Energetics of interaction of oligopeptide (lac 53–57) with DNA base sequences and origin of sequence-specific recognition

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

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

Received 6 October 1985

Computer model building with a dynamic energy minimization procedure is used here to study the interaction of a pentapeptide sequence from the *lac* repressor headpiece (lac 53–57) with different base sequences of DNA. The peptide fragment for this purpose was considered in the classical β-antiparallel as well as the β-associated conformation. The model of its interaction with DNA was optimised for various binding positions and base sequences. Partitioning of energy is analysed for different dielectric constant values and the main contributing factors to sequence-specific binding are discussed.

DNA recognition *lac 53–57* *Energy partitioning*

1. INTRODUCTION

The mechanism by which the *lac* repressor recognises the 20 bp region of *E. coli* and has a very tight binding has been under investigation for several years [1–10]. It has been established that the N-terminus headpiece of 59 residues is really necessary for recognition [2,3]. The region of amino acid residues from 53 to 59 has been specially identified as 'recognising' on the basis of genetic studies [4]. Physico-chemical studies on the interaction of amino acid residues from this region with DNA fragments showed that it does not modify the DNA conformation, assumes a β-associated structure upon binding to DNA and has sequence-specific interaction with DNA [5]. Increase in the melting temperature of poly[d(I-C)]·poly[d(I-C)] and poly(dG)·poly(dC) has also been observed [9].

The aim of this paper is to look into the nature of the intramolecular forces responsible for recognition of DNA by a pentapeptide fragment from this region (sequence 53–57). We use computer model building along with a dynamic energy minimization procedure to observe the effect of peptide conformation, binding position, and dielectric constant changes on the recognition of DNA base sequences.

2. EXPERIMENTAL

Pentanucleotide duplexes with 4 different base sequences $d(T)_5 \cdot d(A)_5$, $d(TATAT) \cdot d(ATATA)$, $d(CGCGC) \cdot d(GCGCG)$, and $d(C)_5 \cdot d(G)_5$ (DNA I, DNA II, DNA III, and DNA IV, respectively) were generated in B-form on the basis of fiber X-ray diffraction data of Arnott et al. [11]. Their conformation was kept fixed throughout these computations. Pentapeptide in the classical antiparallel β-conformation was generated on the basis of IUPAC data [12]. All energy calculations were done on the basis of atom-atom potential consisting of non-bonded (attractive and repulsive), electrostatic, polarization and hydrogen bonding contributions [13]. CNDO/ON charges [14,15] were used for calculations.

As a first step towards model building, the peptide backbone was considered to be rigid. Side

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)

chains were allowed to rotate around C_α–C_β bonds. Peptide was allowed to approach DNA in the major as well as minor groove, in parallel or antiparallel orientation, with side chains wrapped around DNA or penetrating the grooves (position I and II, respectively). Geometry optimization was done allowing freedom to roll, slide, and vibrate in the grooves by a procedure described by us earlier [16].

In the next step, the backbone of the peptide was allowed to be deformed. We also allowed rotations around all single bonds in the side chains. The geometry of the peptide DNA complex was optimised in the minor groove for 2 base sequences (DNA I and DNA IV).

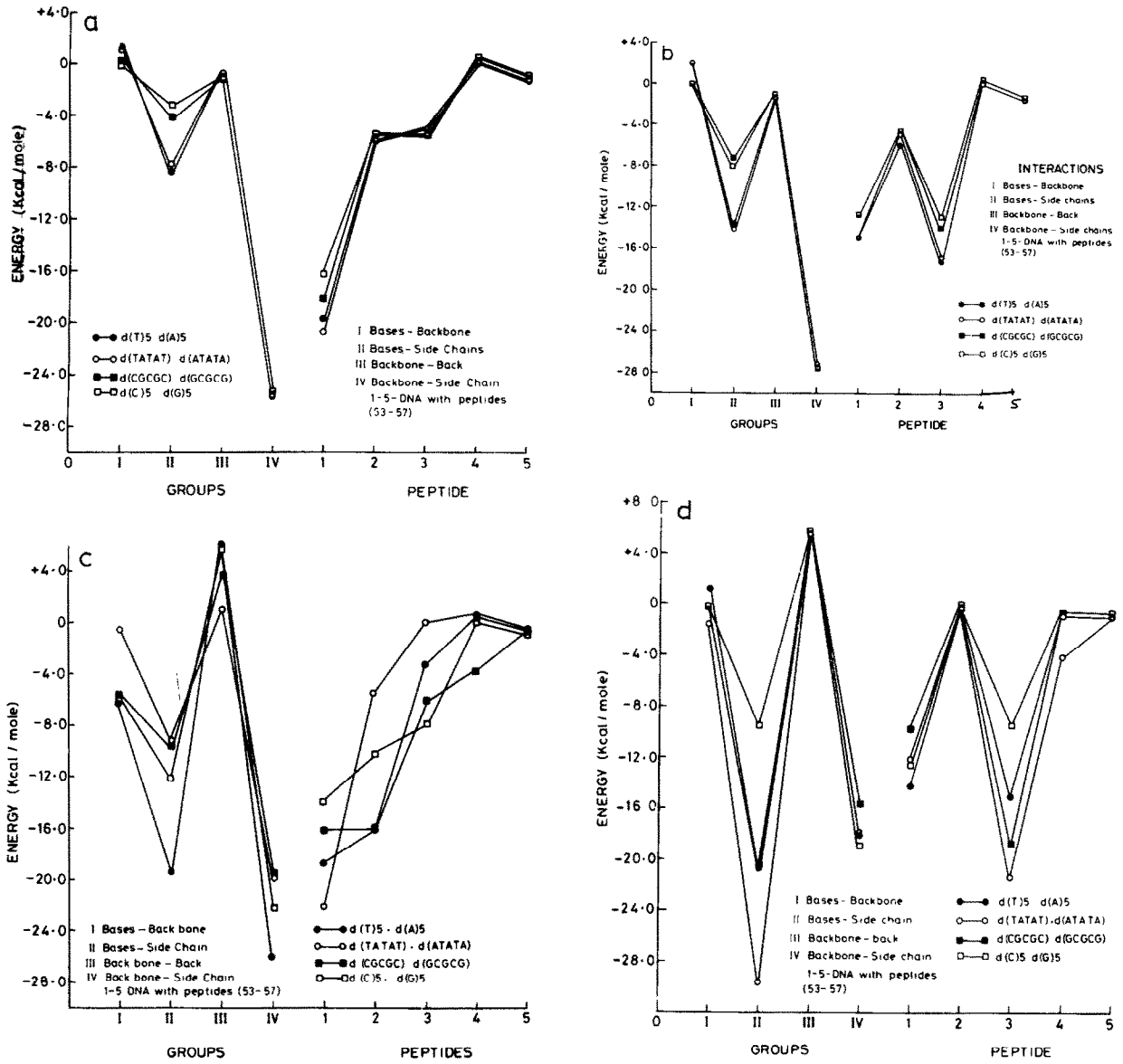

Fig.1. Partitioning of interaction energy of lac (53–57) in classical β-antiparallel conformation, (a) minor groove wrapped around, (b) minor groove penetrating, (c) major groove wrapped around and (d) major groove penetrating position. Average values for ϵ = 4–20 are shown.

3. RESULTS AND DISCUSSION

In the classical β-antiparallel conformation, the presence of 3 large side chains, Gln 54, Gln 55 and Leu 56, restricts the entry of the peptide fragment in the minor groove. As a result of this, DNA bases contribute less to the stabilization energy (fig.1a). In the penetrating position only one of the side chains (Gln 55) has an appreciable interaction with the bases (fig.1b). The maximum energy variation is 12 and 13% in positions I and II (table 1). A-T containing sequences showed preference because of the electrostatic interaction in the wrapped around position and hydrogen bonding in the penetrating position (with N3 of adenine or O2 of thymine).

The percentage variation of the interaction energy with different base sequences in the major groove is about 20 in the wrapped around position and 33 in the penetrating position. This increase in specificity is due to the large size of the major groove which allows more penetration. However, the binding affinity of classical β-peptide fragment in both the grooves is much smaller ($\Delta\bar{E}$ = −39.63 kcal/mol) than those of the non-intercalating antitumour antibiotics netropsin or distamycin which had $\Delta\bar{E}$ = −237.25 and −160.7 kcal/mol respectively at ϵ = 4.0 [16]. Change in the dielectric constant has no appreciable effect on sequence specificity. Peptide can bind in both parallel and antiparallel positions.

A dramatic increase in the binding affinity ($\Delta\bar{E}$ = −113.868 kcal/mol) is observed when backbone and side chain flexibility is allowed for the peptide. This is because the peptide in this case can penetrate deeper in the grooves and make specific contacts with DNA bases and backbone (fig.2). We observe that the peptide backbone can make a beautiful concave surface of radius 4.75 and 5.3 Å in the case of DNA I and DNA IV with alternate amides pointing towards the floor of the groove (fig.2). The net increase in conformational

Table 1

Partitioning of interaction energy of lac (53–57) for different base sequences at ϵ = 4.0

Peptide conformation	Groove	DNA base sequence	Peptide with strand I (5′)	Peptide with strand II (3′)	Nonbonded	Electrostatic	Hydrogen bonding	Total
β deformed	minor	DNA I	−76.854	−43.226	−84.685	− 9.845	−25.550	−120.080
		DNA IV	−59.523	−48.033	−80.348	−14.290	−12.918	−107.556
Classical β	minor I	DNA I	−15.872	−23.904	−19.060	−17.732	− 2.984	− 39.776
		DNA II	−16.604	−23.638	−19.469	−18.281	− 2.493	− 40.242
		DNA III	−14.704	−21.614	−19.956	−14.787	− 1.577	− 36.319
		DNA IV	−14.999	−21.127	−19.901	−16.226	0.0	− 36.126
	minor II	DNA I	−23.307	−22.064	−22.870	−10.178	−12.322	− 45.371
		DNA II	−24.213	−21.055	−22.844	−10.677	−11.747	− 45.267
		DNA III	−20.892	−19.106	−23.614	− 9.820	− 6.564	− 39.998
		DNA IV	−22.811	−17.424	−23.555	− 8.946	− 7.733	− 40.235
	major I	DNA I	−17.913	−30.644	−23.432	−15.680	− 9.447	− 48.559
		DNA II	− 8.857	−28.173	−15.227	−21.802	0.0	− 37.030
		DNA III	−11.063	−31.471	−23.857	−13.323	− 5.354	− 42.534
		DNA IV	−17.513	−22.228	−20.082	−15.330	− 4.330	− 39.741
	major II	DNA I	− 7.570	−29.206	−14.180	−12.860	− 9.736	− 36.776
		DNA II	−20.255	−24.607	−18.152	− 7.779	−18.930	− 44.861
		DNA III	−25.419	−21.686	−12.292	− 5.796	−15.315	− 33.402
		DNA IV	− 6.600	−21.456	−10.154	−10.839	− 7.063	− 28.056

All energies are in kcal/mol

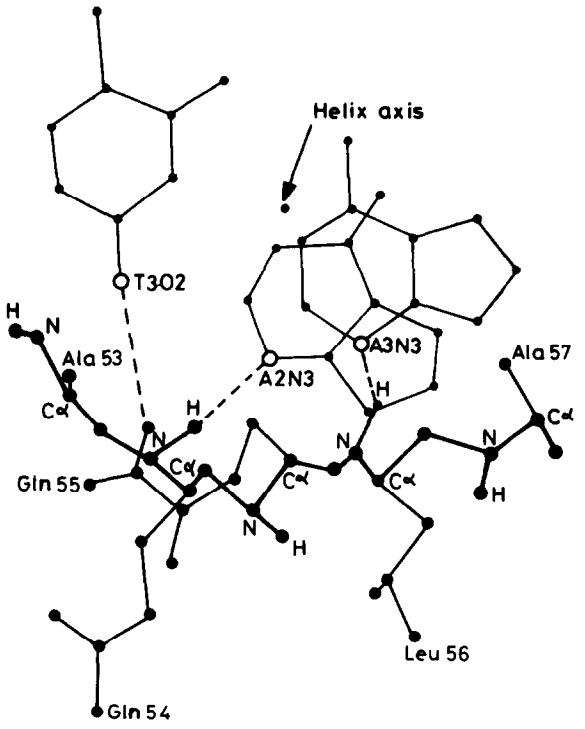

Fig.2. Top view (perpendicular to the helical axis) of lac (53–57) with deformed β-conformation in the minor groove of $d(T)_5 \cdot d(A)_5$. Peptide backbone is shown in bold lines, hydrogen bonding by dashed lines.

energy of this model is less than 10 kcal/mol. In agreement with the detailed model proposed by Gursky et al. [17] for peptide DNA recognition and the crystallographic structure of triostin-DNA [18], we observe hydrogen bonding between backbone amides (of the 2nd and 4th peptide) with the 2nd and 3rd adenine N3 on the 5′-strand. The direction N–C_α–C runs parallel to C3′-C5′ of the adenine strand. The obtained ϕ,ψ values for the 5 residues (144; 51, 156; 171, 140; −165, 152; −121°) lie in the allowed region in Ramachandran's (ϕ,ψ) plot [19]. We observe deviation from planarity of the peptide bond (w = 180, −142, −150; −163°). Such deformation can be induced by the DNA environment and has been noted in the triostin-DNA complex [18]. The side chains can also make specific contacts. Gln 55 could be twisted within the groove and hydrogen bonds with thymine O2.

Partitioning of the interaction energy for 2 strands in the case of $d(T)_5 \cdot d(A)_5$ was asymmetric

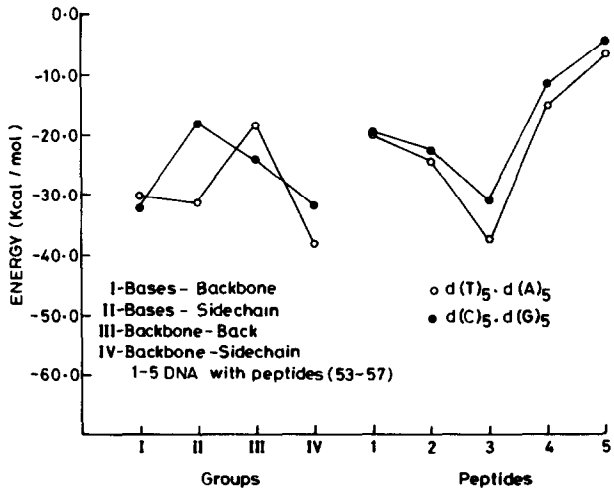

Fig.3. Partitioning of interaction energy of lac (53–57) in the deformed β-conformation.

(table 1). In the case of $d(C)_5 \cdot d(G)_5$ the presence of NH2 of guanine reduces its binding with the 5′-strand and symmetric partitioning is observed (table 1). The difference in the interaction energy with these 2 sequences is 12.524 kcal/mol. This is derived from the hydrogen binding interaction between the peptide backbone and DNA bases and from the electrostatic interaction between side chains and bases (fig.3). Some amount of sequence specificity is also shown by the side chain-backbone interaction which is a major contributing factor to the stabilization energy. It arises because of a better orientation of the side chains in the case of DNA I. No substantial change in specificity is observed due to change in the dielectric constant of the medium.

4. CONCLUSION

Our results demonstrate that the peptide in the classical β-antiparallel conformation binds weakly to DNA without much groove, orientation or sequence specificity. It seems to be ill-suited for the recognition process in the *lac* repressor. In the deformed β-conformation, a series of specific contacts are possible between the peptide backbone and side chains with DNA, which can lead to very tight binding as well as sequence specificity. Thus, in agreement with the literature [8,17,20,21], we feel that recognition of peptide in the minor groove

involves the β-associated species. More rigorous model building with peptide-DNA complexes is in progress. This is the first preliminary report.

ACKNOWLEDGEMENTS

We thank the Department of Science and Technology for financial assistance and award of fellowships to A.K., A.K.R. and M. One of the authors (V.K.) is thankful to G.V. Gursky and A.S. Zasedatelev for some helpful discussions.

REFERENCES

[1] Miller, J. and Reznikoff, W.S. (1978) The Operon, Cold Spring Harbor Laboratory, NY.
[2] Ogata, R.T. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA 75, 5851–5854.
[3] Kania, J. and Brown, D.T. (1976) Proc. Natl. Acad. Sci. USA 73, 3529–3533.
[4] Muller-Hill, B., Gronenborn, B., Kania, J., Schlotmann, M. and Beyreuther, K. (1977) in: Nucleic Acid Protein Recognition (Vogel, H.J. eds) p.219, Academic Press, New York.
[5] Gursky, G.V., Zasedatelev, A.S., Zhuze, A.L., Khorlin, A.A., Grokhovsky, S.L., 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.
[6] Hogan, M., Wemmer, D., Bray, R.P., Wade Jardetzky, N. and Jardetzky, O. (1981) FEBS Lett. 124, 202–203.
[7] Maurizot, J.C. and Charlier, M. (1977) FEBS Lett. 83, 107–110.
[8] Rein, R., Kieber-Emmons, T., Haydock, K., Garduno-Jurez, R. and Shibata, M. (1983) J. Biomol. Struct. Dyn. 1, 1051–1079.
[9] Kaptein, R., Zuiderweg, E.R.P., Scheek, R.M., Boelens, R. and Van Gunsteren, W.F. (1985) J. Mol. Biol. 182, 179–182.
[10] Rao, M.V.R., Atreyi, M., Kumar, S., Kumar, G.S. and Chauhan, V.S. (1985) in: Proceedings of the International Symposium on Biomolecular Structure and Interactions, to be published.
[11] Arnott, S., Smith, P.J.C. and Chandrasekaran, R. (1976) in: CRC Handbook of Biochemistry and Molecular Biology (Fasman, G.D. eds) vol.2, pp.411–422, CRC, Cleveland.
[12] IUPAC-IUB Commission on Biological Nomenclature (CBN), Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains (1970) Eur. J. Biochem. 193–201.
[13] Momany, F.A., McGuire, R.F., Burgers, A.W. and Scheraga, H.A. (1975) J. Phys. Chem. 79, 2361–2366.
[14] Momany, F.A., Carruthers, L.M., McGuire, R.F. and Scheraga, H.A. (1974) J. Phys. Chem. 78, 1595–1620.
[15] Rein, R., personal communication.
[16] Mrigank, Royyru, A.K. and Kothekar, V., to be published.
[17] Gursky, G.V., Tumanyan, V.G., Zasedatelev, A.S., Zhuze, A.L., Grokhovsky, S.L. and Gottikh, B.P. (1976) Mol. Biol. Rep. 2, 413–425.
[18] Wang, A.H.J., Ughetto, G., Quigley, G.J., Hakoshima, T., Vander Marel, G.A., Van Boom, J.H. and Rich, A. (1984) Science 225, 1115–1121.
[19] Ramachandran, G.N. and Sasisekharan, V. (1968) Adv. Protein Chem. 23, 283–437.
[20] Carter, C.W. and Kraut, J. (1977) Proc. Natl. Acad. Sci. USA 71, 283–287.
[21] Chruch, G.M., Sussman, J.L. and Kim, S.H. (1971) Proc. Natl. Acad. Sci. USA 74, 1458–1462.