500 MHz ^{1}H NMR STUDY OF THE ROLE OF LYSINES AND ARGININES IN THE BINDING OF GENE-5 PROTEIN TO OLIGOADENYLIC ACIDS

N. C. M. ALMA, B. J. M. HARMSEN, C. W. HILBERS*, G. VAN DER MAREL+ and J. H. VAN BOOM+

Department of Biophysical Chemistry, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen and +*Department of Organic Chemistry, University of Leiden, Wassenaarseweg 64, 2333 RA Leiden, The Netherlands*

Received 1 October 1981

1. Introduction

Gene-5 protein encoded by the bacteriophages M13 fd and f_1 is a DNA-helix destabilizing protein. It plays an essential role in the replication process of the phage, where the binding of gene-5 protein to the viral DNA strand is required for single-stranded viral DNA synthesis [1–3]. The dimeric protein [4,5] with subunit M_r 9690 [6,7] binds cooperatively [1,5] to single stranded DNA thereby covering ~4 nucleotides/protein subunit [1,4,5,8]. There is a wealth of spectroscopic evidence that aromatic residues are involved in the interaction with DNA [4,8,9]. The molecular surroundings of 2 tyrosines and 1 phenylalanine are influenced by binding of DNA [10,11]. This phenylalanine and ≥1 of these tyrosines are located in close proximity to the DNA [10]. However, the interaction of gene-5 protein with DNA is strongly dependent on ionic strength [1]. One therefore expects that electrostatic interactions are also important in the binding. The most likely candidates for the interaction with the negatively-charged phosphate groups of the DNA are lysines and arginines. To investigate the influence of DNA binding on these residues we have performed ^{1}H NMR experiments at 500 MHz with gene-5 protein and its complex with oligoadenylic acids.

2. Experimental

2.1. *Materials*

The isolation and purification of the gene-5 protein were essentially as in [12]. The octanucleotide $d(A)_8$ was synthesized using the phosphotriester method [13]; $d(A)_{25-30}$, a mixture of oligoadenylic acids from 25–30 residues was purchased from P. L.-Biochemicals. Oligonucleotides were used as the Na^+-salt. Excess salt was removed by passage over a Sephadex G-10 column. Concentrations were determined from UV absorption using the extinction coefficients $\epsilon_{276\ nm} = 7100\ M^{-1} \cdot cm^{-1}$ for the gene-5 protein [8], $\epsilon_{260\ nm} = 10\,325\ M^{-1} \cdot cm^{-1}$ for $d(A)_8$ and $\epsilon_{260\ nm} = 9500\ M^{-1} \cdot cm^{-1}$ for $d(A)_{25-30}$. The extinction coefficients for the oligonucleotides were determined by enzymatic degradation of the oligonucleotides with venom phosphodiesterase (Worthington) [14] using an extinction coefficient for AMP of $15\,300\ M^{-1} \cdot cm^{-1}$. The concentrations are given in mol mononucleotide/l.

2.2. *Instrumentation*

Ultraviolet absorption measurements were done with a Zeiss PMQ II spectrophotometer. The pH-values reported are the uncorrected pH meter readings in D_2O solutions. 500 MHz ^{1}H NMR spectra were recorded on a Bruker WM-500 spectrometer operating in the Fourier transform mode. In decoupling experiments, resonances were selectively saturated by irradiating at the appropriate frequency in the pulsed homonuclear decoupling mode. Chemical shifts are quoted relative to DSS. Downfield shifts are defined as positive. Spectra were resolution-enhanced by a Lorentzian–Gaussian transformation [15,16].

3. Results

The overall 500 MHz ^{1}H NMR spectrum of gene-5 protein is shown in fig.1. We had assigned the resonances

* To whom correspondence should be addressed

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.75

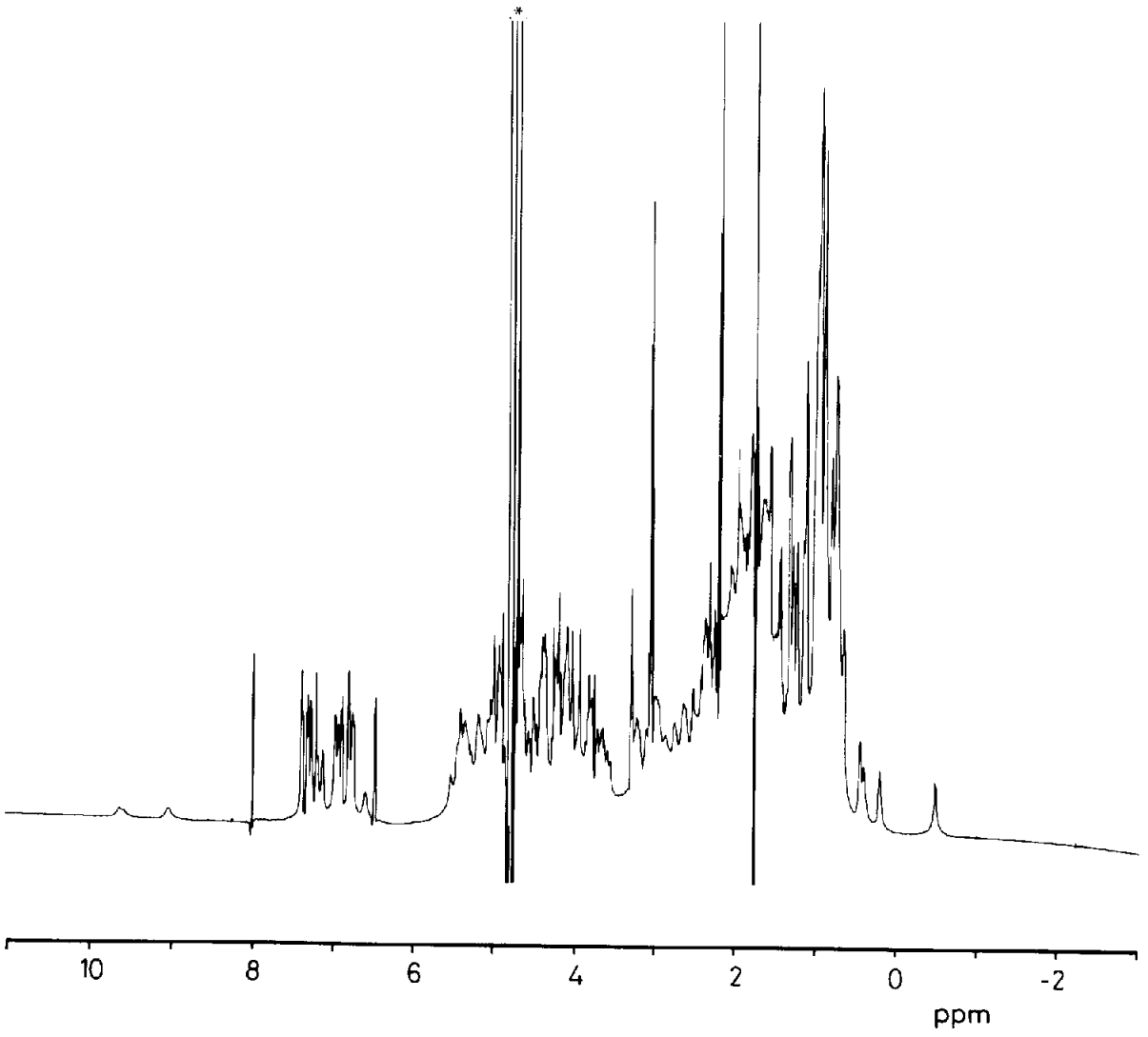

Fig.1. 500 MHz ^{1}H NMR spectrum of 1 mM gene-5 protein in 50 mM NaCl–1 mM cacodylate (pH 7.0) recorded at 28°C; the large resonance at ~1.7 ppm (*) originates from cacodylate.

in the aromatic part of the spectrum (8.1–6.4 ppm downfield from DSS) to the protons of the single histidyl the 5 tyrosyl and the 3 phenylalanyl residues in the protein [10]. The resonances of 2 tyrosyl and 1 phenylalanyl residues also undergo large shifts upon binding of oligonucleotides to the protein [10,11] and ≥2 of these residues are located close to the DNA, in the protein–DNA complex. Here, attention is given to the aliphatic part of the 500 MHz ^{1}H NMR spectrum, in particular to 3.5–2.9 ppm region, where the Lys-ϵ and Arg-δ protons are expected to resonate. This region of the spectrum is expanded in fig.2A. In addition to the resonances of the arginyl and lysyl protons mentioned above, this region is expected to contain several β-proton resonances and probably the resonances of the δ-protons of proline. Indeed photochemically induced nuclear polarization studies have indicated that the β-protons of the tyrosines are situated ~3 ppm [17]. Prominent in fig.2A are 2 triplets centered at 3.052 and 3.297 ppm, respectively. The presence of an extra narrow peak ~0.03 ppm downfield from the triplet at 3.052 ppm and the intensity distribution in this triplet suggests, that there is a third triplet centered at 3.065 ppm overlap-

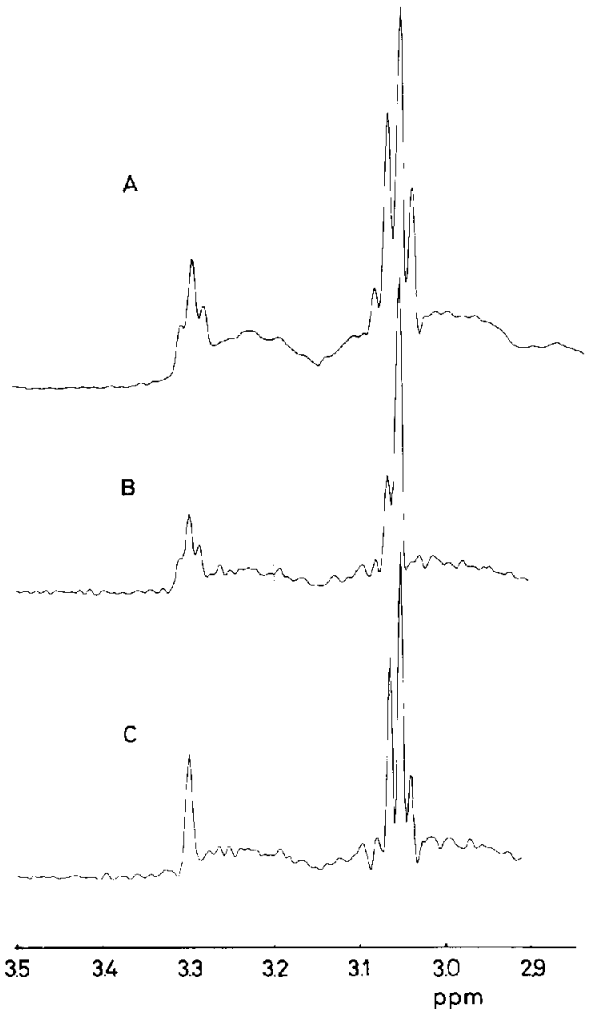

Fig.2. Expansions of the 500 MHz ^{1}H NMR spectra of gene-5 protein (1 mM) showing the 2.9–3.5 ppm spectral region downfield from DSS. Conditions are as given in fig.1: (A) undecoupled spectrum; (B) spectrum recorded during irradiation at 1.73 ppm; (C) spectrum recorded during irradiation at 1.79 ppm.

ping with the triplet at 3.052 ppm. This was verified by performing a decoupling experiment which is shown in fig.2B. As a result of irradiation at 1.73 ppm the 2 overlapping triplets have become single lines, appearing at 3.052 and 3.065 ppm, respectively. The resonance position of the smaller decoupled line demonstrates that the narrow peak observed at 3.082 ppm in fig.2A is indeed part of a triplet. Comparison of fig.2B with fig.2A shows furthermore that there is a small influence of the irradiation at 1.73 ppm on the triplet at 3.297 ppm. This is understandable in view of the experiment shown in fig.2C where the decoupling frequency was adjusted to 1.79 ppm. In this experiment the triplet at 3.297 is fully decoupled whereas the triplets at higher field are slightly influenced (the triplet at 3.065 ppm is somewhat more perturbed in this experiment than the triplet at 3.052 ppm). Of the resonances expected between 2.9 and 3.5 ppm only those belonging to the Lys-ϵ, Arg-δ and (in some cases) Pro-δ protons may appear as triplets coupled to resonances at 1.8 ppm. The narrowness of the triplets makes an assignment to Pro-δ protons improbable. In view of this and on the basis of the observed chemical shifts we assign the triplet at 3.297 ppm to Arg-δ protons and the triplets at 3.052 and 3.065 ppm to Lys-ϵ protons. It is difficult to determine the exact intensity of the triplets because they are superimposed on a background of broad resonances. However, taking the isolated methyl resonance at –0.51 ppm (see fig.1) as an internal standard, a measure of their intensities can be obtained. We estimate that the 2 lysyl triplets account for all 12 [6,7] Lys-ϵ protons present in the protein. On the basis of a stimulation of this spectral region we estimate that the smaller Lys-ϵ triplet contains 2–4 proton resonances while the larger triplet accounts for 8–10 protons. Thus calibrated 4–6 Arg-δ protons (out of the total of 8) contribute to the intensity of the triplet at 3.297 ppm. The intensities of the resonances in the decoupled spectra shown in fig.2B and in fig.2C yield no reliable values for the number of protons resonating in the different peaks, since the application of the decoupling power influences the intensity of the lines (Nuclear Overhauser Effect).

Upon binding of $d(A)_8$ to the gene-5 protein the most remarkable change in the spectral region of interest is the broadening and 0.055 ppm (27.5 Hz) ppm downfield shift of the arginyl resonance (fig.3B). In addition the smaller triplet of the Lys-ϵ protons is influenced by the binding. This is best seen in the decoupled spectrum shown in fig.3C where irradiation was applied at 1.73 ppm. From comparison of fig.3C with fig.2B it is clear that the smaller Lys-ϵ resonance broadens and shifts somewhat to low field upon binding of $d(A)_8$. By varying the irradiation frequency we verified that the increase in linewidth is not an artefact caused by incomplete decoupling. The broadening of the arginyl resonance in this complex is such that it proved impossible to sharpen this resonance by decoupling.

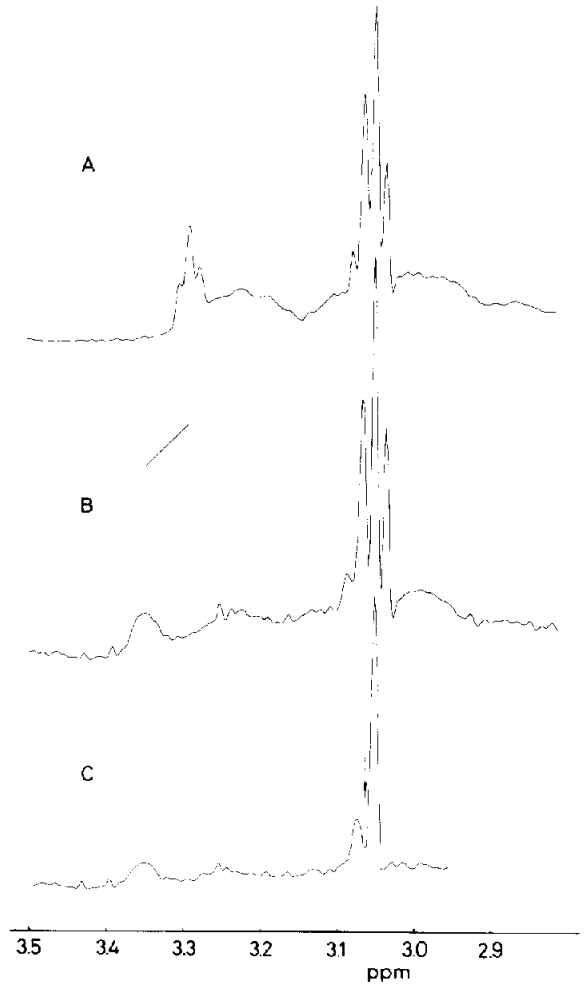

Fig.3. Expansions of the 500 MHz ^{1}H NMR spectra of gene-5 protein (1 mM) and of gene-5 protein (1 mM) with excess $d(A)_8$ added (4.8 mM in mononucleotides). Conditions are as given in fig.1: (A) same spectrum as shown in fig.2A; (B) spectrum of the gene-5 protein $d(A)_8$ complex undecoupled in which the shift of the arginyl δ protons is indicated; (C) spectrum of the gene-5 protein $d(A)_8$ complex recorded during irradiation at 1.73 ppm. The arrow indicates the spectral position of the smaller Lys-ε resonance in the free protein spectrum (compare fig.2B).

The broadening effects observed upon binding of gene-5 protein to $d(A)_8$ are more dramatic upon complexation of the protein to oligo $d(A)_{25-30}$. The relevant part of the ^{1}H NMR spectrum of this complex is displayed in fig.4A. The Arg-δ proton resonances have broadened beyond detection. This is also true for the low field triplet of the Lys-ε protons and can be observed in fig.4B in which again the Lys-ε protons

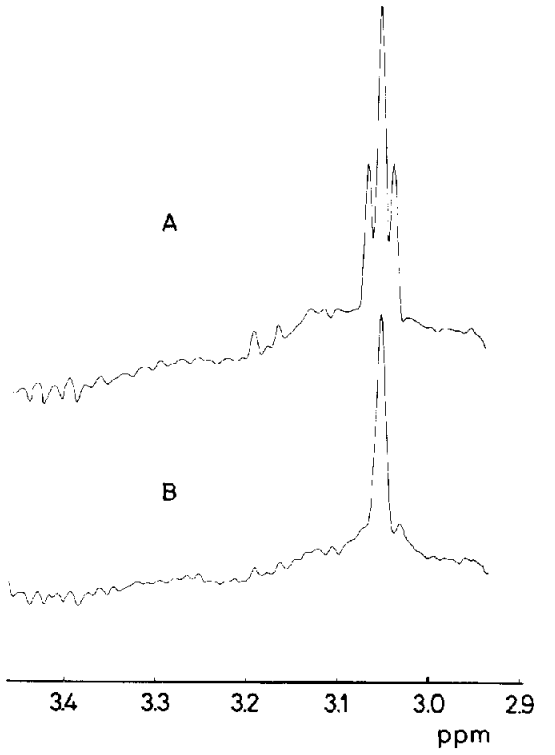

Fig.4. Expansions of the 500 MHz ^{1}H NMR spectra of gene-5 protein (1 mM) with excess $d(A)_{25-30}$ added (~6 mM in mononucleotides). Conditions are as given in fig.1: (A) undecoupled spectrum; (B) spectrum recorded during irradiation at 1.74 ppm.

have been decoupled from the Lys-δ protons. The small peak has disappeared from the spectrum while the large peak remains relatively narrow indicating that the Lys-ε protons contributing to this peak are to a considerable extent still free to move.

4. Discussion

The long side chains of lysines are known to lend the ε-CH_2 group a considerable amount of motional freedom, allowing the hydrophylic NH_3^+ group to interact with the solvent. Therefore most of the Lys-ε protons in a protein are expected to resonate at the position characteristic of a free lysyl residue. Lysines that are involved in (intramolecular) interactions with other residues appear at a different resonance position as was observed for Lys-41 in the basic pancreatic trypsin inhibitor which is interacting with a tyrosine

[18]. The shifted position of the smaller Lys-ϵ proton triplet in the gene-5 protein indicates that these protons are in an environment slightly different from that of the protons of the larger triplet. The small downfield shift of this triplet upon binding to $d(A)_8$ and the broadening observed for the complex with oligo $d(A)_{25-30}$ also differentiate the corresponding lysyl residues from the remaining lysines.

These remaining lysines of which the ϵ-protons resonate in the peak at 3.052 ppm retain their motional freedom even in the complex with $d(A)_{25-30}$ which has est. M_r of 140 000. Therefore we conclude, in contrast with earlier conclusions obtained at lower resolution [19] that 1 or 2 lysyl residues are in fact immobilized by the binding of oligonucleotides. Our experiments also indicate that the arginines in the free protein can be divided into 2 groups of residues. The first group consisting of 2 or 3 residues has an important amount of motional freedom in the free protein as witnessed by the narrowness of their δ-proton resonances, the triplet at 3.297 ppm. The remaining arginines give apparently rise to broad resonances. This is most likely the result of a restriction in their motion as a consequence of their involvement in intramolecular interactions. Upon binding of oligonucleotides the triplet at 3.297 ppm shifts downfield and broadens significantly. The broadening of this resonance was observed in a 270 MHz study [19] but the shifted resonance in the spectrum of the protein–oligonucleotide complex was not resolved. Just as for the smaller triplet of the Lys-ϵ protons we interpret the increased linewidth to indicate that the corresponding arginyl residues are immobilized as a result of complex formation.

Combination of these findings with earlier experimental data makes it plausible to interpret the shifts and broadenings of the lysyl and arginyl resonances as to result from interaction of these residues with the DNA. From X-ray diffraction data [20] it was concluded that 2 lysyl residues (24 and 26) and 3 arginyl residues (21,80 and 82) may be situated at the inner surface of the DNA binding groove. In [21] DNA binding was inhibited upon acetylation of the lysyl residues. The minor shift of the smaller Lys-ϵ triplet upon complexation with $d(A)_8$ is consistent with an 100 MHz ^{1}H NMR study [22] where broadening but no detectable shift for the Lys-ϵ proton resonances was found upon binding of lysine containing oligopeptides to poly(A). Interaction between a phosphate and a lysyl-NH_3^+ group may cause a change in the pK-value of the lysyl group [23]. Unfortunately, this cannot be checked by titration experiments because of the instability of the protein at high pH.

Since the observed changes in the spectral parameters of the Arg-δ and Lys-ϵ protons form only indirect evidence for the proposed interaction with DNA it is important to consider alternative explanations for these changes. The cooperativity of the protein–DNA binding is largely determined by protein–protein interactions and it could be that lysines and arginines are involved in these interactions. This could explain the observed spectral changes equally well. However, we consider this possibility rather unlikely because in an extensive series of fluorescence binding studies we found that the cooperativity of the binding is largely ionic strength-independent. On the other hand the intrinsic binding constant is strongly salt-dependent which means that ionic protein–DNA interactions are formed. Since the gene-5 protein has been shown to cover ~3 nucleotides/protein monomer when bound to short oligonucleotides [11] and ~4 nucleotides/protein monomer when bound to polynucleotides [1,4,5,8] in the binding groove ~4 positive charges are needed to neutralize the negatively charged phosphate groups.

In 500 MHz ^{1}H NMR spectra, different groups of basic residues can be distinguished and their involvement in protein nucleic acid interactions studied.

Acknowledgements

This research was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for the Advancement of Pure Research (ZWO). The NMR measurements were performed on the national (SON) 500 MHz facility at the University of Nijmegen. We thank Mr P. van Dael and Dr C. A. G. Haasnoot for keeping the instrument in excellent condition. Professor Dr J. G. G. Schoenmakers and Dr R. H. N. Konings are acknowledged for discussions and Miss C. Prinse for technical assistance in the preparation of the gene-5 protein.

References

[1] Alberts, B., Frey, L. and Delius, H. (1972) J. Mol. Biol. 68, 139–152.
[2] Oey, J. L. and Knippers, R. (1972) J. Mol. Biol. 68, 125–138.
[3] Ray, D. S. (1977) in: Comprehensive Virology (Fraenkel-

Conrat, H. and Wagner, R. eds) vol. 7, 105–178, Plenum, New York.
[4] Pretorius, H. T., Klein, M. and Day, L. A. (1975) J. Biol. Chem. 250, 9262–9269.
[5] Cavalieri, S. J., Neet, K. E. and Goldthwait, D. A. (1976) J. Mol. Biol. 102, 697–711.
[6] Cuyper, T., Van der Ouderaa, F. J. and De Jong, W. W. (1974) Biochem. Biophys. Res. Commun. 59, 557–564.
[7] Nakashima, Y., Dunker, A. K., Marvin, D. A. and Koningsberg, W. (1974) FEBS Lett. 40, 290–292; FEBS Lett. 43, 125.
[8] Day, L. A. (1973) Biochemistry 12, 5329–5339.
[9] Garssen, G. J., Tesser, G. I., Schoenmakers, J. G. G. and Hilbers, C. W. (1980) Biochim. Biophys. Acta 607, 361–371.
[10] Alma, N. C. M., Harmsen, B. J. M., Hull, W. E., Van der Marel, G., Van Boom, J. H. and Hilbers, C. W. (1981) Biochemistry 20, 4419–4428.
[11] Alma, N. C. M., Harsen, B. J. M., Van Boom, J. H., Van der Marel, G. and Hilbers, C. W. (1981) submitted.
[12] Garssen, G. J., Hilbers, C. W., Schoenmakers, J. G. G. and Van Boom, J. H. (1977) Eur. J. Biochem. 81, 453–463.
[13] Arentzen, R., Van Boeckel, C. A. A., Van der Marel, G. and Van Boom, J. H. (1979) Synthesis 137–139.
[14] Cassani, G. R. and Bollum, F. J. (1969) Biochemistry 8, 3928–3936.
[15] Ernst, R. R. (1966) Adv. Magn. Res. 2, 1–135.
[16] Ferrige, A. G. and Lindon, J. C. (1978) J. Magn. Res. 31, 337–340.
[17] Garssen, G. J., Kaptein, R., Schoenmakers, J. G. G. and Hilbers, C. W. (1978) Proc. Natl. Acad. Sci. USA 75, 5281–5285.
[18] Brown, L. R., De Marco, A., Wagner, G. and Wuthrich, K. (1976) Eur. J. Biochem. 62, 103–107.
[19] Coleman, J. E., Anderson, R. A., Ratcliffe, G. and Armitage, I. M. (1976) Biochemistry 15, 5419–5430.
[20] McPherson, A., Jurnak, F., Wang, A., Kolpak, F., Rich, A., Molineux, I. and Fitzgerald, P. (1980) Biophys. J. 32, 155–173.
[21] Anderson, R. A., Nakashima, Y. and Coleman, J. E. (1975) Biochemistry 14, 907–917.
[22] Dimicoli, J. L. and Helene, C. (1974) Biochemistry 13, 714–723.
[23] Brown, L. R. and Bradbury, J. H. (1975) Eur. J. Biochem. 54, 219–227.