Volume 238, number 2, 289–294 FEB 06334 October 1988

A proton nuclear magnetic resonance study of the conformation of bovine anaphylatoxin C5a in solution

Jutta Zarbock[+], Renato Gennaro, Domenico Romeo, G. Marius Clore[+°] and Angela M. Gronenborn[+°]

[+] *Max-Planck Institut für Biochemie, D-8033 Martinsried bei München, FRG, Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Universita di Trieste, Via A. Valerio 32, I-34127 Trieste, Italy* and [°] *Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Disorders, National Institutes of Health, Bethesda, MD 20892, USA*

Received 5 August 1988

The solution conformation of bovine anaphylatoxin C5a has been investigated by nuclear magnetic resonance (NMR) spectroscopy. The ^{1}H-NMR spectrum is assigned in a sequential manner using a variety of two-dimensional NMR techniques. A qualitative interpretation of the short range nuclear Overhauser enhancement data involving the NH, $C^{\alpha}H$ and $C^{\beta}H$ protons suggests that C5a has four helices comprising residues 5–11, 15–25, 33–39 and 46–61, and is composed of a globular head (residues 5–61) and a C-terminal tail. The polypeptide fold was determined by hybrid distance geometry-dynamical simulated annealing calculations on the basis of 203 approximate interproton distance restraints, 22 distance restraints for 11 intrahelical hydrogen bonds (identified on the basis of the pattern of short range NOEs and slowly exchanging backbone amide protons) and restraints for the 3 disulfide bridges. The overall polypeptide fold is similar to that of the sequence related human recombinant anaphylatoxin C5a [(1988) Proteins 3, 139–145].

C5a; NMR; NOE; Interproton distance; Secondary structure; Polypeptide fold

1. INTRODUCTION

The anaphylatoxin C5a is a small protein of 74 residues which is released during the activation of the complement system by the action of a specific convertase on the parent complement protein C5 [1]. Like the two related anaphylatoxins C3a and C4a, C5a plays an important role in the inflammatory process inducing multiple cellular effects, and may also be involved in the pathogenesis of a number of inflammatory diseases [1]. The amino acid sequences of C5a from a variety of species have been determined [2–5] and circular dichroism studies have suggested a helical content of 40–50% [1,5]. Very recently, the solution structure of human recombinant C5a has been investigated by NMR [6,7] and shown to be similar to that of the crystal structure of C3a [8]. The structures differed, however, at the N-terminus which was helical in solution but disordered in the crystal. In the present paper we have examined the solution conformation of bovine C5a by NMR. The sequence homology between the bovine and human forms of C5a is 68% [5]. We show that the secondary structure and overall polypeptide fold are similar to that of human recombinant C5a.

Correspondence address: G.M. Clore and A.M. Gronenborn, Laboratory of Chemical Physics, Building 2, Room 123, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; HOHAHA, two-dimensional homonuclear Hartmann-Hahn spectroscopy

2. EXPERIMENTAL

2.1. *Purification and biological activity of bovine C5a*

C5a was purified from 48 litres of yeast activated bovine

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

serum (~3.9 kg protein). The initial purification steps up to the ammonium sulfate precipitation steps were as described in [5]. After dialysis, the precipitate (42 g of protein) was loaded in 12 batches on an immunoaffinity column (60 ml), prepared by coupling the total protein A-purified IgG fraction of rabbit anti-C5a serum [5] to Affi-Gel 10 (Bio-Rad). After extensive washing with 10 mM sodium phosphate (pH 7.4)/0.5 M NaCl/0.1% NP-40, the protein bound to the affinity column was eluted with 0.1% trifluoroacetic acid (total yield 177 mg). This was followed by reverse-phase chromatography through a PepRPC HR 10/10 column (Pharmacia) eluted with a 0–100% gradient of acetonitrile in 0.1% trifluoroacetic acid. This purification procedure yielded 58 mg of C5a, which was homogeneous as shown by SDS and urea polyacrylamide gel electrophoresis. This represents a recovery of 40–45% of C5a, as calculated from the molecular mass and serum concentration of the fifth complement component. The purified C5a exhibited maximal stimulation of neutrophil chemotaxis and superoxide production at concentrations of 10^{-9} and 10^{-7} M, respectively [9].

2.2. *NMR spectroscopy*

All NMR spectra were recorded on a Bruker AM-500 spectrometer. The samples for NMR contained 6 mM C5a in either 90% H_2O/10% D_2O or 99.996% D_2O, pH 4.6. All two-dimensional spectra were recorded in the pure-phase absorption mode using the time proportional incrementation method [10,11]. The following spectra were recorded both in H_2O and D_2O: NOESY [12] with mixing times of 100, 200 and 300 ms, and HOHAHA [13,14] with mixing times ranging from 20 to 60 ms. For measurements in H_2O, the water resonance was suppressed by replacing the last 90° pulse in the NOESY sequence by a semi-selective jump-return (90°_x–τ–90°_{-x}) pulse [15] and by adding the sequence 90°_x–H–Δ–90°_x–τ–90°_{-x} after the MLEV17$_y$ sequence in the case of the HOHAHA sequence [16] (where H is a homospoil pulse of 3 ms, Δ a 3 ms delay to allow for recovery from the homospoil, and τ a 80 μs delay).

2.3. *Calculations*

Structure calculations on the basis of NOE derived interproton distance data were carried out using the hybrid distance geometry-dynamical simulated annealing approach [17], which makes use of the programs DISGEO [18] and XPLOR [19,20]. Computations were carried out on a CONVEX C1-XP computer.

3. RESULTS AND DISCUSSION

The ^{1}H-NMR spectrum of C5a was assigned in a sequential manner [21] using HOHAHA spectra recorded at several mixing times ranging from 20 to 60 ms to demonstrate direct, single and relayed through-bond connectivities and NOESY spectra to delineate through-space (<5 Å) connectivities. The most useful NOEs in this respect involve the NH, C^αH and C^βH protons and are of the type NH(i)-NH($i+1$), C^αH(i)/C^βH(i)-NH($i+1$), C^αH(i)-NH($i+3$)/C^βH($i+3$). In order to resolve ambiguities arising from cross-peak overlaps, we found it essential to record spectra at several temperatures, in particular 15°C, 25°C and 35°C. Examples of NOESY spectra in H_2O displaying the NH-NH and NH-C^αH/C^βH region of the spectrum are shown in fig.1. Sequential connectivities extended from residue 5 to residue 66. No sequential NOEs could be detected for the first 4 N-terminal residues and the 8 C-terminal residues, presumably due to a higher degree of mobility.

A summary of NOEs between residues less than five apart in the sequence involving the NH, C^αH and C^βH protons is shown in fig.2. The secondary structure of C5a can be deduced on the basis of a qualitative interpretation of this data [21,22]. In particular, the stretches of NH(i)-NH($i+1$) NOEs together with the presence of C^αH(i)-NH($i+3,4$) and C^αH(i)-C^βH($i+3$) NOEs and slowly exchanging backbone amide protons reveal the presence of four helices comprising residues 5–11, 15–25, 33–39 and 46–61. This corresponds to a helical content of 55% which is consistent with the results from circular dichroism measurements [1,5].

To obtain the approximate polypeptide fold, a set of 203 interproton distance restraints, comprising 134 sequential ($|i-j| = 1$), 44 medium range ($1 < |i-j| \leq 5$) and 25 long range ($|i-j| > 5$) distances were derived from the NOESY spectra recorded with mixing times of 100 ms and 200 ms at 15°C, 25°C and 35°C. These were classified into three distance ranges, 1.8–2.7 Å, 1.8–3.3 Å and 1.8–5.0 Å, corresponding to strong, medium and weak NOEs, respectively [23,24]. The upper limits of distances involving methyl and methylene protons were corrected for centre $\langle r_c \rangle$ averaging (which is equivalent to the use of pseudo-atoms) as described in [25]. The interproton distance restraints were supplemented by two additional sets of distance restraints: (i) 22 restraints for the 11 intrahelical NH($i+4$)-O(i) hydrogen bonds identified on the basis of slowly exchanging amide protons and the presence of C^αH(i)-NH($i+3,4$) and/or C^αH(i)-C^βH($i+3$) NOEs [26] (for each hydrogen bond the N($i+4$)-O(i) and NH($i+4$)-O(i) distances were constrained to <3.3 Å and <2.3 Å, respectively); and restraints for the three disulfide bonds between Cys-21 and Cys-47, Cys-22 and Cys-54, and Cys-34 and Cys-55 [8,27]. In addition, weak restraints in the form of square-well potentials [24] were imposed on selected ϕ and ψ

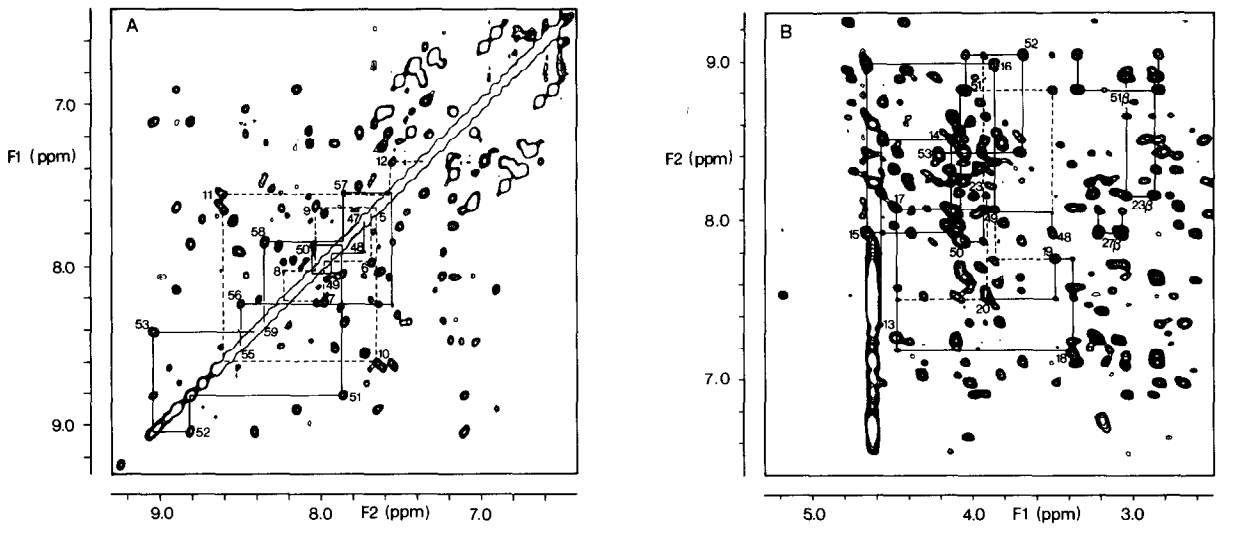

Fig.1. (A) NH(F1 axis)-NH(F2 axis) and (B) NH(F2)-$C^{\alpha}H/C^{\beta}H$ (F1 axis) regions of the 200 ms NOESY spectrum of C5a in 90% H_2O/10% D_2O at 35°C. In (A) stretches of NH(i)-NH($i+1$) connectivities for helices 1 and 4 are indicated by dashed (– – –) and continuous (——) lines, respectively, with the labelling of the cross-peaks at the F1 frequency of the corresponding NH proton. In (B), some $C^{\alpha}H(i)$-NH($i+1$) and $C^{\alpha}H(i)$-NH($i+3$) NOE connectivities are indicated by continuous (——) and dashed (– – –) lines, respectively. Also indicated in (B) are a few $C^{\beta}H(i)$-NH($i+1$) connectivities. The labelling in (B) is at the $C^{\alpha}H(i)$-NH(i) and $C^{\beta}H(i)$-NH(i) intraresidues cross-peaks, the latter indicated by the residue number followed by the letter β.

backbone torsion angles as follows. First all ϕ backbone torsion angles outside the helical regions, with the exception of those for the Gly and Asn residues, were restrained between $-10°$ and $-180°$ on the basis of a survey of high resolution X-ray structures in the Brookhaven protein data bank. (Note that the probability for ϕ values outside this range is <2%; [28].) Second, within the helical regions, the ϕ and ψ angles of residues 5–11, 17–24, 33–39 and 47–60 were restrained to the ranges listed in table 1 using the method of Sherman et al. [28] to determine the most probable values of the ϕ,ψ angles for a given set of observed NH(i)-NH($i+1$), $C^{\alpha}H(i)$-NH($i+1$) and $C^{\beta}H(i)$-NH($i+1$) NOEs. The introduction of these ϕ,ψ restraints has no effect on the determination of the overall polypeptide fold but helps circumvent problems associated with local mirror images.

Five structures which satisfied the experimental restraints, displayed very small deviations from idealized covalent geometry, and had good non-bonded contacts, were computed using the hybrid distance geometry-dynamical simulated annealing approach described in [17]. The orientation of the four helices with respect to each other was well defined by the data. The details of the local backbone structure as well as the side chain conformations could not be determined on the basis of the present data. This will require further experimental work to obtain a much larger number of interproton distance restraints as well as stereospecific assignments of the β-methylene protons. Consequently, only a schematic representation of the polypeptide fold is shown in fig.3.

The structure is dominated by a globular head (residues 5–61) and a tail (residues 62–74). The main stabilizing factors within the globular head are the disulfide bridges between helices 2 and 3 on the one hand and helix 4 on the other, supplemented by a number of hydrophobic interactions. In this respect we note that of the 7 aromatic residues, only two, Tyr 13 and Phe 51, are partially buried within the globular core. This is consistent with the observation that, of the two tyrosine residues present in C5a, only Tyr 23 is susceptible to chemical modification [29]. Helices 1 and 3 are approximately antiparallel to helix 4, while helix 2 is at an angle of ~70° to helix 4. The orientation of helix 1 with respect to helix 4 is determined by NOEs from Tyr-13 to Ala-50, Phe-51 and Ser-53, and from Ile-6 to Ala-58.

The polypeptide fold of bovine C5a is similar to

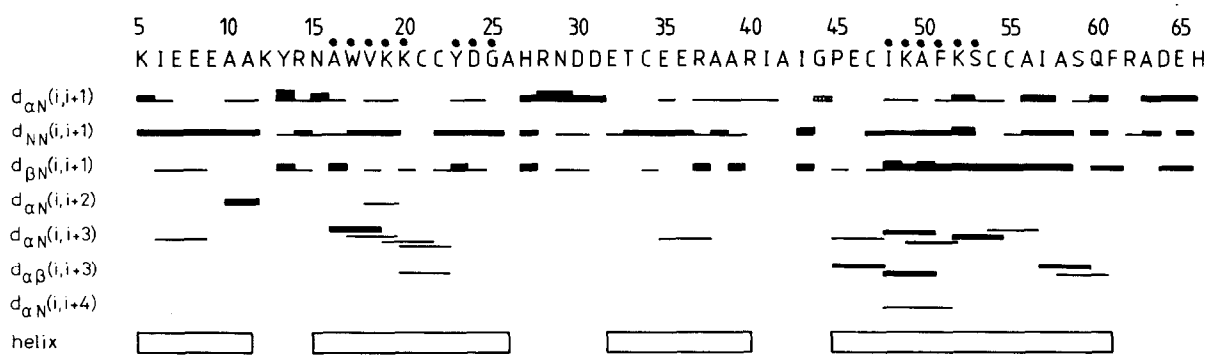

Fig.2. Summary of NOEs involving the NH, $C^{\alpha}H$ and $C^{\beta}H$ protons of residues less than five apart in the sequence observed for C5a together with the secondary structure derived from them. The closed circles indicate slowly exchanging (>24 h) backbone amide protons. The $C^{\alpha}H(i)$-$C^{\delta}H(i+1)$ NOE between Gly-44 and Pro-45 is indicated by a hatched box along the same line as the $C^{\alpha}H(i)$-$NH(i+1)$ NOEs.

that of both human recombinant C5a [7] and C3a [8]. There are a number of differences though between the solution conformation of human and bovine C5a on the one hand and the crystal structure of C3a on the other. Helix 4 extends way beyond the globular head in the crystal structure of C3a and is 10 residues longer than that of C5a. It appears that this helix is stabilized by packing interactions since C3a is a dimer in the crystal with the interface formed by the C-terminal helices arranged in an anti-parallel manner [8]. The first 12 residues of C3a were not visible in the electron density map, while in the case of C5a there is a N-terminal helix (residues 5–11). The location of this helix with respect to the rest of the protein is in agreement with the prediction of Greer based on model building studies [30]. Interestingly, although the polypeptide fold of C3a in solution has not yet been determined, its secondary structure is almost identical to that of C5a, particularly with regard to the presence of helix 1 and the length of helix 4 [31]. It would thus appear that the solution structures of the anaphylatoxins all display the same differences with respect to the crystal structure of C3a.

Table 1

ϕ and ψ restraints used within the helical regions derived on the basis of the pattern of observed $NH(i)$-$NH(i+1)$, $C^{\alpha}H(i)$-$NH(i+1)$ and $C^{\beta}H(i)$-$NH(i+1)$ NOEs[a]

Class	NOEs			ϕ (°)	ψ (°)
	$NH(i)$-$NH(i+1)$	$C^{\beta}H(i)$-$NH(i+1)$	$C^{\alpha}H(i)$-$NH(i+1)$		
Q	+			−85 ± 60	−5 ± 60
QR	+				
or	+	+		−70 ± 70	−20 ± 90
R	+	+		−65 ± 60	−35 ± 90
K	+	+	+	−75 ± 60	−10 ± 60
TK	+		+		
or	+	+	+	−83 ± 80	5 ± 130

[a] The classes Q, R, T and K are those of Sherman et al. [18]. The classes QR and TK are combined classes where the absence or presence of a particular NOE cannot be assessed due to spectral overlap. The values of the ranges of classes Q, R and K are approximately three times the values of the standard deviations from the most probable ϕ or ψ value given by Sherman et al. [28]

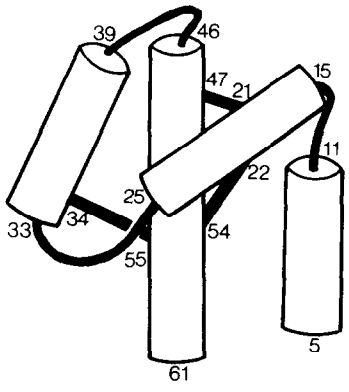

Fig.3. Schematic representation of the polypeptide fold of C5a as derived from the NOE distance data and hybrid distance geometry-dynamical simulated annealing calculations. Helices are represented as cylinders and the other residues as thick lines. The disulfide bridges are indicated by straight lines.

Alignment of the primary sequences of C3a, C4a and C5a shows that, in addition to the six Cys residues, there is total conservation of the following residues [32]: Tyr-13(15), Gly-25(26), Arg-37(39), Phe-51(53) and Arg-62(64) (numbers in parentheses refer to C3a and C4a). Do these residues have a common role in the structures of C5a and C3a? Tyr-13 and Phe-51 are buried in C5a and are probably involved in the stabilization of the globular head. Phe-53 in C3a is likewise buried but Tyr-15 is exposed. This, however, is because the N-terminus is not folded in the crystal structure [8]. Gly-25 in C5a and Gly-26 in C3a both occur at the end of helices and may therefore act as helix breakers. Arg-37(39) is directed towards the interior and may be involved in stabilizing electrostatic interactions in both C5a and C3a. Arg-62 is one residue removed from the C-terminus of helix 4 and may form part of the active site (see below).

Despite the overall structural similarity, C5a and C3a display distinct differences with respect to structure-function relationships [1]. In the case of C3a, all biological actions are mimicked by the C-terminal synthetic octapeptide 70–77 [33]. In contrast, analogous C-terminal synthetic peptides of C5a exhibit no activity [34]. While the C-terminal Arg-74 appears to be essential for spasmogenic activity, it is not required for chemotactic activity and lysosomal release which can still be induced by C5a desArg-74 [34]. Further, the C5a(1–69) fragment is sufficient for receptor interaction although it does not stimulate a cellular response [5,34]. At the same time, it appears that the N-terminus may also play a role in C5a activity as the C5a(18–74) fragment is active, albeit a factor of ~1000 less than C5a desArg-74 [35]. This is supported by the close proximity of the C-terminal end of helix 4 and the N-terminal end of helix 1. Tyr-23 which is located on the exposed surface of helix 2, also appears to participate in ligand-receptor interactions [29]. Indeed, photoreactive derivatives of C5a, prepared by specifically modifying Tyr-23, fail to interact with the granulocyte C5a receptor, while retaining their binding to a specific anti-C5a monoclonal antibody [29]. This suggests that there are two areas on C5a involved in the interaction with the cellular C5a receptor(s). The first, like that of C3a, probably comprises the C-terminal tail, while the second may be made up of the contiguous polar surface formed by helices 1 and 2 comprising Glu-7, Lys-20, Tyr-23 and His-27, possibly supplemented by Arg-62.

Acknowledgements: This work was supported by the Max-Planck-Gesellschaft, grant no.321/4003/0318909A from the Bundesministerium für Forschung und Technologie (G.M.C. and A.M.G.) and by the Italian Ministry of Education and grant no.86.01613.52 from the Italian National Research Council, Progetto finalizzato 'Controllo delle malattie da infezione' (D.R. and R.G.).

REFERENCES

[1] Hugli, T.E. (1981) Crit. Rev. Immunol. 1, 321–366.
[2] Fernandez, H.N. and Hugli, T.E. (1978) J. Biol. Chem. 253, 6955–6964.
[3] Gerard, C. and Hugli, T.E. (1987) J. Biol. Chem. 255, 4710–4715.
[4] Cui, L.-X., Ferrei, K. and Hugli, T.E. (1985) Complement 2, 18–19.
[5] Gennaro, R., Simonic, T., Negri, A., Mottola, C., Secchi, C., Ronchi, S. and Romeo, D. (1986) Eur. J. Biochem. 155, 77–86.
[6] Zuiderweg, E.R.P., Mollison, K.W., Henkin, J. and Carter, G.W. (1988) Biochemistry 27, 3568–3580.
[7] Zuiderweg, E.R.P., Henkin, J., Mollison, K.W., Carter, G.W. and Greer, J. (1988) Proteins 3, 139–145.
[8] Huber, R., Scholze, H., Paques, E.P. and Deisenhofer, J. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 1389–1399.
[9] Gennaro, R., Pozzan, T. and Romeo, T. (1984) Proc. Natl. Acad. Sci. USA 81, 1416–1420.
[12] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) J. Chem. Phys. 71, 4546–4553.
[13] Braunschweiler, L. and Ernst, R.R. (1983) J. Magn. Reson. 53, 521–558.
[14] Davis, D.G. and Bax, A. (1985) J. Am. Chem. Soc. 107, 2821–2822.
[15] Plateau, P. and Gueron, M. (1982) J. Am. Chem. Soc. 104, 7310–7311.
[16] Bax, A., Sklenar, V., Clore, G.M. and Gronenborn, A.M. (1987) J. Am. Chem. Soc. 109, 6511–6513.
[17] Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) FEBS Lett. 229, 317–324.
[18] Havel, T.F. (1986) DISGEO, Quantum Chemistry Program Exchange, Program No.507, Indiana University.
[19] Brünger, A.T., Clore, G.M., Gronenborn, A.M. and Karplus, M. (1986) Proc. Natl. Acad. Sci. USA 83, 3201–3205.
[20] Brünger, A.T., Kuryan, J. and Karplus, M. (1987) Science 235, 458–460.
[21] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[22] Wagner, G., Neuhaus, D., Wörgötter, E., Vasak, M., Kägi, J.H.R. and Wüthrich, K. (1986) J. Mol. Biol. 187, 131–135.

[23] Williamson, M.P., Havel, T.F. and Wüthrich, K. (1985) J. Mol. Biol. 182, 295–315.
[24] Clore, G.M., Nilges, M., Sukumaran, D.K., Brünger, A.T., Karplus, M. and Gronenborn, A.M. (1986) EMBO J. 5, 2729–2735.
[25] Wüthrich, K., Billeter, M. and Braun, W. (1983) J. Mol. Biol. 169, 949–961.
[26] Wagner, G., Braun, W., Havel, T.F., Schauman, T., Go, N. and Wüthrich, K. (1987) J. Mol. Biol. 196, 611–640.
[27] Zimmermann, B. and Vogt, W. (1984) Hoppe Seyler's Z. Physiol. Chem. 365, 151–158.
[28] Sherman, S.A., Andrianov, A.M. and Akhrem, A.A. (1987) J. Biomolec. Struct. Dyn. 4, 869–884.
[29] Johnson, R.J. and Chenoweth, D.E. (1985) J. Biol. Chem. 260, 10339–10345.
[30] Greer, J. (1985) Science 228, 1055–1060.
[31] Nettesheim, D.G., Edalji, R.P., Mollison, K.W., Greer, J. and Zuiderweg, E.R.P. (1988) Proc. Natl. Acad. Sci. USA 85, 5036–5040.
[32] Greer, J. (1986) Enzyme 36, 150–163.
[33] Hugli, T.E. and Erickson, B.W. (1977) Proc. Natl. Acad. Sci. USA 74, 1826–1830.
[34] Chenoweth, D.E. and Hugli, T.E. (1980) Mol. Immunol. 17, 151–161.
[35] Gerard, C., Showell, H.J., Hoeprich, P.D., Hugli, T.E. and Stimler, N.P. (1985) J. Biol. Chem. 260, 2613–2616.