Theor Chem Acc (2001) 106:83–90
DOI 10.1007/s002140100260

Theoretical
Chemistry Accounts

Regular article

Protein-induced DNA bending: the role of phosphate neutralisation

Raphael Gurlie, Krystyna Zakrzewska

Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received: 1 September 2000 / Accepted: 5 January 2001 / Published online: 3 May 2001

Abstract. An important component of protein–nucleic acid interactions is the formation of salt bridges between cationic amino acid side chains and the anionic phosphate groups of the nucleic acid. We have used molecular mechanics to study the energetic and conformational impact of such interactions. Firstly, crystallographic protein–nucleic acid complexes from the Protein Data Bank were analysed in terms of DNA curvature and the presence of salt bridges. For complexes where the DNA is significantly bent, the contribution of salt bridges to this curvature was modelled by studying the effect of neutralising the appropriate phosphate groups. The number and the distribution of salt bridges vary widely for different DNA binding motifs and appear to have very different effects on DNA. In the case of homeodomain, bZIP and helix–loop–helix proteins, salt bridges induce DNA bending, whereas for prokaryotic helix–turn–helix proteins the number of salt bridges is much smaller and little bending is found. By analysing the components of the DNA deformation energy involved in protein binding we show that salt bridges consistently increase the flexibility of the DNA backbone.

Key words: DNA–protein complexation – Phosphate neutralisation – Molecular recognition – Junction minimisation of nucleic acids – DNA curvature

1 Introduction

An important challenge in structural molecular biology is to understand the molecular mechanisms involved in protein–DNA recognition. How does a protein distinguish a particular sequence amongst a large background of similar sequences? A first tentative answer to this question was based on the "direct readout" of sequence information via the formation of specific hydrogen bonds between the two partners. However, it has become clear that other factors must be important and, more recently, considerable emphasis has been put on "indirect readout" mechanisms involving the sequence-dependent structural and dynamic properties of the nucleic acid target. One of the most important features of DNA sequences seems to be their bendability, since axis deformation is a common feature of the protein–DNA complexes whose structures have been solved in recent years. A good, although extreme, example of this is the IHF complex, where DNA forms an elegant superhelix around the protein with a bending angle of the order of 80° per crystallographic unit [1]. Another well-known example is the TATA-box binding protein (TBP), where the induced DNA bending reaches 100° [2, 3, 4].

Two mechanisms have been evoked to explain such strong bending. The first involves the intercalation of protein side chains between the DNA bases. This mechanism was proposed for proteins which bind in the DNA minor groove, such as TBP, SRY and LEF-1 [5, 6, 7]. By partially intercalating amino acid side chains on the minor groove side of DNA, the double helix is bent away towards the major groove. The second mechanism is electrostatic in nature and was proposed 20 years ago by Mirzabekov and Rich [8]. It is based on the reduction in phosphate–phosphate repulsion on one side of the double helix as the result of interactions with cationic amino acid side chains. This concept has been confirmed experimentally by studying the bending of modified DNAs which contain patches of neutral methylphosphonate groups [9]. In recent years, the mechanism of electrostatically induced bending has been refined by looking at the effect on curvature of individual charges belonging to the basic segment of bZIP proteins. For GCN4 it has been shown that curvature can be created by replacing neutral amino acids (Pro-Ala-Ala) in the N-terminal of the basic region with either positively or negatively charged residues [10, 11, 12]. The cationic or anionic character of these residues and the net charge of the protein surface has been found to correlate with

Correspondence to: K. Zakrzewska
Contribution to the Symposium Proceedings of Computational Biophysics 2000

the bending direction: towards the protein for positive charges and away for negative charges. It has also been shown that DNA interactions with anionic side chains can be reproduced by substitution of the corresponding phosphate groups with neutral methylphosphonates [13]. DNA bending is found to be a linear function of the number of neutralised phosphates, with roughly 3.5° of bending per group. Related studies have treated the role of phosphate neutralisation within the PU.1 complex [14]. While there is only 8° of bending in the crystallised complex of this protein, substitution of phosphates involved in salt bridges with methylphosphonates increases the bending to 28° according to electrophoretic measurements [14]. This work led to the conclusion that electrostatic forces were not sufficient to account for DNA bending within the complex and that the pattern of neutralisation and the DNA sequence were important factors in modulating DNA deformation.

The present studies aim at analysing such effects with the help of molecular mechanics. Our work in this area has already enabled us to show that modelling can demonstrate clear effects of phosphate neutralisation for both CAP and HNF3γ complexes [15, 16]. These two proteins, however, act rather differently. In the case of CAP, the DNA binding sequence is intrinsically curved in the direction appropriate for complexation and neutralising the phosphates which make salt bridges within the complex increase this curvature without changing its direction. In the case of HNF, the DNA binding sequence is bent in a direction very different to that observed in the complex and the main effect of phosphate neutralisation is to change this direction towards that found in the presence of the protein. In both cases, we found that the effect of neutralising the phosphates involved in salt bridges (the ''neutralisation pattern'') was strongly dependent on the base sequence, both in terms of magnitude and direction of bending.

To extend this work and reach more general conclusions, we have now analysed the phosphate neutralisation patterns of a wide range of protein–DNA complexes taken from the Protein Data Bank (PDB) [17]. We discuss here only those complexes where the bound DNA is significantly curved (20° or greater measured between the ends of the helical axis generated by CURVES [18]) and where salt bridges between DNA phosphates and cationic amino acid side chains are present. By energy-optimising the DNA sequences targeted, with or without neutralisation of the phosphates belonging to salt bridges, we try to understand the general rules governing their effect on DNA bending. Using a new method for forcing the DNA target to adopt its protein-complexed conformation we have also been able to look at the effect of phosphate neutralisation on deformation energetics.

2 Methodology

2.1 Conformational analysis

The analysis of the conformation of irregular DNAs clearly depends on the technique employed and a consensus has been slow to emerge in this field. The CURVES algorithm has an advantage for significantly bent DNAs where it is important to define the nature and extent of axis bending [18]. CURVES determines an optimal curvilinear helical axis via the minimisation of a least-squares function which describes departures from ideal helical symmetry. These departures are optimally distributed between deformation of the helical axis (bending and dislocation) and the rotational and translation positioning of successive bases or base pairs with respect to this axis.

We used CURVES to analyse the crystallised protein–DNA complexes presently available in the PDB. The complexes in which DNA is curved by at least 20° (measured between the ends of the helical axis) were selected. For each complex, we searched for protein atoms lying within 4 Å of the phosphate anionic oxygens. If these atoms were side chain cationic nitrogens belonging to arginine, lysine or histidine they were treated as forming salt bridges with the corresponding phosphate group. In agreement with the literature [19] many contacts are also observed with peptide backbone nitrogens. It is remarked that we also searched for, and found, a considerable number of anionic oxygens belonging to glutamic or aspartic acid side chains within 4 Å of phosphate oxygens, but their effects were not studied. Well-resolved complexes containing salt bridges were selected for further study (Table 1).

2.2 Molecular mechanics calculations

All energy minimisations were performed using the junction minimisation of nucleic acids (JUMNA) program (version 10.0) [20]. In order to limit the number of degrees of freedom representing the flexibility of a given nucleic acid fragment, this program uses a combination of helicoidal and internal variables. Single bond torsions and valence angles (within the sugar rings and along the phosphodiester backbone) are used to model the flexibility of each nucleotide, while the nucleotides are positioned with respect to a reference axis system using helical rotations and translations. All bond lengths are kept fixed. The junctions between successive (3′-monophosphate) nucleotides and the closure of the sugar rings are ensured by quadratic constraints on the O5′–C5′ and C4′–O4′ distances. This choice enables rapid and effective energy minimisation and strongly reduces the problems associated with local energy minima.

Conformational energies are calculated using the Flex force field, which includes parameters specifically developed for nucleic acids [20, 21]. In addition to atomic point charge electrostatics, Lennard-Jones, torsion and valence angle terms, this force field includes angular-dependent hydrogen-bonding contributions. Solvent and counterion electrostatic damping is treated with a sigmoidal distance-dependent dielectric function [20] coupled with reduced phosphate net charges ($-0.5e$). Previous studies from several laboratories have shown that this simple approach reproduces quite well both the conformational and mechanical properties of helical nucleic acids [22, 23, 24].

In order to model the presence of salt bridges we reduce the net charge of appropriate phosphate groups to zero by adding a fractional positive charge to each of the anionic oxygens, as described in our previous work [15, 16].

We calculate the DNA deformation energy associated with the passage from a free oligomer conformation to the DNA conformation adopted in the protein complex, using the program CONTACT recently developed in our laboratory. This program determines which DNA atoms belong to the protein–DNA interface and creates a set of spatial restraints which can be used to ''mold'' a free DNA and reproduce the conformation adopted by the protein binding site. In the present calculations, all atoms within 4.5 Å of protein atoms within the complex were included in the restraint file. In order to allow for experimental imprecision and for the limited conformational flexibility of the JUMNA model, restraints only became active when the DNA atoms fell beyond 0.5 Å of their relative positions within the protein complex.

3 Results and discussion

The list of protein–DNA complexes studied is given in Table 1. The length of the protein binding sites vary from

Table 1. Protein–DNA complexes studied grouped by family. (Complexes pdt009 and pdt025 have no equivalent Protein Data Bank, *PDB*, entry and have been coded as TBP1 and TBP2)

Family	NDB	PDB	Protein	Origin	Ref.
HTH	pdr010	1LMB	Lambda repressor	Virus	25
	pdr016	1LL1	Lambda repressor	Virus	26
	pdr009	1TRO	Trp repressor	Bacteria	27
	pdr004	2OR1	434 repressor	Virus	28
	pdr015	1PER	434 repressor	Virus	29
	pdr028	1BDH	Purine repressor	Bacteria	30
Homeodomain	pdt028	1YRN	MATa1/MATα1	Yeast	31
	pdr049	1AKH	MATa1/MATα1	Yeast	32
	pdr036	1MNM	MATα2/MCM1	Yeast	33
	pdt019	1OCT	OCT-1/POU	Human	34
HLH/bZIP	pdt047	1A0A	PHO4	Yeast	35
	pdr042	1HLO	Max	Human	36
	pdt064	1SKN	Skn-1	Nematode	37
TBP	pdt009	TBP1	TBP	Plant	2
	pdt025	TBP2	TBP	Plant	3
	pdt034	1CDW	TBP	Human	38
	pdt032	1VOL	TBP/TFIIB	Human	4
	pdr031	1AIS	TBP/TFIIB	Archea	39
	pdt036	1YTF	TBP/TFIIA	Yeast	40
Other	pdt040	1IHF	Integration host factor	Bacteria	1
	pdr032	1A3Q	NF-κB/P52	Human	41
	pdr012	1PAR	Arc repressor	Bacteria	42
	pdv001	2BOP	Bovine papillimavirus-1	Virus	43
	pdr021	2NLL	Retinoic acid receptor	Human	44

10 to 25 base pairs. The DNA oligomers from each of these complexes are shown in Fig. 1 and the presence of salt bridges involving cationic amino acid side chains (by a red dot) and of contacts with anionic amino acid side chains (by a black dot) are indicated. Note that salt bridges made to the single-stranded ends of DNA within some complexes were not taken into account in the present studies. While salt bridges to cationic amino acids clearly dominate in most cases, there is an exception in the case of helix–turn–helix (HTH) proteins (see later).

3.1 Phosphate neutralisation

For each complex, we energy-optimised the structure of the free DNA oligomer and then reminimised it after having neutralised the phosphate groups involved in salt bridges. The impact on DNA bending is shown in Table 2. On the basis of these calculations the 24 complexes were divided into two groups: those where neutralisation increases bending (top of Table 2) and those where it does not (bottom of Table 2). For the first group of oligomers, Fig. 2 demonstrates that the increase in bending occurs in the direction of curvature observed within the corresponding protein–DNA complex. Why some oligomers do not bend upon neutralisation is less immediately clear, but, as Table 2 shows, this is associated neither with the total number of phosphates neutralised (from 2 to 8) nor with the presence or absence of salt bridges outside the binding region which has been modelled.

These results can be better understood if we look more closely at the various families of protein complex that were treated.

3.1.1 Homeodomain

The first three complexes (1YRN, 1AKH, 1MNM) correspond to homeodomain systems. The first two complexes (MAT a1/MAT α2) are very similar, the DNA fragments have the same length and only differ in sequence in the central segment. They both present the same neutralisation pattern, with two patches of three contacted phosphates distant by roughly one turn of the double helix. In addition, for MAT a1, there are two neutralised phosphates in the middle of the binding site where the arm of the protein lies in the minor groove of the TAAT sequence. The disposition of the neutralised phosphates around the double helix is illustrated in Fig. 3. The third case (MAT α2, MCM1) is a ternary complex containing the MCM protein positioned at, and overlapping, the 5′ end of the oligomer. The DNA targets of these three complexes all bend considerably upon phosphate neutralisation with increases of 24°, 13° and 50°, respectively.

3.1.2 bzip/helix–loop–helix

The 1A0A complex contains a helix–loop-helix (HLH)/bZIP binding motif. The two α helices interact with successive major grooves, producing a very strong concentration of neutralised contacts. A similar situation is observed for 1HLO, the human Max protein complex. Although the binding sites in these complexes are much shorter than with the former family, neutralisation again increases bending by 27° and 35°, respectively. A modified protein with a single α helix in the major groove (1SKN) also increases bending, but only by 6°.

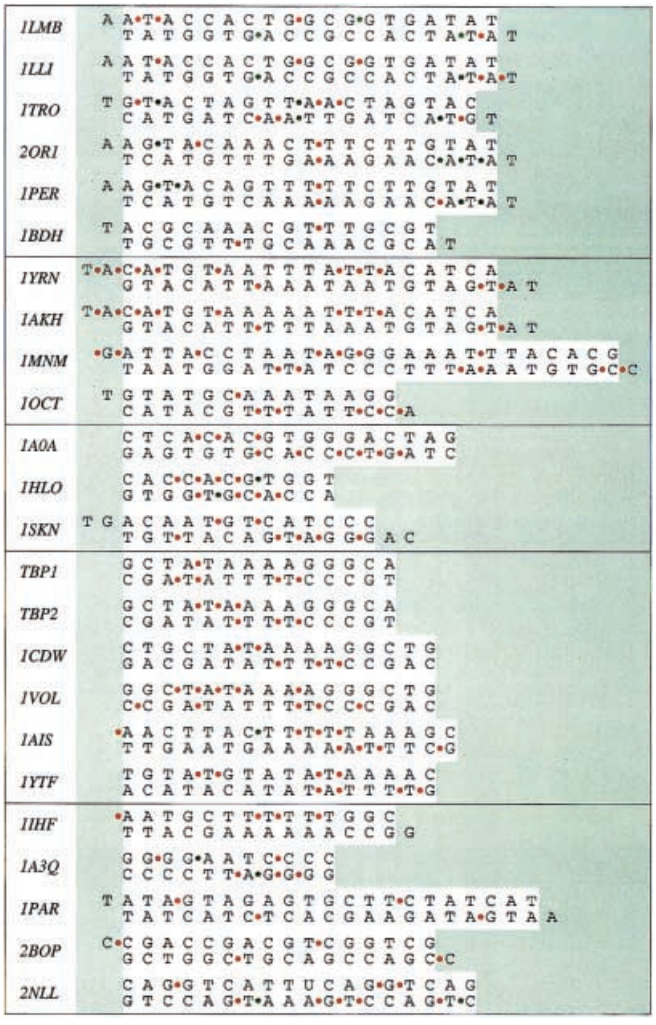

Fig. 1. DNA oligomers from the complexes studied. The *red dots* indicate the salt bridges involving cationic amino acid side chains and the *black dots* indicate the contacts with anionic side chains. The *white zone* denotes the oligomers modelled using JUMNA

3.1.3 TBP

The complexes TBP1, TBP2, 1CDW and 1VOL all correspond to complexes of TBP. This protein represents a special case since it interacts in the minor groove of DNA, producing a very strong deformation of the DNA helix and inducing an A-like conformation within the DNA target sequence [45]. The strong bending towards the major groove and away from the protein is mainly the result of junctions between the TATA-box conformation of the binding site and the B-like ends of the rest of the oligomer. Phosphate neutralisation does not reproduce the details of this very particular structure, but it does, however, lead to strong bending. The increases are almost identical for the first three complexes (24°–25°), and are roughly double this value (53°) for 1VOL, which has two proteins bound to the target oligomer. Moreover, it should be noted that the bending induced is directed towards the major groove (Fig. 2) although the protein is on the minor groove side of the duplex. Interestingly, the bending which occurs is not

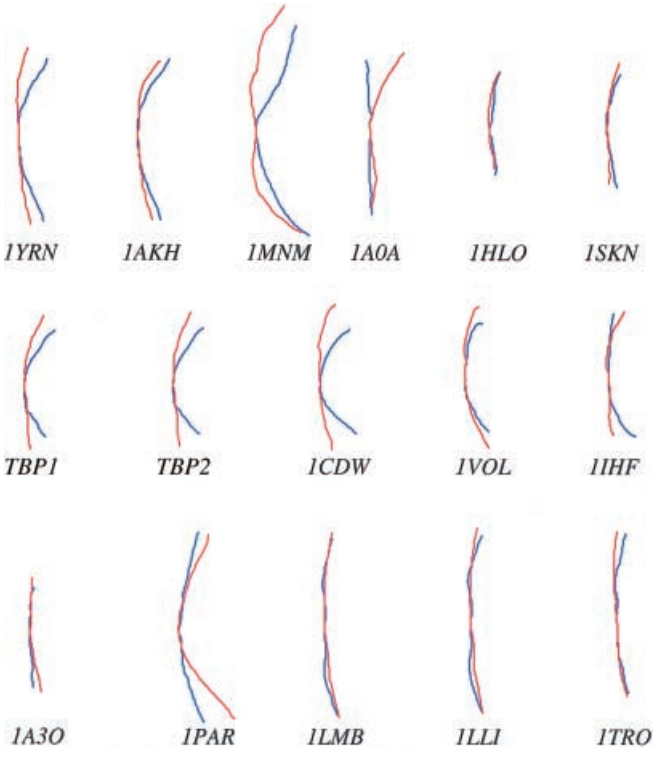

Fig. 2. Comparison of the helical axis calculated with CURVES for the DNA from crystallised complexes (*blue*) and with neutralised phosphates (*red*)

localised at the position of the salt bridges with the protein (in the central TATA motif), but rather involves the junction with adjacent GC base pairs.

3.1.4 IHF/NF-κB,P52/arc repressor

The IHF complex resembles TBP to some extent, although the protein now lies on the major groove side of the bent DNA, rather than on the minor groove side. The bending occurs almost precisely in a single plane (the root mean square of the points defining the helical axis with respect to the best plane is only 0.35 Å). It is limited to the first few nucleotides of the binding sequence. The neutralised phosphates occur further along the binding site within the run of six Ts and are not associated with local bending. Modelling the effect of the neutralisation reproduces this situation quite well and gives an overall curvature within 17° of that seen in the crystallised complex.

In the 1A3Q complex, DNA is contacted by five loop residues localised at the edges of large NF-κB and P52 proteins. The DNA curvature is rather gentle, with a bending angle just exceeding our limiting value of 20°. Phosphate neutralisation introduces small kinks which reproduce rather well those in the complex. The first is localised at the junction A5pA6, which is not neutralised, and the others are localised in the region between C8 and C10, which contains four salt bridges.

The recognition by beta strands is also observed in the arc repressor complex (1PAR). A good correlation is found between most important kinks in the helical axis and the position of neutralised phosphates.

3.1.5 Helix–turn–helix

The complexes 1LMB, 1LLI and 1TRO belong to the HTH family. The first two complexes correspond to the native and a mutant λ repressor bound to identical DNA sequences. These structures are gently curved, with the largest deformation occurring at C12pG13, which is not involved in a salt bridge. This step, however, lies between G10pG11, which makes a cationic contact, and G13pG14, which has an anionic residue nearby. The localisation of the bend may be the result of the greater flexibility of CpG steps [46]. The 1TRO (Trp repressor) complex is also gently curved, with two symmetric deformations occurring on either side of the central neutralised zone. These three HTH complexes have few salt bridges and an almost equal number of negatively charged residues. It is surprising that neutralisation of so few phosphates produces significant bending towards the protein in our modelling, although this is in line with the crystallographic results.

We now turn to the complexes where phosphate neutralisation did not lead to increased bending.

We begin with three further HTH complexes 2ORI, 1PER and 1BDH. For this group, the number of cationic and anionic side chains falling within 4 Å of phosphate oxygens is almost equal. On the basis of our present results, the interactions with cationic side chains do not appear to be involved in bending the DNA upon complexation (although it should be added that when the full protein–DNA complex was modelled in the case of the 434 repressor the correct curvature was found [47]). It is possible that salt bridges in these cases are mainly limited to positioning the recognition helices in the major grooves of DNA, as other authors have proposed [48, 49].

3.1.6 Ternary complexes of TBP

Two ternary complexes involving TBP with other transcription factors were studied. 1AIS is a bacterial equivalent of the human ternary complex with TFIIB (1VOL), while 1YTF is a ternary complex with TFIIA. The binding sequence of 1VOL, TATAAA, becomes $TU^{I}U^{I}AAA$ in 1AIS, where U^{I} signifies a 5-iodouracil. The absence of TpA steps in the latter sequence and the fact that the patches of salt bridges lie close together in the 1AIS complex may be linked to its lack of response to phosphate neutralisation. One should also note that its overall bending is much smaller than that of 1VOL. The 1YTF complex has three patches of salt bridges spaced by roughly half turns, which may explain why this structure also shows no significant bending upon neutralisation apart from a small kink at the 3′ end of the structure.

3.1.7 POU domains

The Oct complex 1OCT has three neutralised phosphates in the middle of the oligomer and two further on. Phosphate neutralisation only slightly increases local curvature. It has been noted [46] that Oct and Pit-1 POU domains make similar contacts to the DNA backbone and that these contacts are related to those of the 434 repressor, the λ repressor and 434 cro. There is, however,

Table 2. Comparison of DNA curvature in the crystallised complexes compared to that in the free oligomer before and after neutralisation of the phosphates involved in salt bridges. The curvature was calculated with CURVES [18] and is given in degrees

Family	Complex	Curvature			Salt bridges	Length (bp)	External salt bridges
		Complexed oligomer	Free oligomer	Neutralised oligomer			
Homeo-domain	1YRN	49	9	33	8	19	+
	1AKH	53	5	55	8	19	+
	1MNM	67	18	31	9	25	+
HLH/bZIP	1A0A	30	14	41	9	17	–
	1HLO	24	2	37	8	11	–
	1SKN	36	15	21	7	13	–
TBP	TBP1	83	16	41	5	14	–
	TBP2	93	16	40	6	14	–
	1CDW	95	22	47	5	16	–
	1VOL	107	17	70	10	16	–
Other	1IHF	77	9	60	3	14	+
	1A3Q	22	14	21	5	11	–
	1PAR	42	8	45	4	19	–
HTH	1LMB	36	14	33	4	19	–
	1LLI	37	14	26	4	19	+
	1TRO	32	14	22	5	18	–
HTH	2OR1	34	16	17	3	19	–
	1PER	27	14	6	3	19	–
	1BDH	40	2	5	2	16	–
TBP	1AIS	66	9	9	7	17	+
	1YTF	44	7	15	8	16	–
Homeo-domain	1OCT	39	16	10	5	14	+
Other	2BOP	34	12	6	2	16	+
	2NLL	37	34	28	8	18	–

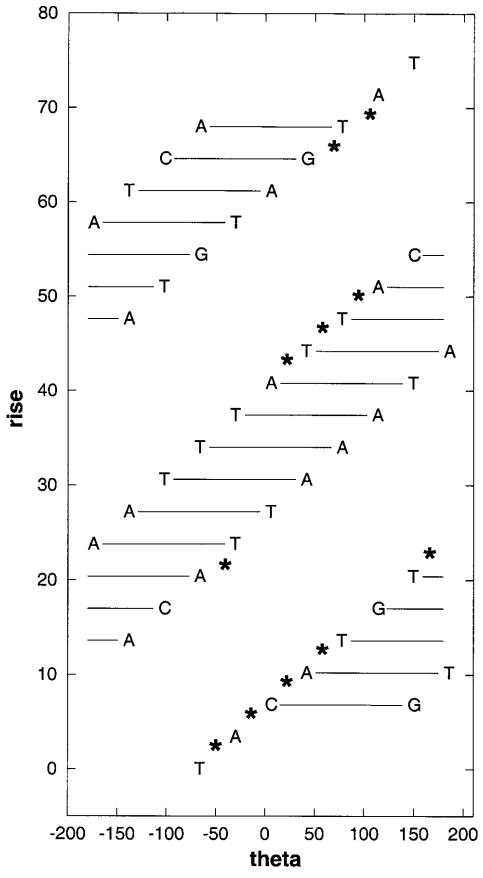

Fig. 3. Pattern of the neutralised phosphates for 1YRN shown for the regular helix in cylindrical coordinates. The phosphates neutralised are localised on one face of the double helix

an important difference between the prokaryotic proteins, which involve protein dimerisation, and the eukaryotic proteins, where the two binding domains are attached through a linker [48]. This difference may be linked to the observed change in the impact of salt bridges.

3.1.8 Beta barrel

2BOP is associated with only two salt bridges and it is thus not surprising that these interactions alone are not found to be responsible for the bending of DNA around the β-barrel binding motif.

3.1.9 Retinoid receptor

The last complex in Table 2, 2NLL (the retinoic acid receptor), has eight salt bridges, but also two contacts with anionic side chains. In the crystallised complex, the curvature is mainly localised at the ends of the oligomer, whereas the salt bridges are localised in the centre of the complex, where an α helix binds (and where the sequence contains an iodothymine which may hinder bending). Neutralisation in this case effectively does not increase bending.

This study of a range of different DNA-binding proteins suggests that phosphate neutralisation need not necessarily lead to bending. Its effect is a function of the positioning of the salt bridges and also of the underlying base sequence. Thus, although there are cases where axis kinks and salt bridges are clearly correlated, it is also possible for bending to occur outside the neutralised zone – notably, when this zone contains adenine or thymine runs (e.g. TBP, IHF). In several families, neu-

Table 3. DNA deformation energy upon protein binding (kcal/mol) before and after taking salt bridge neutralisation of phosphates into account. Each energy (*Tot*) is decomposed into its Lennard-Jones (*LJ*), electrostatic, valence angle and torsional components

DNA	Before phoshate neutralisation					After phosphate neutralisation					ΔE
	LJ	Elec	Ang	Tor	Tot	LJ	Elec	Ang	Tor	Tot	
1YRN	18	19	7	36	80	15	20	8	32	75	−5
1AKH	20	12	1	43	76	27	21	−1	34	81	5
1MNM	34	21	−8	23	70	42	39	−10	10	81	11
1A0A	50	2	7	24	83	55	22	8	13	98	15
1HLO	10	7	−9	20	28	15	−13	−10	10	2	−26
1SKN	4	24	5	18	51	5	19	5	11	40	−11
TBP1	39	20	−9	39	89	39	21	−11	35	84	−5
TBP2	36	20	−12	32	76	35	15	−14	30	66	−10
1CDW	31	28	−12	29	76	31	19	−12	26	64	−12
1VOL	43	29	−5	30	97	45	13	−9	27	76	−21
1IHF	27	20	3	12	62	36	22	0	9	67	5
1A3Q	3	25	−7	11	32	6	18	−8	−2	14	−18
1PAR	21	20	−5	36	72	33	37	−10	21	81	9
1LMB	16	14	−1	24	53	17	13	−1	22	51	−1
1LLI	27	13	6	28	74	28	12	7	25	72	−1
1TRO	8	7	−1	14	28	7	5	0	13	25	−3
2OR1	24	17	−1	10	50	10	18	4	10	42	−8
1PER	10	16	−5	19	40	24	28	−10	9	51	11
1BDH	35	1	−2	28	62	35	9	−3	24	65	3
1AIS	41	27	−14	64	118	60	39	−22	46	123	5
1YTF	36	18	−7	50	97	36	8	−4	43	83	−14
1OCT	9	19	−1	18	45	8	15	0	14	37	−8
2BOP	12	16	−4	29	53	10	15	−4	27	48	−5
2NLL	13	19	−4	38	66	25	7	−5	21	48	−18

Table 4. Influence of phosphate neutralisation on the components of the DNA deformation energy (kcal/mol)

DNA	LJ	Elec	Ang	Tor	ΔE
1YRN	−3	1	1	−4	−5
1AKH	7	9	−2	−9	5
1MNM	8	18	−2	−13	11
1A0A	5	20	1	−11	15
1HLO	5	−20	−1	−10	−26
1SKN	1	−5	0	−7	−11
TBP1	0	1	−2	−4	−5
TBP2	−1	−5	−2	−2	−10
1CDW	0	−9	0	−3	−12
1VOL	2	−16	−4	−3	−21
1IHF	9	2	−3	−3	5
1A3Q	3	−7	−1	−13	−18
1PAR	12	17	−5	−15	9
1LMB	1	−1	1	−2	−1
1LLI	1	−1	2	−3	−1
1TRO	−1	−2	1	−1	−3
2OR1	−14	1	5	0	−8
1PER	14	12	−5	−10	11
1BDH	0	8	−1	−4	3
1AIS	19	12	−8	−18	5
1YTF	0	−10	3	−7	−14
1OCT	−1	−4	1	−4	−8
2BOP	−2	−1	0	−2	−5
2NLL	12	−12	−1	−17	−18

tralisation has a very strong effect on bending and this includes the homeodomain proteins where neutralised patches occur at separations corresponding to the helical pitch and HLH proteins where α helices occupy consecutive major grooves (the single-helix SKN protein can be placed in this group).

TBP might seem to be a case apart since the protein lies in the minor groove; however, the neutralised phosphates are not in such different locations with respect to the major groove bend as the patterns seen in other complexes. It should also be noted that for TBP the salt bridges often reach around the helix and involve both anionic oxygens of the corresponding phosphate groups. For complexes where TBP is bound alone, neutralisation enhances bending in the correct direction, as if it were a major groove binder. In the case of ternary complexes with TFIIA or TFIIB the situation is clearly more complicated. It should be recalled that Elcock and McCammon [51] have proposed another explanation of TBP-induced bending on the basis of increased phosphate–phosphate repulsion on the minor groove side of the duplex owing to the presence of the low dielectric protein. This effect is, however, absent in our modelling.

3.2 DNA deformation upon protein binding

We now turn to the energetic aspects of phosphate neutralisation upon DNA deformation. Using a program recently developed in our laboratory (see Methodology) we are able to deform freely minimised DNA oligomers to reproduce the conformation adopted within a protein complex. The deformation energies calculated and their components were compared with those obtained after neutralising the phosphate groups involved in salt bridges. The results are given in Tables 3 and 4. Without phosphate neutralisation (left half of Table 3) the values of the DNA deformation energy vary from roughly 30 to 120 kcal/mol, in a manner that is qualitatively consistent with the degree of distortion caused by the protein. When phosphate neutralisation is taken into account (right half of Table 3) the deformation energy decreases in 15 out of 24 cases. However, it should be noted that the overall energy change should be treated with caution since we are considering an isolated DNA oligomer and thus are not taking into account the remaining electrostatic and dielectric effects of the bound protein.

If we limit ourselves to what is happening inside DNA, it is striking to note that phosphate neutralisation leads to a decrease in the torsion component of the deformation energy in virtually all cases (Table 4) and independently of the change in the total deformation energy. This implies that apart from its eventual role in enhancing DNA curvature, phosphate neutralisation consistently increases the flexibility of the DNA phosphodiester backbones and enables them to adjust more easily to the presence of a bound protein. Note, however, that this effect refers to the internal conformation of the backbones and not necessarily to the overall flexibility of the oligomer, since our earlier studies using normal-mode analysis pointed to a reduction in global deformations such as bending, stretching and twisting [15, 16].

4 Conclusions

We have carried out a systematic analysis of the role of salt bridges on protein complexes of known conformation. Several interesting points have come to light:

1. There are important differences between the complexes of prokaryotic HTH transcription factors and other proteins. The binding site of these complexes have roughly constant lengths, relatively few salt bridges and an almost equal number of negative side chains placed in the vicinity of phosphate oxygens. In our modelling, neutralisation of phosphates involved in salt bridges in most cases does not induce DNA curvature. It seems that in these complexes curvature has a different origin.
2. For certain DNA binding motifs, neutralisation of the phosphates involved in salt bridges indeed induces DNA curvature. This is the case for homeodomains where the neutralised phosphates form localised patches separated by roughly a helical turn and also for the HLH/bZIP family where two α helices grip DNA in consecutive major grooves. Interestingly sufficient neutralisation of salt bridges also induces curvature in most TBP complexes, despite the fact that the double helix is bent away from the protein.
3. Calculating the energy needed to deform DNA to its protein-bound conformation shows that the effect of phosphate neutralisation often decreases the overall deformation energy and virtually always leads to a significant decrease in the torsion energy component.

This suggests that, apart from bending DNA, salt bridges can also favour protein binding by increasing the flexibility of the phosphodiester backbones of the nucleic acid.

Acknowledgement This work was supported by Indo-French Centre for the Promotion of Advanced Research, project 1804–1.

References

1. Rice PA, Yang S-W, Mizuchi K, Nash H A (1996) Cell 87: 1295
2. Kim JL, Nikolov DB, Burley SK (1993) Nature 365: 520–527
3. Kim JL, Burley SK (1994) Nature Struct Biol 1: 638–653
4. Nikolov DB, ChenH, Haley ED, Usheva AA, Histake K, Lee DK, Roeder RG, Burley SK (1995) Nature 377: 119–128
5. Werner MH, Burley SK (1997) Cell 88: 733–736
6. Crothers DM (1993) Curr Biol 3: 675–676
7. Werner MH, Gronenborn AM, Clore GM (1996) Science 271: 778–784
8. Mirzabekov AD, Rich A (1979) Proc Natl Acad Sci USA 76: 1118–1121
9. Strauss J, Maher LJ (1994) Science 266: 1829–1834
10. Strauss-Soukup JK, Maher LJ (1997) Biochem 36: 10026–10032
11. Paolella DN, Yichin L, Fabian MA, Schepartz A (1997) Biochem 36: 10033–10038
12. Strauss-Soukup JK, Maher LJ (1998) Biochem 37: 1060–1066
13. Tomky LA, Strauss-Soukup JK, Maher LJ (1998) Nucleic Acids Res 26: 2298–2305
14. Strauss-Soukup JK, Maher LJ (1997) J Biol Chem 272: 31570–31575
15. Gurlie R, Ha Duong T, Zakrzewska K (1999) Biopolymers 49: 313–327
16. Gurlie R, Zakrzewska K (1998) J Biomol Struct Dyn 16: 605–618
17. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) J Mol Biol 112: 535–542
18. Lavery R, Sklenar H (1989) J Biomol Struct Dyn 6: 655–667
19. Mandel-Gutfreund Y, Schueler O, Margalit H (1995) J Mol Biol 253: 370–382
20. Lavery R, Zakrzewska K, Sklenar H (1995) Comput Phys Commun 91: 135–158
21. Hingerty B, Richie RH, Ferrel TL, Turner JE (1985) Biopolymers 24: 427–439
22. Lavery R (1994) Adv Comput Biol 1: 69–145
23. Fritsch V, Westhof E (1991) J Am Chem Soc 113: 8271–8277
24. Jayaram B, Swarminathan S, Beveridge DL, Sharp K, Honig B (1990) Macromolecules 23: 3156–3165
25. Beamer LJ, Pabo CO (1992) J Mol Biol 227: 177–196
26. Lim WA, Hodel A, Sauer RT, Richards FM (1994) Proc Natl Acad Sci USA 91: 423
27. Otwinowski Z, Schevitz RW, Zhang R-G, Lawson CL, Joachimiak AJ, Marmorstein R, Luisi BF, Sigler PB (1988) Nature 335: 321–329
28. Aggarwal AK, Ridgers DW, Drottar M, Ptashne M, Harrison SC (1988) Science 242: 899–907
29. Rodgers DW, Harrison SC (1993) Structure 1: 227–240
30. Glasfeld A, Schumacher MA, Choi KY, Zalkin H, Brennan RG (1996) J Am Chem Soc 118: 13073
31. Li T, Stark MR, Jonson AD, Wolberger C (1995) Science 270: 262–269
32. Li T, Jin Y, Vershon AK, Wolberger C (1998) Nucleic Acids Res 26: 5707–5718
33. Tan S, Richmond TJ (1998) Nature 391: 660–666
34. Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO (1994) Cell 77: 21–32
35. Shimizu T, Toumoto A, Ihara K, Shmizu M, Kyogoku Y, Ogawa N, Oshima Y, Hakoshima T (1997) EMBO J 16: 4689–4697
36. Schumacher MA, Choi KY, Zalkin H, Brennan RG (1994) Science 266: 763–770
37. Rupert PB, Daughdrill GW, Bowerman B, Mattews BW (1998) Nature Struct Biol 5: 484–491
38. Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK (1996) Proc Natl Acad Sci USA 93: 4862–4867
39. Kosa PF, Ghosh G, Dedecker BS, Sigler PB (1997) Proc Natl Acad Sci USA 94: 6042–6047
40. Tan S, Hunzinker Y, Sargent DF, Richmond TJ (1996) Nature 381: 127
41. Cramer P, Larson CJ, Verdine GL, Muller CW (1997) EMBO J 16: 7078–7090
42. Raumann BE, Rould MA, Pabo CO, Sauer RT (1994) Nature 367: 754–757
43. Hedge RS, Grossman SR, Laimins LA, Sigler PB (1992) Nature 359: 505–512
44. Fastinejad F, Perlmann T, Evans RM, Sigler PB (1995) Nature 375: 203–211
45. Guzikevich-Guerstein G, Shakked Z (1996) Nature Struct Biol 3: 32–37
46. Dickerson RE (1998) Nucleic Acids Res 26: 1906–1926
47. Ha Duong T, Zakrzewska K (1998) J Mol Biol 280: 31–39
48. Brennan RG (1992) Curr Opin Struct Biol 2: 100–108
49. Suzuki M, Yagi N and Gerstein M (1995) Protein Eng 8: 329–338
50. Jacobson EM, Li P, Leon-del-Rio A, Rosenfel MG, Aggarwal AK (1997) Genes Dev 11: 198–212
51. Elcock AH, McCammon JA (1996) J Am Chem Soc 118: 3787–3788