Origin of sequence-specific recognition of DNA by non-intercalating anti-tumor antibiotics

Mrigank, A.K. Royyuru and V. Kothekar*

Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India

Received 7 October 1985

Partitioning of energy in the interaction of non-intercalating antibiotics (netropsin, netropsin without its cationic ends and two analogs of distamycin A) with different base sequences of B-DNA is studied here by the atom-atom potential technique and geometry optimization procedures. The results show that electrostatic forces contribute substantially to the stabilization energy as well as to the sequence specificity. The hydrogen-bonding term is also sequence specific and is significant in properly orienting the drug molecule. Relative roles of the hydrogen bonding and electrostatic interactions depend on the dielectric property of the medium.

DNA recognition *Netropsin* *Distamycin A analog* *Energy partitioning* *Base sequence specificity*

1. INTRODUCTION

Antitumour antibiotics of the netropsin, distamycin class bind to DNA molecule in the minor groove and exhibit marked conformation and base specificity [1–10], which is dependent on the surrounding medium. For example, under high salt conditions, poly $d(GC)_n \cdot d(GC)_n$ does not bind to netropsin whereas lowering the salt concentration enables it to bind [4]. These antibiotics have been the target of extensive physicochemical studies in the last couple of years because they serve as ideal 'models' for studying recognition of DNA by repressor proteins and enzymes [11–15].

The aim behind this work is to study the mechanism of sequence-specific recognition of DNA by non-intercalating antitumour antibiotics in the minor groove in changing medium. Rigorous computer model building with energy minimization based on atom-atom potentials is used. Partitioning of energy contributions and its variation with dielectric constant are studied here for netropsin without its cationic tails, netropsin, distamycin A analogs – distamycin 2 and another in which the formamide is replaced by nitro and the propylaminidinium chain is modified as shown in fig.1 (Net 1, Net 2, Dist A An., Dist 2); interactions with different base sequences $d(A)_5 \cdot d(T)_5$; $d(TATAT) \cdot d(ATATA)$; $d(AGAGA) \cdot d(TCTCT)$; and $d(C)_5 \cdot d(G)_5$ (DNA 1, DNA 2, DNA 3, DNA 4).

* To whom correspondence should be adressed

2. EXPERIMENTAL

2.1. *Data preparation*

Starting geometries for the drug molecules were taken from their crystallographic data [16,17]. DNA sequences in the B-form were generated on the basis of fiber X-ray diffraction data obtained by Arnott et al. [18]. Partial atomic charges were generated by the CNDO/II method [19]. Other parameters for the energy calculations were taken from Momany et al. [20].

2.2. *Geometry optimization*

First, 3 target points (TPs) were fixed with

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

respect to the DNA molecule. They can be specific hydrogen-bonding positions. Three points were then chosen on the drug molecule as attacking points (APs). AP_1 was first superimposed on TP_1 and AP_2–AP_1 were then made colinear with TP_2–TP_1. In the next step AP_3 was brought closest to the plans described by TP_1, TP_2, TP_3. Undesirable short contacts between the drug molecule and DNA were then removed by giving small rotations around single bonds in the drug molecule. This provides the starting drug geometry. This was followed by series of geometry operations to minimize the interaction energy. These operations consist of (i) rotation around helical axis, (ii) rotation around a bond, of either whole or part of the molecule, (iii) gliding along the helix, (iv) moving in or out of the groove, (v) moving towards a particular strand of the helix. The potential energy function consisting of conformation as well as interaction terms was calculated using the monopole approximation method and non-bonded, electrostatic, polarization and hydrogen-bonding contributions between interacting groups (phosphate sugar, bases, amides, rings, cationic tails, etc.) were computed. The group-group potential was then reorganised to observe individual contributions.

3. RESULTS AND DISCUSSION

3.1. *Drug-DNA model*

Fig.1 depicts the top view (in a plane perpendicular to the helical axis) for Net 1, Net 2, Dist A An. and Dist 2, whereas the drug position along with DNA 1 in the B-form is shown in fig.2.

It can be seen that all 4 drug molecules are snugly fitted in the minor groove with C_5H, $C_{11}H$ and amides N_4H, N_6H, N_8H of Net and C_4H, $C_{10}H$ and amides N_3H and N_5H of distamycin pointing to the floor of the groove and forming a concave

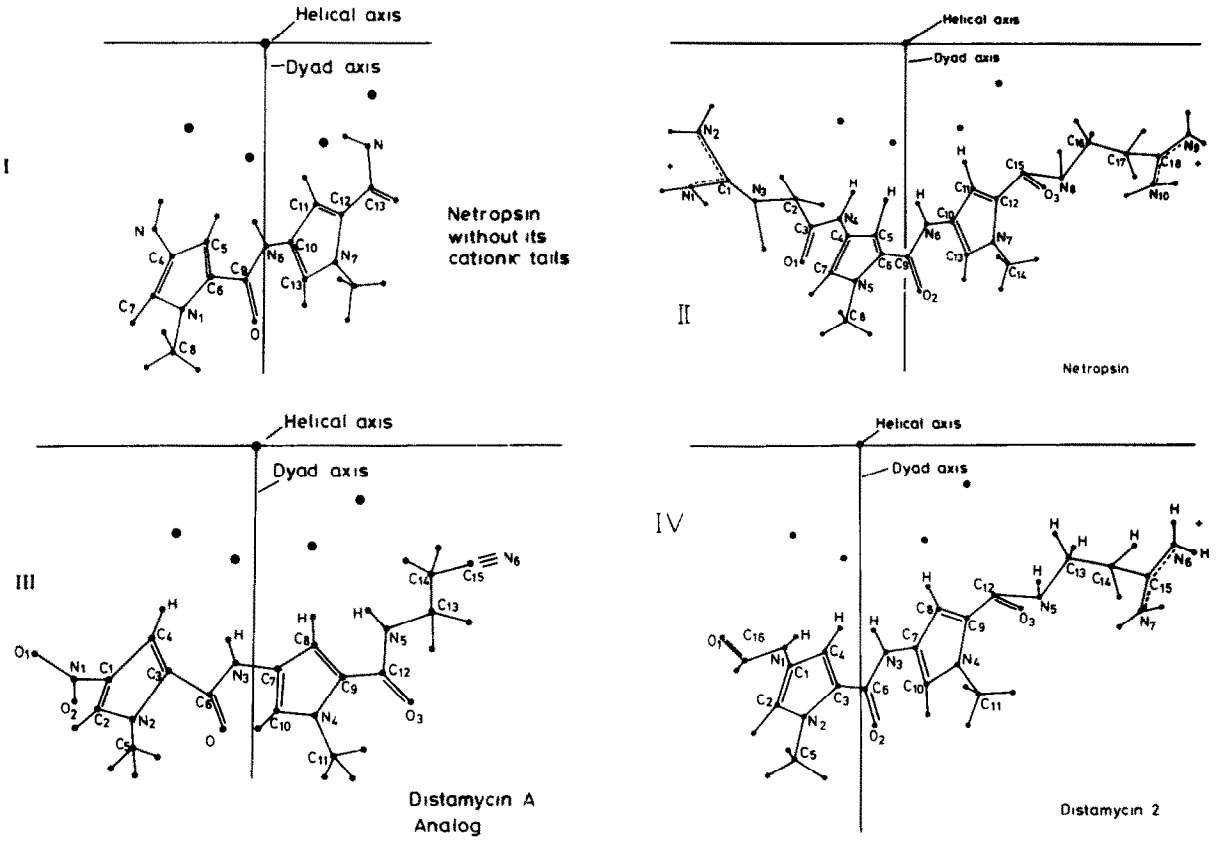

Fig.1. Top view of (I) Net 1, (II) Net 2, (III) Dist A An., (IV) Dist 2. (●) Position of hydrogen bonding.

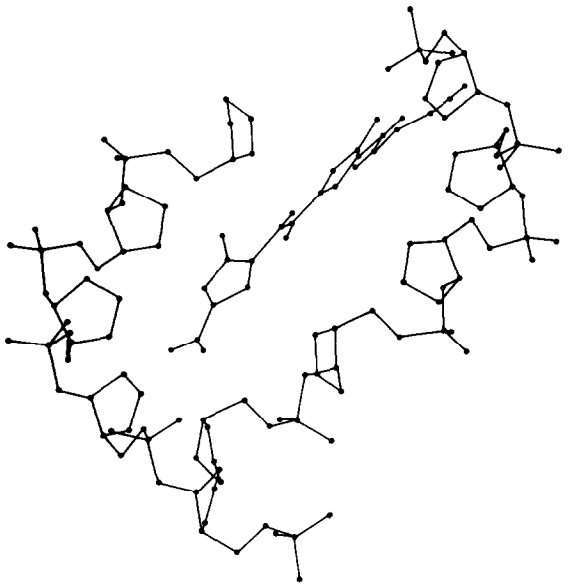

Fig.2. Side view of Dist A An. with DNA 1. Bases and hydrogens are omitted for the sake of clarity.

surface. The 2 pyrrole rings are oriented with respect to each other at an angle ranging from 18 to 20° which can be compared with about 20 and 4.5° in crystallographic studies of netropsin and Dist A An. [16,17] and 33° in a complex of netropsin with CGCGAATTCGCG [13]. The hydrogen bonds are observed between N_4H, N_6H of netropsin and N_3H, N_5H of distamycin with N_3 of purines and O_2 of pyramidines. Bifurcated hydrogen bonds were observed for all 4 drugs with DNA 1 and DNA 2. The drug molecule could bind parallel or antiparallel to the backbone with a very small (4.5 kcal/mol in the case of Net 2 with DNA 1) expenditure of energy. The general geometric features of our model are quite similar to the crystallographic model of Dickerson et al. [13] and agree with the NMR data [14], which shows the validity of the adopted procedure. The only difference between our model and that of Dickerson et al. was in the placement of the drug about the dyad axis. Drug was slightly shifted towards one strand in our model. This was because the minor groove in the B-DNA fiber diffraction data of Arnott et al. [18] is slightly wider compared to that of the model of Dickerson et al.

3.2. *Total interaction energy changes*

A summary of different energy contributions and energy break up for 2 strands is shown in table 1 for Net 1, Net 2, Dist A An. and Dist 2 for binding with DNA 1, DNA 2, DNA 3 and DNA 4 at $\epsilon = 4.0$. It was observed that the drugs could also bind at the periphery of the groove where its position becomes locked in the local minima. Multiple binding positions were observed in some cases. Energy break up for these additional positions is also shown in table 1. The total interaction energy follows the order: DNA 1 ≃ DNA 2 < DNA 3 < DNA 4, which is consistent with the binding constant for Net and Dist 2 obtained by Wartell et al. [4] and Gursky et al. [15]. It is also consistent with enthalpy measurements made by Zimmer et al. [21]. The relative binding order was: Net 2 > Dist 2 > Dist A An. > Net 1. Its direction agrees with the free energy change values for netropsin and Dist 2, viz. —12.3 and —2.91 kcal/mol [15, 22].

3.3. *Analysis of energy partitioning for different values of dielectric constants*

The relative contributions by bases-amides, bases-rings, bases-tails, backbone-amides, backbone-rings and backbone-tails for Net 1, Net 2, Dist A An. and Dist 2 with DNA 1, DNA 2, DNA 3, DNA 4 are depicted in fig.3.

The results show that in the case of netropsin, the major part (60%) of the stabilization energy arises from the electrostatic interactions between the tails and backbone. Its contribution at $\epsilon = 4.0$ in Dist 2 and Dist A An. is 42 and 0.5%, respectively. This term is susceptible to changes in dielectric constant and affects the binding constant. Depending on its weightage, it can be a major controlling factor in environment-based changes in the association constant as observed in the case of netropsin [4]. However, it is not sequence specific. Removal of the cationic tails will considerably reduce the changes in the binding energy caused by salt concentration variations.

Sequence specificity arises due to systematic differences in the electrostatic contribution from base-ring and base-tail interactions. Hydrogen bonding assists in proper orientation of the drug. However, its net contribution to the free energy of interaction is small at $\epsilon = 4.0$. It shows a sequence-dependent variation from DNA 1 to DNA 4. Its relative role in sequence specificity depends on the nature of the drug and the environment. For those

Table 1

Interaction energy (kcal/mol) Net 1, Net 2, Dist A An., Dist 2 with DNA 1, DNA 2, DNA 3, DNA 4 at $\epsilon = 4.0$

	Strand		Non	Ele-pol	Hydg	Total
	I	II				
Net 1 with						
DNA 1	− 36.544	− 31.588	−45.995	− 9.972	−12.408	− 68.132
DNA 2	− 35.512	− 32.940	−46.516	− 9.373	−12.563	− 68.452
DNA 3	− 31.418	− 19.786	−31.719	− 10.823	− 8.662	− 51.204
DNA 4	− 29.719	− 22.627	−35.380	− 10.426	− 6.541	− 52.346
Net 2 with						
DNA 1-I	−145.955	−110.769	−29.998	−215.009	− 8.357	−253:365
DNA 1-II	−112.813	−104.452	−15.027	−202.139	-	−217.266
DNA 2-I	−146.644	−106.947	−30.175	−215.115	− 8.301	−253.591
DNA 2-II	− 87.663	−116.197	−19.284	−184.576	-	−203.860
DNA 3-I	−137.006	− 96.355	−28.539	−197.131	− 7.692	−233.362
DNA 3-II	− 71.147	− 98.995	−13.658	−156.485	-	−170.143
DNA 4-I	−128.628	− 81.686	−33.503	−171.020	− 5.791	−210.315
DNA 4-II	− 74.466	− 90.966	−19.549	−145.882	-	−165.432
Dist A An. with						
DNA 1-I	− 42.320	− 31.001	−44.923	− 20.871	− 7.536	− 73.330
DNA 1-II	− 39.794	− 32.686	−43.699	− 19.419	− 9.361	− 72.480
DNA 2	− 41.821	− 34.132	−46.220	− 20.267	− 9.466	− 75.953
DNA 3	− 36.450	− 33.676	−44.090	− 18.252	− 7.784	− 70.126
DNA 4	− 34.708	− 34.260	−46.960	− 16.228	− 5.780	− 68.968
Dist. 2 with						
DNA 1	−106.157	− 61.419	−42.835	−121.137	− 3.605	−167.576
DNA 2	−107.868	− 60.075	−42.873	−121.388	− 3.682	−167.943
DNA 3	−110.661	− 48.445	−45.262	−110.668	− 3.176	−159.106
DNA 4	−105.171	− 44.121	−49.753	− 97.409	− 2.130	−149.292

Strand I and strand II are total interaction energies with first and second strands. Other terms give respectively nonbonded (Non), electrostatic with polarization (Ele-pol), hydrogen bonding (Hydg) and total energies

drugs which do not have cationic tails (Dist A An. and Net 1) it contributes 22 and 35%, respectively, to sequence specificity ($\epsilon = 4.0$). This increases with increase in the dielectric constant. In contrast, for drugs with cationic tails such as Net 2 or Dist 2, electrostatic interaction plays a central role in sequence-specific recognition as suggested by Lavery and Pullman [23].

Interesting interaction is shown by the amide groups. Negatively charged nitrogens show repulsion by both phosphate oxygens as well as N_3 and O_2 of bases. Phosphate repulsion probably assists in guiding the drug to the floor of the groove, where it is held by other interactions.

To sum up differences in the binding modes of different drugs with DNA sequences arise due to an intricate balance between the energy contributions mentioned above, each of which comprises a series of atom-atom contributions. Thus in the case of netropsin it is basically the electrostatic interaction with the DNA backbone and bases which leads to the free energy of association and specificity. The same holds true in the case of Dist 2 with reduced magnitude. In Dist A An. the relative weightage of hydrogen bonding is greater. Geometric differences in the binding of the ring portions of these drugs are negligible. The tails have high mobility and adjust themselves according to their structure. Conformational changes in the drug molecule have been found to

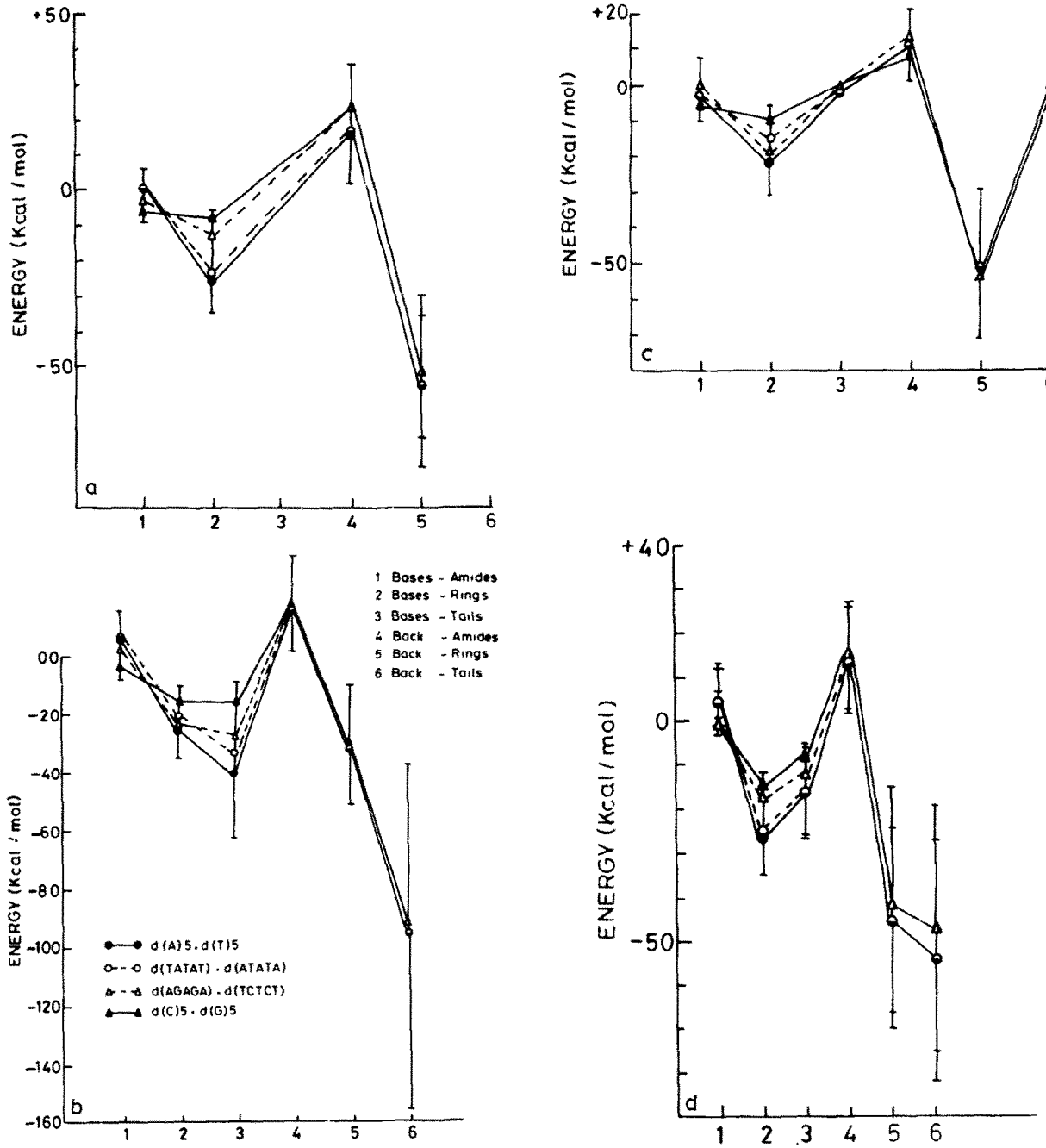

Fig.3. Partitioning of interaction energies of (a) Net 1, (b) Net 2, (c) Dist A An., (d) Dist 2 with DNA sequences. The curves depict average values for the change in dielectric constant between 4–20. Variations in electrostatic and polarization terms are shown by vertical bars.

affect these contributions significantly. Hence the conformation plays an equally important role in sequence specificity.

Work on the effect of DNA backbone deformations and of different ionic media on DNA recognition by these drugs is in progress.

ACKNOWLEDGEMENTS

The authors are thankful to Department of Science and Technology for financial assistance, to Professor R.K. Mishra for his encouragement and to Dr A. Kalia for her critical comments.

REFERENCES

[1] Van Dyke, M.W., Hertzberg, R.P. and Dervan, P.B. (1982) Proc. Natl. Acad. Sci. USA 79, 5470–5474.
[2] Zimmer, C., Reinert, K.E., Luck, G., Wahnert, U., Lober, G. and Thrum, H. (1971) J. Mol. Biol. 58, 329–348.
[3] Luck, G., Triebel, H., Waring, M. and Zimmer, C. (1974) Nucleic Acids Res. 1, 503–530.
[4] Wartell, R.M., Larson, J.E. and Wells, R.D. (1974) J. Biol. Chem. 249, 6719–6731.
[5] Zimmer, C. (1975) Prog. Nucleic Acids Res. Mol. Biol. 15, 285–318.
[6] Lafer, E.M., Moller, A., Nordheim, A., Stollar, B.D. and Rich, A. (1981) Proc. Natl. Acad. Sci. USA 78, 3546–3550.
[7] Kolchinsky, A.M., Mirzabekov, A.D., Zasedatelev, A.S., Gursky, G.V., Grokhovsky, S.L., Zhuze, A.L. and Gottikh, B.P. (1975) Mol. Biol. (USSR) 9, 14–20.
[8] Reinert, K.E. and Thrum, H. (1970) Stud. Biophys. 24, 319–325.
[9] Gursky, G.V., Tumanyan, V.G., Zasedatelev, A.S., Zhuze, A.L., Grokhovsky, S.L. and Gottikh, B.P. (1977) in: Nucleic Acid-Protein Recognition (Vogel, H.J. ed.) pp. 189–217, Academic Press, New York.
[10] Zimmer, C., Marck, C., Schneider, C. and Guschlbauer, W. (1979) Nucleic Acids Res. 6, 2831–2837.
[11] Zakrzewska, K., Lavery, R. and Pullman, B. (1983) Nucleic Acids Res. 11, 8825–8839.
[12] Zakrzewska, K., Lavery, R. and Pullman, B. (1984) Nucleic Acids Res. 12, 6559–6574.
[13] Kopka, M.L., Pjura, P., Yoon, C., Goodsell, D. and Dickerson, R.E. (1984) in: Structure and Dynamics of Nucleic acids, Proteins and Membranes (Clementi, E. and Sarma, R.H. eds) Adenine Press, New York.
[14] Gupta, G., Sarma, M.H., Sarma, R.H. (1984) J. Biomol. Struct. Dyn. 1, 1457–1472.
[15] Gursky, G.V., Zasedatelev, A.V., Zhuze, A.L., Khorlin, A.A., Grokhovsky, S.A., Streltsov, S.A., Surovaya, A.N., Nikitin, S.M., Krylov, A.S., Retchinsky, V.O., Mikhailov, M.V., Beabealashvili, R.S. and Gottikh, B.P. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 367–378.
[16] Berman, H.M., Neidle, S., Zimmer, C. and Thrum, H. (1979) Biochim. Biophys. Acta 561, 124–131.
[17] Gurskaya, G.V., Grokhovsky, S.L., Zhuze, A.L. and Gottikh, B.P. (1979) Biochim. Biophys. Acta 563, 336–342.
[18] Arnott, S., Smith, P.J.C. and Chandrasekaran, R. (1976) in: CRC Handbook of Biochemistry and Molecular Biology (Fasman, G.D., ed.) vol. 2, pp. 411–422, CRC, Cleveland.
[19] Pople, J.A. and Segal, G.A. (1966) J. Chem. Phys. 44, 3289–3296.
[20] Momany, F.A., Carruthers, L.M., McGuire, R.F. and Scheraga, H.A. (1974) J. Phys. Chem. 16, 1595–1620.
[21] Zimmer, Ch., Luck, G., Lang, H. and Burckhardt (1978) 12th FEBS Meet., Dresden.
[22] Marky, L.A., Blumenfeld, K.S. and Breslauer, K.J. (1983) Nucleic Acids Res. 11, 2857–2870.
[23] Lavery, R., Pullman, B. (1981) Nucleic Acids Res. 9, 4677.