, the singlet near 7.2 ppm shows up much more clearly than in spectrum B. The two singlets at lower fields are considerably broader and reduced in intensity, when compared with the other two (sharp, intense lines near 7.0 ppm). The singlet near 7.7 ppm is readily assigned to H2 of His-22, because of its chemical shift and its exchangeability against deuterium from the solvent (fig. 3B: at 45°C, pH 7.9, the singlet has disappeared within a few hours). Since line broadening effects are quite commonly observed on His signals near the pH range of protonation–deprotonation of the imidazolyl ring, the singlet at 7.2 ppm can be confidently assigned to the H4 of His-22, and, consequently, the sharp singlets near 7.0 ppm to the H2s from Trp-49 and Trp-50. This assignment was confirmed by a spectrum recorded at pH 5.8 (not shown), in which the two singlets near 7.0 ppm remained virtually unchanged, whereas those at 7.2 and 7.7 ppm were shifted to about 7.4 and 8.5 ppm, respectively. The tyrosyl resonances were identified by homonuclear decoupling, and are indicated Tyr-I to Tyr-V in Fig. 1E. The doublets marked with the black dots at 7.7 and 7.9 ppm, coupled to each other, have been assigned to an impurity from some *p*-disubstituted benzoderivative present in the commercial sample (see below).

Fig. 2A and 2B show a reference spectrum of EGF at 360 MHz, and a photo-CIDNP difference spectrum obtained at the same frequency. Although the spectral region between 6.6 and 7.3 ppm is rather crowded, in spectrum 2B various types of aromatic resonances can be identified, with their characteristic polarization patterns [15]. The multiplet near 7.2 ppm contains two triplets (absorptive, positive in the difference spectrum), one per Trp. The two triplets are coupled to the neighbour doublets (at 7.4 ppm), which do not give any photo-CIDNP response. The assignment shown in fig. 1C and 1D, is then justified, in view of the fact that the normal trend for Trp proton resonances in the conventional light-minus-dark representation is: H2(positive)–H4(positive)–H5 (no response)–H6(positive)–H7(no response) [15].

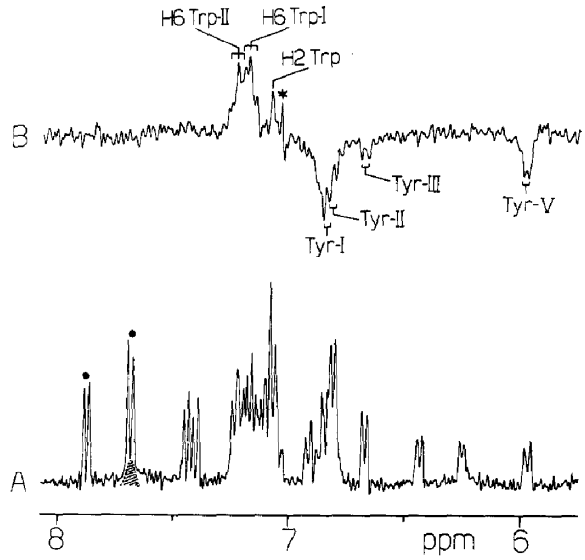

Fig. 2. (A) 360 MHz ^{1}H-NMR spectrum of 0.2 mM mouse EGF in ^2H_2O; (B) 360 MHz photo-CIDNP difference spectrum (light-minus-dark: see section 2). The dashed area near 7.7 ppm indicates the position of the broad H2 singlet from His-22; (•) represents the same impurity as in Fig. 1; (*) denotes a spectral artefact; T = 25°C, pH = 8.1, 0.5 mM flavin dye (3-*N*-carboxymethyllumiflavin).

Of the other tryptophanyl resonances expected to give signals in the photo-CIDNP difference spectrum, only one H2 singlet at 7.0 ppm is neatly visible. A spectral artefact (* in Fig. 2B) and negative interference with Tyr signals near 6.8 ppm mask the presence of the second H2 and of both the H4 doublets.

At 360 MHz, pH 8.1, the His-22 H2 singlet (dashed in spectrum 2A), is much broader than it was at 270 MHz (pH 7.9) and slightly shifted to higher fields, so that it overlaps with the impurity doublet at 7.7 ppm. Since a similar broadening is expected for H4 at the same residue, the lack of photo-CIDNP response from H2 and the uncertainty about a possible response from H4 (broad and overlapping with Trp triplets which produce signals of the same sign) do not allow the drawing of any safe conclusions about the dye exposure of the imidazolyl ring.

In a photo-CIDNP difference spectrum, the tyrosyl protons exposed to the flavin dye usually exhibit a strong emissive (negative) signal from the ϵ-resonances, and a weak response from the δ-resonances (due to cross-polarization), positive or negative, according to the local correlation time [15]. Tyr-I and Tyr-II ring doublets appear near the random coil positions [16], which makes it reasonable to assign the lower and higher field resonances to the δ and ϵ protons, respectively. This already suggests that the two aromatic side chains are exposed, and such a hypothesis is corroborated by the presence of strong, negative photo-CIDNP signals in correspondence with the ϵ resonances. Tyr-III exhibits a small but neat photo-CIDNP effect on its higher field doublet near 6.7 ppm (fig. 2B), and no effect on the doublet at lower fields. Tyr-V, although the most conformationally shifted, gives a strong emissive signal for the higher field doublet, and no appreciable signal for that at lower fields (fig. 2B). Hence, the photochemical experiment provides for both tyrosines, not only information on the side chain exposure, but also the specific assignment of the aromatic resonances. Regarding Tyr-IV, an unambiguous assignment of the two doublets was achieved on the basis of the variable temperature experiments, owing to the lack of photo-CIDNP effect on the signal at 6.25 ppm and to spectral degeneracy near 6.8 ppm.

Fig. 3 shows the aromatic spectral region of

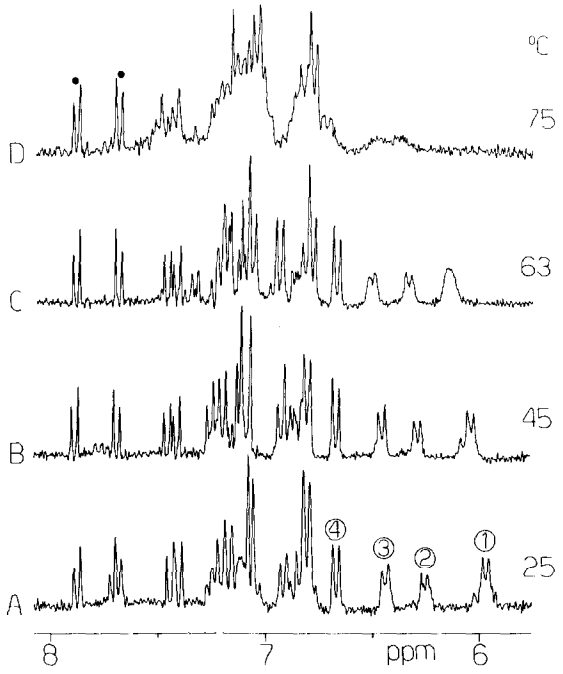

Fig. 3. (A–D) 270 MHz ^{1}H-NMR spectra of 0.2 mM mouse EGF in ^2H_2O at selected temperatures, as indicated on the right-hand side (pH = 7.9).

EGF over 25–75°C. A progressive line broadening and disappearance of native features are observed with increasing temperature. In particular, the intensity of the 4 tyrosyl doublets at highest field (labelled 1–4 in fig. 3A) decreases uniformly, whereas a concomitant growth is observed near 6.8 and 7.1 ppm; i.e., at the random coil positions for Tyr ϵ and δ protons, respectively [16].

At 63°C unfolding starts to take over significantly and the 4 doublets are variously broadened, because of exchange between native and denatured forms of EGF. Neglecting possible effects of the coupling on the lineshape, the linewidths of doublets 1–4 were evaluated by simulating the resonances in the absence of any digital filtering (fig.4A,B). The excess broadening $\Delta_{1/2}$ was then calculated by subtracting the linewidth in the absence of exchange, measured on a number of unperturbed resonances. The differences ($\Delta\nu$) between the actual positions of the doublets and the random coil shifts of δ and ϵ Tyr protons [16] were also calculated. The relation between $\Delta_{1/2}$ and $\Delta\nu^2$ can be found analytically (with some approximations) if the non-exchanging

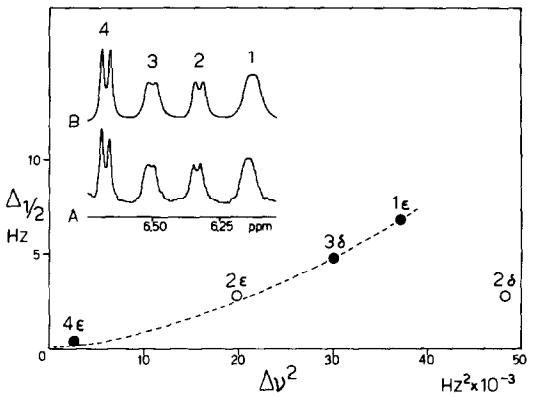

Fig. 4. $\Delta_{1/2}$ *vs* $\Delta\nu^2$ plot for tyrosyl doublets labelled 1–4 as in fig. 3A. The excess broadening ($\Delta_{1/2}$) and the relative chemical shift $\Delta\nu$ refer to the spectrum at 63°C shown in the upper insert, experimental (A) and calculated (B). For symbol explanation, see text; (---) an arbitrary interpolation between the experimental points.

linewidth is neglected [17]. Since, however, a correlation has to be expected in any case, we plotted $\Delta_{1/2}$ *vs* $\Delta\nu^2$ for doublets 1, 3 and 4 (fig. 4 (●)). For doublet 2 we calculated $\Delta\nu^2$ according to both possible assignments (δ or ϵ (○)). It is clear that only the assignment of doublet 2 to Tyr resonances of ϵ-type fits the trend of doublets 1, 3 and 4, indicated by the dotted line.

The complete lack of photo-CIDNP effect on doublet 2 in the experiment shown in fig. 2B can hence be interpreted as due to non-exposure of the Tyr-IV side chain to the flavin probe, and the effects detected near 6.8 ppm as arising only from the ϵ protons of Tyr-I and Tyr-II.

The two doublets at 7.7 and 7.9 ppm (fig.3D (●)) remain unchanged throughout the temperature range explored and were hence assigned to an impurity (see above). This was confirmed by the disappearance of the two doublets in spectra of more purified samples, which became later available, as will be shown in a following paper.

By using a set of invariant signals as reference, the ratios between denatured and native species $K_d = d/n$ were calculated at 7 temperatures (over 25–75°C) from the averaged integrated intensity of the 1–4 Tyr doublets. A linear fitting was then performed, according to the Van't Hoff's relationship:

$$-R \ln K_d = \Delta G°/T = \Delta H°/T - \Delta S°$$

and gave $\Delta S° = 0.17$ u.e., $\Delta H° = 59$ kcal/mol, with a correlation coefficient $r^2 = 0.99$. This corresponds to a denaturation temperature $T_d = \Delta H°/\Delta S° = 73$°C. A $\Delta G°$ (25°C) = 8.2 kcal/mol was also extrapolated from the fitting.

4. CONCLUSIONS

By looking at the chemical shifts and linewidths in the spectrum of fig. 1A, it can already be inferred that 2 out of 5 tyrosines in EGF (Tyr-IV and Tyr-V) are in a highly anisotropic environment and probably restricted in mobility. A third tyrosine (Tyr-III) appears to be in an intermediate state, whereas the remaining 2 (Tyr-I and Tyr-II) are near the random coil positions predicted for this type of residue in a fully extended conformation [16]. With the exception of Tyr-V, the photo-CIDNP response follows the same trend quite closely: the more a residue is displaced from an isotropic environment, the less it is exposed to the flavin dye used to generate spin polarization. Tyr-IV is completely unaccessible to the dye; Tyr-III gives a weak response, which becomes more pronounced for Tyr-II and strongly enhanced for Tyr-I (fig.3). The situation of Tyr-V is rather peculiar: although its resonances are the most shifted from their free side chain positions (0.8 and 0.9 ppm for the δ and ϵ protons, respectively), the ϵ doublet at 5.9 ppm gives a strong emissive signal. Since aromatic/methyl interactions in EGF are monitored by the presence of numerous CH_3 resonances below 0.8 ppm (not shown), it seems resonable to deduce the presence of hydrophobic domains surrounding Tyr-IV and Tyr-V. Whilst the environment of the latter aromatic ring would appear to be loose, the lack of CIDNP effect from Tyr-IV could be due to a closer contact between the aromatic side chain and its neighbour groups: an H-bond involving the Tyr-OH would, for instance, make the oxydryl proton unavailable to the flavin dye. However, as far as thermal decay of the native structure is concerned, Tyr-III, Tyr-IV and Tyr-V resonances all follow the same trend, suggesting that EGF unfolds as a whole.

The alleged high conformational stability of EGF was mainly judged from unfolding experiments in the presence of denaturants, such as GdmCl at various concentrations [8]. In [8], $\Delta G°$ (25°C) = 16 ± 7 kcal/mol was extrapolated in the

absence of denaturant. A CD-monitored study of thermal stability [8] was also performed in the presence of GdmCl at two concentrations. From the reported plots, denaturation temperatures of 40–45°C can be estimated. In these conditions, the unfolding appears to be completely reversible. The present data do not fully match those from the CD studies: the value of $\Delta G°$ at 25°C (8.2 kcal/mol) quoted here is at the lowest limit of 16 ± 7 kcal/mol [8]. Moreover, here, T_d in the absence of denaturant increases to 70°C, but the refolding observed is not complete, as inferred from the appearance of the ^{1}H-NMR spectrum after re-cooling the sample (not shown).

In conclusion, from the present NMR evidence, EGF appears to be highly folded in some regions. However, the very fast NH → ND exchange at pH 7.9 suggests that the conformation is not quite compact: in the basic pancreatic trypsin inhibitor (BPTI) at pH 8, after 3 h at 45°C, at least 6 NH proton resonances are still present in the spectrum at almost full intensity, and some are still visible after keeping the protein in solution for > 10 days, at the same pH and temperature [18]. Accordingly, the photo-CIDNP experiment shows that only one out of 5 tyrosines is completely unaccessible to the flavin dye. The two neighbour tryptophans at sites 49 and 50 are both exposed and are in similar isotropic environments. This suggests that all or part of the protein tail, including the residues 43–53 (after the last disulfide bond), is highly extended.

ACKNOWLEDGEMENTS

This work was supported by the 'Progetto Finalizzato Chimica Fine e Secondaria' of the Italian CNR. Photo-CIDNP measurements were by courtesy of Professor R. Kaptein and his collaborators.

REFERENCES

[1] Birnbaum, J.E., Sapp, T.M. and Moore, J.B. (1976) J. Invest. Dermatol. 66, 313–318.
[2] Ho, P.C., Davis, W.H. and Elliott, J.H. (1974) Invest. Ophthalmol. 13, 804–809.
[3] Sundell, H., Serenius, E.S., Barthe, P., Friedman, J., Kanarek, K.S., Escobedo, M.B., Orth, D.N. and Stahlman, M.T. (1975) Pediatr. Res. 9, 371.
[4] Hollenberg, M.D. and Cuatrecasas, P. (1975) J. Biol. Chem. 250, 3845–3853.
[5] Barnes, D. and Colowick, S.P. (1976) J. Cell. Physiol. 89, 633–640.
[6] Hollenberg, M.D. and Cuatrecasas, P. (1973) Proc. Natl. Acad. Sci. USA 70, 2964–2968.
[7] Covelli, I., Mozzi, R., Rossi, R. and Frati, L. (1972) Hormones 3, 183–191.
[8] Cohen, S. and Stastny, M. (1968) Biochim. Biophys. Acta 166, 427–437.
[9] Taylor, J.M., Mitchell, W.M. and Cohen, S. (1972) J. Biol. Chem. 247, 5928–5934.
[10] Savage, C.R. jr., Inagami, T. and Cohen, S. (1972) J. Biol. Chem. 247, 7612–7621.
[11] Savage, C.R. jr., Hash, J.H. and Cohen, S. (1973) J. Biol. Chem. 248, 7669–7672.
[12] Holladay, L.A., Savage, C.R. jr., Cohen, S. and Puett, D. (1976) Biochemistry 15, 2624–2633.
[13] Lecomte, J.T.J., De Marco, A. and Llinas, M. (1982) Biochim. Biophys. Acta 703, 223–230.
[14] Ferrige, A.G. and Lindon, J.C. (1978) J. Magn. Res. 31, 337–340.
[15] Kaptein, R. (1982) in: Biological Magnetic Resonance (Berliner, L.J. and Reuben, J. eds) vol. 4, p. 145, Plenum, New York.
[16] Bundi, A. and Wüthrich, K. (1979) Biopolymers 18, 285–297.
[17] Anet, F.A.L. and Basus, V.J. (1978) J. Mag. Res. 32, 339–343.
[18] Richarz, R. (1975) Diploma Thesis, ETH, Zürich.