# Continuum Electrostatic Analysis of DNA Bending

by

**Robert Lester Murry**

B.A., Chemistry (1990)
Rice University

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of
**Master of Science in Chemistry**

at the

**Massachusetts Institute of Technology**

**September 1996**



Signature of Author .....
Department of Chemistry
September 9, 1996

Certified by .....
Bruce Tidor
Assistant Professor of Chemistry
Thesis Supervisor

Accepted by .....
Dietmar Seyferth
Chairman, Departmental Committee on Graduate Students

# Continuum Electrostatic Analysis of DNA Bending

by

**Robert Lester Murry**

Submitted to the Department of Chemistry on September 9, 1996
in partial fulfillment of the requirements for the degree of
Master of Science in Chemistry

ABSTRACT

The binding of proteins to DNA frequently causes significant bending or distortion of the nucleic acid crucial for biological function. The electrostatic energies and forces involved in such distortions are examined with a combination of a molecular mechanics force field and continuum solvation methods. The flexibility of DNA is included by way of its normal modes, which are calculated and described in detail for one structure. The electrostatic force caused by docking a hydrophobic model protein in the major and minor grooves, the force caused by introducing charge into this docked protein, and the force of substituting sets of phosphates with neutral isosters are all computed and analyzed. The strength of the phosphate-phosphate electrostatic interactions of free DNA in solvent is also assessed, and the effects that binding the ETS1 and SRY proteins to their operators has on these interactions is examined.

Thesis Supervisor: Bruce Tidor

Title: Assistant Professor of Chemistry

# Continuum Electrostatic Analysis of DNA Bending

## Introduction

Central to deoxyribonucleic acid's (DNA) role as carrier and controller of genetic information is the molecule's flexibility. The bending and twisting caused by protein binding and revealed with such techniques as electrophoretic gel-shift and cyclization assays plays a crucial role in regulatory sequence recognition and in the formation of multiple-component structures for control of transcription (van der Vliet and Verrijzer, 1993; Harrington, 1992). Despite the fact that detailed analysis of X-ray crystal and NMR structures of both free DNA and protein-DNA complexes (Young *et al.*, 1995; Suzuki and Yagi, 1995; Dickerson *et al.*, 1996; El Hassan and Calladine, 1996) has revealed many of the local, sequence-specific properties of DNA distortions — kinks, rolls, and intercalations — much basic information remains unknown about the energetics of complex formation and exactly how protein binding facilitates or forces the structural distortions observed.

For instance, the importance of the highly charged phosphate backbone has been recognized for some time. Rich was the first to propose that neutralization of phosphates on one side DNA wrapped around the histone proteins may be the source of bending (Mirzabekov and Rich, 1979), and these ideas were seemingly recently confirmed by Strauss and Maher (1994), who observed bending in DNA asymmetrically substituted with neutral phosphate analogs. Manning and coworkers have extensively studied the effects that counterions have in partially neutralizing the negative charges in DNA,

particularly with regard to bending (Manning, 1978; Manning *et al.*, 1989). However, Honig and others have used continuum models (for a review, see Honig and Nicholls, 1995) to reveal the effectiveness of aqueous solvent in screening the interactions of the phosphates with themselves and the rest of the DNA (Jayaram *et al.*, 1989; Friedman and Honig, 1992), possibly lessening their potential to cause curvature upon asymmetric neutralization.

The role of interactions across the binding interface is also unknown. While the surfaces' shape complementarity seems logical, the function of the frequently seen charge complementarity is more poorly understood. The protein-DNA hydrogen bonds and salt bridges might make binding favorable, particularly for certain DNA sequences, but the extent of their structural, bend-inducing effects is unknown and could be central in dictating the final DNA conformation. Unexpected interactions may also be quite significant. Several authors have argued that the displacement of solvent necessitated by binding protein could produce forces that significantly distort the DNA structure (Travers, 1995; Werner *et al.*, 1996), and a recent theoretical study combining continuum solvation with molecular dynamics has substantiated this (Elcock and McCammon, 1996).

In this work, we analyze the electrostatic energies and forces of DNA distortions caused by binding proteins. A 14-base pair length of DNA was studied in an all-atom model using the CHARMM parameters. First, the B-DNA structure was energy minimized using molecular mechanics, and its normal modes calculated. Then, solvent effects were introduced using a continuum model, and the strength of the electrostatic interactions between groups of atoms in the solvated DNA calculated using the DelPhi program. Finally, the electrostatic *force* caused by three types of hypothetical perturbations to DNA — docking an entirely hydrophobic model protein (whose only electrostatic effects

were due to the displacement of high-dielectric solvent by low-dielectric protein), introducing charge into that protein, and directly neutralizing selected phosphates without displacing solvent — were calculated. The forces, projected onto the lowest-frequency normal modes, were computed by finite-differences. The emphasis here is on analyzing the direction and origin of the forces, estimating their effects on bending and distorting the DNA, highlighting the differences between effects due to proteins that bind in the major versus the minor groove, and comparing the three types of perturbations to one another.

The energetics of two actual protein-DNA complexes in which the proteins bind in the minor groove and bend the DNA to compress the major groove were also analyzed. The effect on inter-phosphate repulsion due to the human *ets1* oncogene product ETS1 and the high mobility group (HMG) domain protein SRY to their operators was calculated via the scheme in Figure 1. Examining the differences due to binding straight versus bent DNA (Fig. 1, changes **1** and **3**) gave insight into how the inter-phosphate repulsion due to solvent displacement on binding is minimized by DNA distortion.

## Methods

### *The 14GC structure and normal modes*

The structure of the 14-base pair alternating sequence d[GCGCGCGCGCGCGC]$_2$ (referred to here as 14GC) was built in the B-DNA conformation using Quanta (version 4.0, distributed by Molecular Simulations Inc., 16 New England Executive Park, Burlington, MA 01803-5297). This length of DNA is one of the shortest in which the central phosphates have both minor and major groove partners. The terminal phosphates were replaced by hydroxyls, resulting in a model with 880 atoms and a total charge of –26*e*.

Molecular mechanics calculations were carried out with a modified version of the CHARMM computer program (version 24b1; Brooks *et al.*, 1983) using standard all-atom DNA parameters (version 6.4; MacKerell *et al.*, 1995). The DNA structure was subjected to energy minimization using a number of different treatments for the non-bonded electrostatic interactions. The default non-bonded scheme (with a switched cutoff between 8 and 12 Å and $\varepsilon=1$) was one treatment used, and $\varepsilon=1$, 4, $r$, and $4r$, all with no non-bonded cutoff were also used. Two thousand steps using the steepest descent algorithm were followed by steps using the ABNR algorithm until the RMS gradient was less than $10^{-5}$ kcal/(mol·Å). Normal-mode eigenvalues and eigenvectors for the $\varepsilon=4r$ minimum were calculated by diagonalizing the mass-weighted second-derivative matrix and normalized in the standard way (Wilson *et al.*, 1980; Brooks and Karplus, 1983). All matrix algebra was done using the numerical computation package Matlab (version 4.2c; distributed by The MathWorks Inc., 24 Prime Park Way, Natick, MA 01760).

Three tools were used to facilitate analysis of the normal modes (and forces, see below). First, trajectory files were created from the Cartesian versions of the eigenvectors to display the motion described by each mode. Structures were also plotted using the Ribbons program (version 2.0; Carson and Bugg, 1986; Carson, 1987) to illustrate structural changes due to each mode, and finally, the Dials and Windows program (version 1.0; Ravishanker *et al.*, 1989; Lavery and Sklenar, 1988) was employed to quantify motion into inter-base, intra-base, and helical axis-junction parameters.

*Electrostatic free energies and forces on DNA*

Solvent effects on electrostatic interactions were described in a continuum model by using solutions to the linearized Poisson–Boltzmann equation,

$$\nabla \cdot (\varepsilon(\vec{r}) \nabla \phi(\vec{r})) + 4\pi\rho^f(\vec{r})/kT - \varepsilon(\vec{r})\kappa^2(\vec{r})\phi(\vec{r}) = 0 \quad (1)$$

where $\vec{r}$ is the position vector, $\varepsilon(\vec{r})$ is the position-dependent dielectric constant, $\kappa(\vec{r})$ is the inverse Debye length ($\kappa = 1/\lambda = \sqrt{8\pi e^2 I/\varepsilon kT}$ where $e$ is the proton charge, $I$ is the bulk ionic strength, $k$ is Boltzmann's constant, and $T$ is the absolute temperature), $\phi(\vec{r})$ is the dimensionless electrostatic potential ($\phi = \varphi e/kT$, where $\varphi$ is the electrostatic potential), and $\rho^f(\vec{r})$ is the charge density due to fixed charges (i.e., the point charges of the DNA). This equation was solved numerically on a 65×65×65-point 3-dimensional grid using finite-difference techniques (Warwicker and Watson, 1982) and a modified version of the DelPhi computer program (version 3.0; Gilson and Honig, 1988; Gilson *et al.*, 1988; Sharp and Honig, 1990a). The DNA interior was assigned a dielectric constant of 4, the solvent was assigned 80, and an ionic strength of 0.145 M was used beyond a 2-Å Stern ion-exclusion layer (Bockris and Reddy, 1973; Gilson and Honig, 1987) outside the macromolecule and zero elsewhere. The total electrostatic energy was obtained from the expression,

$$\Delta G^{electrostatic\ total} = \frac{1}{2}\sum_{i=1}^{N} q_i \cdot \varphi(i) = \frac{1}{2}\sum_{i=1}^{N} q_i \cdot \left\{ \varphi_i^{solvation}(i) + \sum_{j \neq i}^{N} \left[ \varphi_j^{solvation}(i) + \varphi_j^{coulomb}(i) \right] \right\} \quad (2)$$

where $N$ is the total number of atoms, $q_i$ is the charge on atom $i$, $\varphi(i)$ is the total electrostatic potential at atom $i$, $\varphi_{i\,or\,j}^{solvation}(i)$ refers to the potential at point $i$ caused by the continuum solvent's reaction (frequently called the reaction field) to the charge of atom $i$ or $j$, and $\varphi_j^{coulomb}(i)$ is the potential at point $i$ calculated by the Coulomb's law (with $\varepsilon = \varepsilon_{in} = 4$) interaction with atom $j$.

CHARMM was used to calculate all $\varphi_j^{coulomb}$ terms, which excluded interactions between atoms sharing a covalent bond or bond angle. Trial runs in which 1–2 and 1–3 coulombic interactions were included produced similar overall conclusions (see Results); 1–4 electrostatic interactions were always included in the CHARMM energy function. DelPhi was used to calculate the

(sum of the) $\varphi_{i\,or\,j}^{solvation}$ terms, and the "grid energy" (a non-zero self-energy of each charged atom even when there is no dielectric boundary) was canceled by subtracting the total energies (or $\varphi(i)$'s where appropriate) from two DelPhi calculations with the same grid placement — usually [$\varepsilon_{in}$=4, $\varepsilon_{ext}$=4, ionic strength=0.0 M] from [4, 80, 0.145 M] to obtain the solvation energy. Dummy atoms were placed in a cube just outside the macromolecules to ensure the placement of the charged atoms on the grid for these two runs was identical.

The electrostatic force caused by a perturbation to the DNA was calculated as the difference between the forces acting on the DNA molecule before and after the change. Using protein docking as an example,

$$\bar{F}(protein\ docking) = \bar{F}(docked) - \bar{F}(free) \tag{3}$$

As the total electrostatic energy has two components (Eq. 2), so does each force:

$$\begin{aligned}\bar{F}(protein\ docking) &= \bar{F}^{coulomb}(docked) + \bar{F}^{solvation}(docked) - \bar{F}^{coulomb}(free) - \bar{F}^{solvation}(free) \\ &= \bar{F}^{coulomb}(protein\ docking) + \bar{F}^{solvation}(protein\ docking)\end{aligned} \tag{4}$$

The $\bar{F}^{coulomb}$ were evaluated by CHARMM as minus the gradient of the electrostatic energy with $\varepsilon$=4, then projected onto the normal modes. Each $\bar{F}^{solvation}$ was obtained by finite differences (Gilson *et al.*, 1993) using DelPhi along the normal-mode directions,

$$\bar{F}^{solvation} = -\frac{\partial \Delta G^{solvation}}{\partial \bar{r}} \cong \frac{\Delta G^{solvation}(x-\Delta x) - \Delta G^{solvation}(x+\Delta x)}{2\cdot\Delta x} \tag{5}$$

where $\Delta x$ is a small step. Steps along the normal modes were taken such that the maximum any atom moved was 0.05 Å. The response to the force caused by a perturbation was calculated as the vector sum of the components of the force along each normal mode divided by the normal mode's eigenvalue (see Results).

Use of the linearized Poisson–Boltzmann equation permits the calculation of the interaction energy between two atoms or groups of atoms $I$ and $J$:

$$\Delta G_{IJ}^{electrostatic} = \frac{1}{2}\sum_{i\in I} q_i \cdot \sum_{j\in J, j\neq i}\left[\varphi_j^{solvation}(i) + \varphi_j^{coulomb}(i)\right] \tag{6}$$

For calculations of this sort, the *I* group was charged and the solvation component of the potential at the *J* group computed by DelPhi (by the difference of two calculations, see above) and added to the coulombic interaction. It was found that for distant groups (separation >10 Å with the grid spacings used here), the difference in total potential at *J* between this two-step procedure and the result from a single DelPhi calculation was negligible. The effective dielectric constant between groups was calculated as the coulombic interaction with $\varepsilon=1$ divided by the total interaction energy from the above equation (with $\varepsilon_{in}=4$, $\varepsilon_{ext}=80$, etc.).

Trial calculations were also performed solving the non-linear Poisson–Boltzmann equation (Sharp and Honig, 1990b) using the expression for the total energy (Luty *et al.*, 1992),

$$\Delta G^{electrostatic\ total} = \frac{1}{2}\sum_{i=1}^{N} q_i \cdot \varphi(i) + \int\left[\frac{q^m(\vec{r})\cdot\varphi(\vec{r})}{2} + \Delta\Pi(\vec{r})\right]dV$$
$$= \frac{1}{2}\sum_{i=1}^{N} q_i \cdot \varphi(i) + I \cdot V_m \cdot \sum_{\substack{ion\text{-}accessible \\ grid\ points\ m}}\left[e\cdot\varphi(m)\cdot\sinh\phi(m) - kT(2\cosh\phi(m)+2)\right] \quad (7)$$

where the first term in each line is simply Eq. 2 (this is the only term when the linear equation is used), $q^m(\vec{r})$ is the position-dependent net mobile-ion charge concentration, $\Delta\Pi(\vec{r})$ is the excess osmotic pressure, $V_m$ is the volume per grid point in the DelPhi run, and $\varphi(m)$ and $\phi(m)$ are the potential and dimensionless potential at the grid point *m*. These calculations produced results for the forces along the normal modes very similar to those from the linear equation, as did other calculations performed at zero ionic strength, where the linear and non-linear equations are identical (see Results).

For all DelPhi calculations, the reported results represent the average of 11 random rotations and translations of the DNA on the grid, and error bars are twice the standard deviation of the mean, which reflects uncertainty due to the granularity of the grid but not any systematic errors that may be present. As the

finite-difference force calculations require substantial accuracy, DelPhi was modified to run completely in double precision (64-bit representation of all floating point numbers). The dielectric boundary was the molecular "contact" surface (Richards, 1977), i.e., the surface contacted by a probe of radius 1.4 Å rolled across the van der Waals surface, and it was calculated using a program written for this purpose (C. V. Sindelar and B. Tidor, unpublished) according to the method of Connolly (1983a, b). A focusing scheme was employed where the result of a calculation with screened coulombic boundary conditions in which the DNA filled 23% of the grid was used to obtain the boundary conditions for a 92% fill calculation (Klapper *et al.*, 1986). However, some non-linear calculations did not converge at 23% fill, so a 46% fill initial calculation was used instead; for cases where both converged, the difference in total energy for the 92% fill computation was found to be less than 0.1 kcal/mol. A smoothing algorithm was employed to better represent the dielectric boundary on the grid lattice (Mohan *et al.*, 1992; Davis and McCammon, 1991). This involved assigning dielectric lines that cross the molecular boundary a value intermediate between $\varepsilon_{in}$ and $\varepsilon_{ext}$ according the to the prescription,

$$\varepsilon = \frac{\varepsilon_{ext}\varepsilon_{in}}{\alpha\varepsilon_{in} + (1-\alpha)\varepsilon_{ext}} \tag{8}$$

where $\alpha$ is the fraction of the grid line that was in solvent. In trial runs without smoothing, the center of each grid line determined which of the two dielectrics was assigned to it. For all calculations on 14GC, the grid spacing was 0.85 Å/grid-unit.

To test the numerical calculation method implemented here, comparisons were performed with a simple model that was also solved analytically. Values for a system of two charges, $q_1$ and $q_2$, separated by a distance $x$, with $q_1$ constrained to be at the center of a low-dielectric sphere of radius $R$ and dielectric

$\varepsilon_{in}$, in a solvent of dielectric $\varepsilon_{ext}$ and zero ionic strength, were obtained from Kirkwood (1934) and Davis and McCammon (1990).

$$\begin{aligned}\Delta G^{\text{electrostatic total}} &= \frac{q_1 q_2}{\varepsilon_{in} x} + \frac{(q_1+q_2)^2(\varepsilon_{in}-\varepsilon_{ext})}{2R\varepsilon_{ext}\varepsilon_{in}} \\ &\quad + \frac{{q_2}^2}{2\varepsilon_{in}} \cdot \sum_{n=1}^{\infty} \frac{x^{2n}(n+1)(\varepsilon_{in}-\varepsilon_{ext})}{R^{2n+1}(\varepsilon_{ext}(n+1)+n\varepsilon_{in})} \qquad \text{for } x < R \\ &= \frac{q_1 q_2}{\varepsilon_{ext} x} + \frac{{q_1}^2}{2R\varepsilon_{in}}\left(\frac{1}{\varepsilon_{ext}} - \frac{1}{\varepsilon_{in}}\right) \\ &\quad + \frac{{q_2}^2(\varepsilon_{ext}-\varepsilon_{in})}{2\varepsilon_{ext}} \cdot \sum_{n=1}^{\infty} \frac{n \cdot R^{2n+1}}{x^{2n+2}(\varepsilon_{ext}(n+1)+n\varepsilon_{in})} \qquad \text{for } x > R\end{aligned} \tag{9}$$

Notice the first term is these equations is the Coulomb's law interaction energy between the charges. The force on each charge is of equal magnitude and opposite direction; it was obtained for the off-center atom as $q_2 \cdot \vec{E}$ (where $\vec{E}$ is the electric field vector), while the force on the central atom has additional terms because it "creates" the dielectric boundary (Davis and McCammon, 1989; Sharp, 1991). For either charge, the magnitude of the force is therefore:

$$\begin{aligned}F_{\text{total}}^{\text{electrostatic}} &= \frac{q_1 q_2}{\varepsilon_{in} x^2} - \frac{{q_2}^2}{\varepsilon_{in}} \cdot \sum_{n=1}^{\infty} \frac{n(n+1)(\varepsilon_{in}-\varepsilon_{ext})x^{2n-1}}{R^{2n+1}(\varepsilon_{ext}(n+1)+n\varepsilon_{in})} \qquad \text{for } x < R \\ &= \frac{q_1 q_2}{\varepsilon_{ext} x^2} + \frac{{q_2}^2(\varepsilon_{ext}-\varepsilon_{in})}{\varepsilon_{ext}} \cdot \sum_{n=1}^{\infty} \frac{n(n+1)R^{2n+1}}{x^{2n+3}(\varepsilon_{ext}(n+1)+n\varepsilon_{in})} \quad \text{for } x > R\end{aligned} \tag{10}$$

For DelPhi calculations in which $q_2$ was outside the sphere, the grid energy for each charge was canceled separately using three DelPhi calculations for each solvation energy.

*Phosphate-phosphate repulsions in protein-DNA complexes*

The NMR structures of the ETS1 (PDB code 1stw; Werner *et al.*, 1995a) and SRY (1hry; Werner *et al.*, 1995b) protein-DNA complexes were obtained from the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977) and used for the bent DNA-protein complex structures in Fig. 1. The structure of the DNA alone was extracted from these complexes and used for the bent-DNA geometry.

Coordinates for the B-DNA structures for each operator were generated using Quanta as they were for 14GC. For the geometry of the hypothetical unbent DNA-protein complex, the B-DNA was placed contacting the protein in a manner analogous to the (center of the) bent complex. Specifically, the central nucleotides (for ETS1, nucleotides 208-210; for SRY, 4, 5, 12, and 13) were aligned by minimizing the RMS deviation with the corresponding nucleotides in the B-DNA, while the protein remained unmoved. This caused some overlap of the protein and the DNA, but since only intra-DNA interactions were calculated here, this was not problematic.

The CHARMM all-atom radii for the protein atoms were used (version 6.4; see MacKerell *et al.*, 1995), but no atoms in the protein were assigned charge and the protein was assigned a dielectric of 4. The phosphate-phosphate interaction energies for each structure were calculated using DelPhi and Eq. 6 by charging each phosphate in turn. There were 14 phosphates in the SRY complex and 32 in the ETS1 complex. A single set of dummy atoms (arranged in a cube outside the molecules) was used for each complex to ensure reproducible placement on the grid. The final grid spacings were 1.11 Å/grid-unit for the ETS1 complex, and 0.87 Å/grid-unit for SRY.

## Results

### *DNA minimization*

Energy minimizations on a 14-base pair length of DNA ($d[GCGCGCGCGCGC]_2$) using the CHARMM all-atom parameters (MacKerell *et al.*, 1995) were performed starting from a B-DNA structure but varying the electrostatic scheme used. The default electrostatic non-bonded interaction scheme (a switched cutoff between 8 and 12 Å and $\varepsilon$=1) was tried, as well as $\varepsilon$=1,

$\varepsilon=4$, $\varepsilon=r$, and $\varepsilon=4r$, all with no cutoff. The $\varepsilon=4$ and $\varepsilon=4r$ calculations were done for consistency with the continuum model used here, where the interior dielectric was 4. Only the $\varepsilon=4r$ minimization reached a structure similar to the B-DNA double helix (mass weighted RMS for all atoms = 1.86 Å, for the central eight nucleotides = 1.30 Å). The other calculations either reached a minimum whose RMS deviation from the starting structure was greater than 2.5 Å and had considerable distortions from a regular double helix, or were terminated after the RMS deviation from the starting structure exceeded 10 Å. Similar minimizations performed on a d[ATATATATATATAT]$_2$ B-DNA structure with $\varepsilon=4r$ and no non-bonded cutoff produced a minimum with an RMS deviation from B-DNA of 2.0 Å, while the other energy functions again produced larger deviations.

The 14GC structure from this $\varepsilon=4r$ energy minimization was used for subsequent analysis and normal-mode calculations. The RMS gradient was less than $10^{-5}$ kcal/(mol·Å). The minimum was virtually symmetric, with a pseudo-two-fold rotation axis running through the central major and minor grooves, as shown in Figure 2. The overall structure was somewhat shorter than idealized B-DNA (distance between terminal guanine N1 atoms = 41.2 vs. 44.1 Å). The major and minor grooves were also slightly wider than idealized B-DNA (average and standard deviation of six central major groove P-P distances (e.g., 6A-7B) = 18.7±0.1 vs. 17.6 Å; central ten minor groove P-P distances (e.g., 9A-10B) = 12.3±0.3 vs. 11.5 Å). The twist (rotation of one base pair relative to the other around the helical axis) per base-pair step was near the idealized B-DNA value of 36° except near the termini, which were slightly overtwisted.

*Normal modes of DNA*

At this 14GC minimum, the normal modes of the DNA were computed by diagonalizing the mass-weighted second derivative matrix calculated with the same CHARMM parameters and $\varepsilon=4r$ with no non-bonded cutoff. The frequencies of the 2640 normal modes are shown in Figure 3. There were no negative eigenvalues, indicating that a local minimum rather than a saddle point was located on the energy surface. The lowest six frequencies were essentially zero (less than 0.01 $cm^{-1}$) and correspond to the three overall rotations and three overall translations.

A brief description of the motion corresponding to each of the 20 lowest-frequency modes is given in Table 1. Motion along these low-frequency modes corresponded to global deformations, such as twisting, bending, and rocking large sections of DNA. It should be noted that motion along each mode is energetically equivalent in both the forward and backward directions; that is, an infinitesimal bend toward the central major groove along mode 7 requires the same energy as a bend away from it and towards the minor groove. For convenience, Table 1 describes only one of these directions, which will arbitrarily be called the forward or positive here. The motions for two of the modes — mode 10, a twisting mode, and mode 16, which opened and closed the whole minor groove — are shown in Figure 4. In both cases, as for all low-frequency modes, the bases remained well stacked and paired, with some flexibility arising from propeller-twisting (rotation of one base relative to its partner along their long axis) and base buckling (bending of a base pair at a hinge perpendicular to and through the base-pairing hydrogen bonds), while the backbone moved more freely. It is also interesting that, although a uniform cylinder would have two degenerate "bending" modes, in all the low-frequency modes of 14GC bending

occurred along an axis passing through the major and minor grooves, and never along the axis perpendicular to this and the helical axis.

The lowest frequency mode, numbered 7 because it has the next smallest eigenvalue after the six translations and rotations, corresponded to a smooth bend — not a kink — along the whole length of the DNA. Motion along the mode compressed (or expanded) the minor groove and expanded (or compressed) the major groove in the center of the structure. The smooth, un-kinked bend allows a smaller loss of base-stacking interactions per degree of bend (Calladine and Drew, 1992). As this mode represents the easiest way to deform the 14GC minimum, the lowest energy internal motion, it will figure prominently in the subsequent analysis. Changes in the inter-base pair roll (tilting of one pair along its long axis relative to the other) for motion along the forward direction of this mode revealed large negative rolls in the center and smaller positive rolls near the termini; the sign change is due to the fact that positive roll indicates compression toward the *local* major groove. This correlation between roll and overall bend in DNA was first suggested by Dickerson *et al.* (1983). However, while other modes involving bends had large rolls at the appropriate points of the bend in the structure, the converse was not true — a large roll did not always create a bend (mode 10, for instance, had large rolls). This highlights the difficulty in using any single inter-base pair or similar parameter to describe bending in DNA structures. Furthermore, concerted motion of the sugar-phosphate backbone was usually much larger than the local distortions of the bases and more often gave — by visual inspection at least — a better indication as to the type of motion the mode or combination of modes involved.

A number of inter-base pair parameters changed in an alternating fashion for several of the modes. That is, certain parameters changed one direction for

CG base steps and in the opposite for GC steps. This was true, for example, for rise (increasing the base-pair separation along the helical axis) in mode 7 and for twist in mode 10. This alternation ensured that the GC (purine-pyrimidine) steps remained better stacked in the course of the twisting or bending than the CG steps, presumably because of the greater overlap and hence higher base-stacking van der Waals energy of GC steps. Note that this difference in overlap exists in idealized B-DNA as well. Statistical analysis of bending in crystal structures reveals a similar flexibility of pyrimidine-purine steps relative to purine-pyrimidine steps (Suzuki and Yagi, 1995; Gorin *et al.* 1995).

*Electrostatic interactions in free DNA*

Continuum electrostatic methods were used to calculate the effect of solvent on the electrostatic interactions in the 14GC structure. The solvent-screened, effective electrostatic interactions between groups of atoms (phosphates, riboses, and bases) were calculated according to Eq. 6. These values were reduced from their Coulomb's law magnitudes due to the shielding effects of the high dielectric solvent and the ionic strength of 0.145 M.

As an example of typical DNA values, the strongest total interactions between the three groups of nucleotide 7A and the other groups of the 14GC structure are given in Table 2. Note that the charge distribution in this CHARMM parameter set puts a charge of –1.2 on each phosphate, +0.2 on each ribose and 0.0 on each base. Only groups within the immediate vicinity of the nucleotide had significant interactions with it. The strongest interaction was between the bases of the 7A-8B base pair; this was due not only to the highly favorable arrangement of hydrogen bonds but also to the relative solvent *inaccessibility* of the bases, which limits solvent screening. For more solvent-exposed groups, particularly the phosphates, the solvent was very effective at

screening the interactions. The largest interaction energy between two phosphate groups was only about 1.0 kcal/mol, even though the strict Coulomb's law interaction with a dielectric of four was about 18 kcal/mol. Hence the phosphate-phosphate interactions *did not* dominate the electrostatic structure of free DNA in solvent, and stronger interactions were observed for groups closer to one another that were less solvent-exposed, even though these groups carried a smaller formal charge. Note that, with the exception of the interactions within the 7A nucleotide itself, the solvation energy had the opposite sign and was of lesser magnitude than the coulombic energy.

To investigate the phosphate-phosphate repulsions further, the interaction energies between the phosphate groups throughout the DNA structure were computed and are listed in Table 3. Part *a* shows the coulombic interaction between pairs of phosphates, which was maximal (about 17-18 kcal/mol) for nearest neighbors on the same strand. Coulombic interactions across the minor groove (e.g. 8A-11B or 9A-10B) were about 9.5-10 kcal/mol, while those across the major groove (5A-8B or 6A-7B) were about 6.4 kcal/mol. The solvent-screened total interactions are listed in part *b*. Again these show very effective screening; the maximum total interaction was still between nearest neighbors on the same strand but only about 1.0 kcal/mol, followed by interactions across the minor groove (0.25 kcal/mol), then interactions of second neighbors on the same strand (0.15 kcal/mol), and then interactions across the major groove (0.1 kcal/mol). These small interaction energies persisted even at zero ionic strength, as shown in part *c*, where the nearest-neighbor interactions are about 1.4 kcal/mol, the minor-groove interactions about 0.55 kcal/mol, and the major-groove interactions about 0.3 kcal/mol.

Effective dielectric constants are a measure of the strength of solvent screening between groups. For pairs of phosphates in the 14GC structure at zero

ionic strength, the effective dielectrics ranged from about 50 for nearest neighbors on the same strand to above 100 for phosphates on opposite sides of the DNA. Values across the minor groove were about 70 and across the major groove were about 85. The effective dielectric (without salt) acting between phosphate 4A and the other atoms in the 14GC structure is shown in Figure 5. Atoms near the phosphate had low effective dielectrics while those far away or on the opposite side of the molecule had higher ones. The effectiveness of solvent screening depended not only on the solvent exposure, but on the precise geometry between the two groups and the dielectric boundary. Moreover, neither the length nor the solvent exposure of an imaginary line drawn between a pair of groups gave an indication of the value of the effective dielectric between them. For example, the effective dielectric between atoms separated only by low-dielectric interior of the DNA (e.g., phosphate 4A with phosphorus atom 11B or 12B, near it but on the other strand in Fig. 5) was higher than some more distant atoms separated by intervening solvent (e.g. phosphate 4A with phosphorus atom 6B or 7B, across the major groove from it). Finally, salt only *increased* the effectiveness of screening, particularly for long-range interactions, and the effective dielectrics between distant or exposed groups in the presence of salt was in the hundreds or thousands (not shown).

It may seem strange that the effective dielectric constant predicted by continuum solvation methods for many interactions in DNA can rise above 80, the value of the external dielectric. This is possible even for spherical systems (like the one solved analytically here) when the charges are very close to the dielectric boundary, where they interact favorably with (at least) their own reaction field. Honig and coworkers have previously found this effect (Jayaram *et al.*, 1989; Friedman and Honig, 1992; Sharp *et al.*, 1990). The effective dielectrics in DNA calculated here match the values in these studies well, and for the cases

here where the effective dielectric was greater than 80, the total interaction energy between groups was quite small, less than 0.1 kcal/mol.

Although the above results were for the 14GC minimum structure, calculations performed on the idealized B-DNA conformation revealed similar phosphate-phosphate interaction energetics. The only significant difference was that the same-strand, nearest-neighbor phosphate-phosphate interaction energy was about 15% higher in the B-DNA structure, because the groups were slightly closer (this strong repulsion may have been a driving force in the minimization). But the effectiveness of solvent screening resulted in a small magnitude for cross-groove inter-phosphate interactions (0.25 kcal/mol for the minor groove; 0.1 kcal/mol for the major), and the only interactions greater than 0.3 kcal/mol were again between nearest neighbors on the same strand.

*Accuracy of finite-difference forces*

While the efficacy of using numerical, grid-based methods like DelPhi to calculate electrostatic free energies has been well documented, their ability to calculate electrostatic forces on atoms by finite differences is less clearly established. To prove the feasibility of this, results from a DelPhi finite-difference force calculation were compared to the exact results for a simple system that was solved analytically (Eq. 9 and 10).

The system consists of two charges of magnitudes $0.3e$ and $-0.3e$, one of which is centered in a low-dielectric sphere of radius 6 Å and the other a distance $x$ away. The force on each atom was calculated numerically by taking small (±0.001 Å) steps in three orthogonal directions (none of which was on the line connecting the atoms), calculating the total electrostatic energy for each geometry, and then computing the finite-difference force using Eq. 5. Note that moving the central atom also moves the dielectric boundary since it remains at

the center of the sphere. As shown in Figure 6, the finite-difference forces reproduced the analytical values well for this system. The direction of the finite-difference force was on the line connecting the atoms, as the analytical solution demands. Part *a* shows the excellent agreement in magnitude when fine grid spacing was used, and *b* shows a good match even at coarser spacing than used here for DNA. Parts *c* and *d* show the effects of not using smoothing (see Methods) to calculate the solvation energies for the finite-difference forces. While in each case the force on the off-center atom was well calculated, the force on the central atom was wrong unless smoothing was used. The representation of the dielectric boundary without smoothing on the discrete grid is very rough, and hence the small changes in the boundary's location caused by the ±0.001 Å steps go unnoticed by the program (unless a new grid line happens to get incorporated or left out of the moving sphere). The force computed without smoothing for small step sizes, then, is only the $q \cdot \vec{E}$ term of the force, which is correct for the off-center atom since it is the only term for that atom, but incorrect for the central atom, the force on which has other terms because it creates the dielectric boundary (Davis and McCammon, 1989; Sharp, 1991). In DNA there are no free charges (like $q_2$), so smoothing was used for all subsequent calculations.

*Forces due to docking a neutral protein to DNA*

The force caused by binding a "generic" protein — a 10-Å radius low-dielectric sphere with no partial or formal charges — to the DNA structure was computed by finite differences along each low-frequency normal-mode direction. Two dockings were computed, one with the protein in the major groove and one with it in the minor groove, as shown in Figure 7. The center of the protein was 10 Å from the helical axis in either case. The change in the total electrostatic

energy caused by the change in shape of the dielectric boundary was computed for the 14GC structure. The DNA was then moved a small step along each normal mode and the above change in energy re-computed. The finite-difference force along each normal mode was then obtained using Eq. 5. Note that the coulombic component of the force (Eq. 4) for these dockings was exactly zero, since only a change in the shape of the boundary was made.

As an example, the data for calculation of the force for the minor-groove docking along mode 12 is shown in Figure 8. In part *a*, the change in energy caused by the docking is plotted for various forward and backward step sizes along the mode. The steps were taken such that the maximum distance any atom moved was controlled (and is the abscissa in Fig. 8), which required weighting the normal-mode step by a factor. Dividing these energy changes by the step size (via the weighting factor) gave the force along mode 12, in units of energy per unit-step along the mode. The value of this force for various step sizes is shown in part *b*. The finite-difference force was constant for a range of step sizes, and similar consistency was found for the force along other modes, so a maximum single-atom movement of ±0.05 Å was used for all subsequent calculations.

The force calculated in this manner along the each of the lowest 50 modes (excluding the translations and rotations) for the major- and minor-groove dockings is shown in Figure 9a. Many of the modes had no significant component of the force in their direction, while a few modes had large projections. The minor-groove docking produced significant projections along modes 7, 9, 12, 16, and 18, all with negative magnitudes. Thus, docking the hydrophobic sphere in the minor groove caused a force along the following combination of directions: bending away from the minor groove (mode 7), closing the major groove at the center and bending the termini away from the

helical axis (mode 9), opening the minor groove at the center and closing the major groove (mode 12), opening the whole minor groove (mode 16), and opening the minor and major grooves at the center (mode 18). All of these force projections act to *open the minor groove* in the middle of the DNA, where the protein was docked. The corresponding force for the major groove was smaller in magnitude than that for the minor groove and generally had the opposite sign along each mode, with significant components along modes that open the central major groove.

The vector sum of these projections clearly revealed the nature of the protein-docking force. For either docking, the force opened the groove where binding occurred by pushing apart the DNA backbone near the protein, moving it nearly parallel to the helical axis, as shown in Figure 10. The backbone atoms near the termini and the bases had little or no component of the force, in comparison. The displacement of solvent around the DNA by the protein lessened the solvent's screening effect, increasing in particular the strength of the phosphate-phosphate repulsions. These repulsions, as well as the favorable energy to be gained by better solvating the backbone, generated the drive to push the two strands of the backbone apart, opening the appropriate groove. Because more backbone atoms were buried by docking in the minor groove than in the major (Fig. 7b vs. 7d), the force caused was correspondingly greater.

Because the potential energy near equilibrium may be approximated to first order as harmonic along each of the normal modes, the initial response of the DNA to the above forces may be grossly approximated by dividing each projection by the mode's eigenvalue (~ frequency$^2$). This is equivalent to treating the DNA as though it were held together by springs (pulling in the normal mode directions), which counter the docking force by expanding or contracting until the force along each mode is zero. Figure 9b shows the responses to the minor

and major groove dockings along the lowest 50 modes. As the frequencies rise with increasing mode number, only the very lowest modes had significant projections in the response. For the major groove, modes 7, 9, 10, and 12 had positive projections, while the minor groove had nearly the opposite, with modes 7, 9, 12 and a few others having negative magnitudes. In both cases, mode 7 dominated the response, leading to a clear conclusion: the docking of the protein in a groove of the DNA and consequent displacement of solvent created a force which caused the DNA to bend away from the protein. This follows from the fact that bending was the easiest way (lowest frequency) to open the groove, relieving some of the increased phosphate-phosphate repulsion and better solvating the backbone. The minor groove docking caused a stronger magnitude response than the major, about 2.5 times as much.

Figure 11 shows the vector sum over the lowest 50 modes of the projections of these responses. Each groove opened in the middle of the DNA, primarily by bending away from the hydrophobic sphere. This is most easily recognized at the termini in parts *b* and *d*. The severity of the distortions (not the magnitude in Fig. 11, which is arbitrary for display) for the minor groove docking reveals some of the potential for conformational change in DNA when the screening effects of solvent are reduced. That is, the primary effect of propagating the relatively local force (that opens either the major or minor groove) through the somewhat stiff DNA structure was to produce a more non-local bend.

The effect of salt on the force for the minor-groove docking was also examined. The projections of the force along the lowest 50 modes were calculated under four conditions: at zero ionic strength, using the linearized Poisson–Boltzmann equation at 0.145 M salt concentration, using the non-linear equation at 0.145 M but taking (incorrectly) only the first term in Eq. 7, and using

the correct total energy for the non-linear equation at 0.145 M. As shown in Figure 12, the force along each mode under each of these conditions was nearly identical. Trial calculations for other forces produced similar agreement. Hence, the use of the linearized Poisson–Boltzmann equation was valid for the calculation of these forces along the lowest modes, since the effect of salt is minimal. Previous studies have found other contexts in which the use of the linearized equation was justified for DNA (Friedman and Honig, 1992; Zacharias *et al.*, 1992).

*Forces due to adding a charge into the docked protein*

The force due to the introduction of a charge of magnitude +1*e* at the center of the hydrophobic sphere representing the docked protein was calculated along the lowest 50 normal modes. Figure 13 shows this force for both the major- and minor-groove binding proteins. The force for the major groove protein had large components along modes that *close* the major groove, and the reverse was true for the minor groove, which again produced a stronger effect. The vector sums of the projections over the lowest 50 modes (without the translations and rotations) for these two forces are shown in Figure 14. The effect of charging the protein is clear: because the positive charge was quite buried in the center of the protein, its introduction created strong attractive coulombic interactions with the phosphates, which the solvent did not screen completely. This caused a force pulling the DNA backbone buried inside the protein toward the added charge and buckling the nearby base pairs. The vector sum of the responses along the lowest modes bent the DNA towards the protein for both cases (not shown).

The force due to adding a charge of +1 in the center of either protein was smaller than the force due to docking that neutral protein, but the two were

generally opposite in sign along each mode. Docking the hydrophobic protein to the minor groove generated forces that tended to bend the DNA away from the protein, while introducing positive charge generated forces in essentially the opposite direction. Furthermore, the force for the introduction of charges of larger magnitudes was also calculated, and there was a roughly linear relationship between the size of the added charge and the magnitude of the projections of the force; that is, a charge of +4 caused four times the force of a +1 charge (results not shown). Thus, introducing a charge of about +1.5 to +2.5 into the center of the protein *nearly exactly cancels* the force of docking the neutral protein, so the overall force caused by docking the model protein here with a charge of about +2 would be small. This was true for both major- and minor-groove dockings. Docking a protein with a smaller charge at its center would bend the DNA away from it, while docking one with a larger charge would bend the DNA towards it. While the exact placement of the charges within the protein presumably affects the precise nature of the overall force of docking, it is interesting that the charge on the protein necessary to overcome the effects of displacing solvent was so small.

*Forces due to neutralizing phosphates in the DNA*

The force caused by directly neutralizing various sets of phosphates in the free DNA was also calculated using the lowest 50 modes as a basis. For this, the radii were unchanged, but the total charge on each selected phosphate group was reduced from –1.2 (CHARMM charges: P=1.5; O1P, O2P= –0.8; O3′, O5′= –0.55) to –0.2 (P=0.0; O1P, O2P=0.0; O3′, O5′= –0.1) thus making the nucleotide uncharged overall (MacKerell *et al.*, 1995). Four sets of phosphates were neutralized, as shown in Figure 15: four across the major groove, four and six across the minor groove, and two adjacent on the same strand. The two components of the force

— coulombic and solvation (Eq. 4) — were calculated and are shown for each neutralization in Figure 16. The coulombic or solvation force projections individually were of comparable magnitude to the force of docking the hydrophobic protein, but their sum and thus the total force was very small for all of the neutralizations.

The coulombic forces for the major- and minor-groove four-phosphate neutralizations shown in parts *a* and *b* had projections along many of the same modes as the docking of the major and minor groove binding proteins (Fig. 9), but these were generally in the opposite direction. Thus, this component of the force for neutralizing the four minor-groove-spanning phosphates had projections along directions that close the minor groove; Figure 17 (parts *a* and *b*) shows the vector sum of this component over the lowest 50 modes. Loss of the phosphate-phosphate repulsions caused a force that collapsed the groove by motion of the backbone parallel to the helical axis, buckling the central base pairs. Even the termini experienced some noticeable force, revealing the long-range nature of electrostatic interactions. The vector sums of the forces for the other cross-groove neutralizations also tended to close the appropriate grooves (not shown). Interestingly, the coulombic force of the same-strand neutralization acted to close a groove as well — the minor — as shown in Figure 17 (parts *c* and *d*). The neutralized phosphates were affected by the loss of coulombic interactions with their minor-groove partners more than the loss of the major-groove interactions. The un-neutralized strand experienced the larger force in this case, towards the neutralized phosphates as well as the helical axis.

Excluding the 1–2 and 1–3 bonded interactions from the calculation of the coulomb force, as is done in many molecular mechanics force fields, was found to have essentially no effect on the projections along the lowest normal modes, as shown in Figure 18. Both forces were calculated with no non-bonded cutoff

and a dielectric of 4. Only for some modes higher than about 90 does the exclusion of these local interactions affect the projection, and the coulombic component of the forces for other neutralizations revealed similar results.

However, the solvation effects nearly canceled the coulombic forces. The total forces, in fact, were so small that they were within the calculated error for most mode projections. To understand this, consider that while the unfavorable coulombic interaction of a phosphate with its minor-groove partner was large (about 9.5–10 kcal/mol; see Table 3), the solvation component of that total interaction energy, i.e. the *favorable* interaction of the phosphate with its partner's *reaction field* (Eq. 6), was also large (about –9.25 to –9.75 kcal/mol), leaving a total interaction of only about 0.25 kcal/mol. Thus, the loss of these small total interactions did not create a large force.

The vector sum over the lowest 50 modes of the total force for each neutralization did not act to open or close the grooves, but instead acted locally to rearrange the phosphates relative to nearby riboses and bases (not shown), due to the loss of these closer and stronger interactions (see Table 2). The response to these forces was also very small and difficult to analyze, but was mostly local in nature and did not involve bending the DNA toward or away from the neutralization. Thus, relative to the forces of solvent-displacement or addition-of-charge during protein binding, the force of direct phosphate neutralization was small and did not open or close the DNA grooves or cause the molecule to bend in any consistent direction.

*Solvent displacement in the association of protein-DNA complexes*

The phosphate-phosphate interaction contribution to the electrostatic free energy change due to binding two actual proteins to both straight B-DNA and the bent DNA conformation in each complex was calculated. The sums of the

solvent-screened inter-phosphate electrostatic interactions for each corner of the thermodynamic bending/binding cycle of Fig. 1 are listed in Table 4 for both the ETS1 and SRY complexes. All the totals were positive, since each phosphate had a charge of –1.2 in this parameter set, and they repelled one another.

Docking the proteins to DNA displaced solvent and consequently increased the strength of the phosphate-phosphate interactions due to the loss of screening. However, the increase was much greater for the B-DNA docking compared to the bent-DNA. The DNA in the complexes adopts a structure that minimizes the burial of the phosphates — by opening the minor groove and bending away from the protein — which in turn reduces the increase in their mutual repulsion upon docking. The last column in Table 4 lists this difference; that is, it shows the electrostatic "assistance" the proteins' presence gave the DNA in adopting the conformation seen in each complex. Of course, only the effect on the phosphate-phosphate interactions was computed. Still, the amount of energy was large (–17.0 kcal/mol for SRY, –14.4 for ETS1), and the distortion of the DNA made the dockings of both complexes more favorable.

It was also interesting that the phosphate-phosphate contribution to the bending of the free DNA was slightly favorable for SRY (–0.9 kcal/mol) and quite favorable for ETS1 (–7.9 kcal/mol). For SRY, the average nearest-neighbor-phosphate separation in complex was only slightly smaller than B-DNA (P–P distances and standard deviations: 6.3±0.4 Å vs. 6.50 Å), and for ETS1 this average was larger (6.8±0.3 Å). Both complexes have drastic openings of the minor groove (P–P distances: 17.5±1.5 Å for SRY, 18.9±0.6 for TBP, 11.46 for B-DNA). Thus, for SRY the increase in nearest-neighbor interaction energy was offset by the nearly complete loss of cross-minor groove repulsions, and for ETS1 both interactions are reduced, making the overall distortion favorable.

**Discussion**

We have presented a method to calculate the electrostatic force on a DNA structure caused by protein binding or phosphate neutralization. The internal flexibility of the DNA was included by way of its normal modes, along the lowest of which the components of the various forces were computed. This revealed the direction and relative magnitude of the forces. While the resultant forces for protein docking were local and involved opening the DNA groove contacting protein, the response to this force due to the vibrational properties of DNA was predominantly a bend. Moreover, computations on actual protein-DNA complexes showed that the coupled bending-binding events can result in savings of 15-20 kcal/mol in contributions to the bending energy through reduced inter-phosphate repulsions. Of course, there are significant non-electrostatic forces that can distort DNA on protein binding which were not studied here.

The observation that docking a hydrophobic model protein into either the major or minor groove of the DNA can create an electrostatic force that bends the DNA away from the protein has been made by Travers (1995) and Werner *et al.* (1996), who suggested this force may be at work in several minor-groove binding complexes in which the DNA bends away from the protein. The raising of the "local dielectric constant" in the groove caused by the presentation of primarily hydrophobic residues by these proteins to the DNA was posited to be the source of minor-groove widening and bending. A study by Elcock and McCammon (1996), which combined continuum solvation with molecular mechanics for internal deformations, also found DNA distortions similar to those described here. The approach of a large model protein was simulated and found to cause a dramatic opening of the minor groove, allowing better solvation of the phosphates — an "induced fit" of the DNA to the protein binding surface, completely of electrostatic origin. The authors reported no

bending of the DNA, perhaps because their model protein covered half a helical turn of the minor groove, as compared to the more lateral docking of the smaller protein model studied here.

Evidence of the strength of the electrostatic force driving this fitting came from calculating the sum of the phosphate-phosphate interaction energies for two protein-DNA complexes. The displacement of solvent by binding the proteins increases these repulsions, but much more so for straight B-DNA than for the conformation of the bent DNA in each complex. The difference was significant, about 15 kcal/mol. Since the solvent-displacement is required for binding, it is interesting that these complexes use it to their advantage in distorting the DNA. The energy for this bending comes from the loss of solvation energy of the phosphates, and must be overcome by non-electrostatic contributions to binding or by favorable electrostatic interactions between the protein and DNA, neither of which was calculated for the complexes here.

We also found that adding positive charges to the model protein in either groove created a force that acted to close the groove and bend the DNA toward the protein, thus opposing the force due to solvent displacement. Therefore, binding a model protein with charge in its center could cause little or no bending if the two electrostatic forces balance out. A charge of only $+2e$ was enough to mostly cancel the solvent-displacement force in both the major and minor groove. Further experiments to examine this relation between the total charge on the binding protein and the extent of DNA bending might prove interesting, and these could perhaps be performed utilizing protein charge ladders (Gao *et al.*, 1996).

However, we also found that direct neutralization of the phosphates (i.e., substitution with neutral isosters) did *not* cause a large force that would bend or distort the DNA, a confusing result in light of the work on asymmetric

neutralizations by Strauss and Maher (1994). These authors substituted neutral methylphosphonate analogs (Miller and Ts'o, 1987) for the six minor-groove-spanning phosphates in Fig. 15c and found via electrophoretic gel-shift phasing experiments that the DNA bent towards the neutralization. The authors described this for several DNA sequences, none of which was the one used here. We find, however, that the solvent (in the continuum model) is very effective in screening the cross-groove phosphate-phosphate electrostatic interactions — total interactions across the minor groove of only 0.25 kcal/mol per pair — and this would seem to limit the possibility of DNA bending caused by the loss of such interactions. It is possible that continuum electrostatics overestimates the extent of solvent screening in DNA, either because of lack of atomic detail of the water near the molecule, the use of the CHARMM charges and radii here, or perhaps because dielectric saturation is not accounted for (Warwicker, 1994). It is unclear from the present study what value of an effective dielectric between the minor-groove-spanning phosphates would be required for bending upon neutralization; but for the force to be comparable in magnitude to the force of the hydrophobic model protein docking, a very low value (4-20) would be needed (see Figs. 9 and 16). It is also interesting to note that contribution of the phosphate-phosphate interactions to the electrostatic energy change of the drastic bending seen in the ETS1 and SRY complexes was *favorable*, stressing the weakness of long-range interactions in DNA immersed in continuum solvent.

On the other hand, it is possible the bending seen experimentally has non-electrostatic origin or may not be as great as the phasing experiments predict. The authors do note a dependence of the magnitude of the bend on the type of counterion in the buffer solution, strong evidence electrostatics is important for bending. But both crystal structures of very short pieces of DNA (Szabó *et al.*, 1993; Han *et al.*, 1990) and molecular dynamics simulations (Hausheer *et al.*,

1990) suggest a difference in DNA with one stereoisomer (R or S) of methylphosphonate compared to the other, which could have electrostatic or non-electrostatic origin. Strauss and Maher do not specify which isomer or if a mixture was used in their study. Moreover, asymmetric neutralization could, as the authors point out, affect the anisotropic flexibility of the DNA, complicating the interpretation of gel-shift phasing studies. The extent of solvent screening as the DNA moves through the polyacrylamide gel is also unknown, but could be less than in free solution, perhaps affecting the bending.

Additional experiments might prove interesting. First, a selective synthesis of different diastereomers of methylphosphonated DNA and measurement of their bends compared to the unsubstituted DNA would help reveal the effects of the neutralization as opposed to stereospecific distortions. Second, the extent of bending caused by neutralization of phosphates on *one* strand should also be significant, since this removes interactions across both the minor and major grooves; neutralizing the second strand (across the minor groove) only removes the interactions with the next major groove.
Third, a study of the *twisting* of substituted DNA might probe our finding that the interactions of nearest neighbors on the same strand are much larger than cross-groove interactions; a possible scheme would substitute every other phosphate on each strand, which removes all nearest-neighbor interactions. Finally, an X-ray crystal or NMR study of a large piece of substituted DNA would test the conclusions made from the gel-shift experiments, and also provide another (supposed) intrinsic bend for the ongoing comparison of these structural methods with gel-shift studies in estimating DNA bending (e.g., Dickerson *et al.*, 1996).

The use of normal modes with $\varepsilon=4r$ is not entirely satisfying since continuum methods were used for calculating the electrostatic forces. The

prohibitive computational cost of minimizing and calculating the second derivatives on the combined CHARMM-DelPhi energy surface necessitated this, but these limitations are quickly being overcome, as the study by Elcock and McCammon (1996) demonstrates. Nevertheless, for the minimum structure, use of the normal modes effectively filtered out distracting local motion and gave the global character of the forces examined; this property of normal modes has been used to analyze trajectories for proteins in molecular dynamics simulations (Brooks *et al.*, 1995; Janezic *et al.*, 1995a, b). A consequence of this is that finite-difference forces may be calculated only along relatively few modes, as opposed to in every direction for each atom. For comparison, we examined the vector sum of the projections of the coulombic component of the forces of the direct-phosphate neutralizations (which were calculated in Cartesian coordinates and projected onto the normal modes; see Methods) over the lowest 25, 50, 75, 100, 150, 200, 500 and all 2640 of the normal modes; it was found that less than about 50 modes were enough to characterize each force's general action, and inclusion of modes greater than about 200 began to confuse things, with positively charged atoms moving in the opposite direction as their negatively charged neighbors. Furthermore, the projections of the response to these forces dropped off quickly with increasing mode number. The projections along modes 50–100 of the total force of the minor-groove six phosphate neutralization as well as the minor groove hydrophobic-protein docking were also calculated by finite differences, and similar conclusions reached for these forces.

We calculated the modes of the 14AT sequence of DNA as well, and found qualitative similarities with the 14GC modes. Though the exact frequencies were different in 14AT, modes were found that closely matched the motion of modes 7, 8, 9, 10, 12, 13 and 16 for 14GC. Smooth-bending mode 7 was still the lowest-frequency internal motion. This suggests a certain DNA-sequence independence

to many of the lowest modes. However, these modes are also the ones for which the harmonic approximation of the potential energy is weakest. For that reason, care should be used in interpreting the initial *responses* to the forces presented here (which used these frequencies in calculating their magnitudes), for while their overall character was frequently clear, their exact details might change with DNA sequence or with different energy parameters.

## Conclusion

The electrostatic energies and forces involved in DNA bending and binding to proteins were calculated and analyzed in the continuum model. Docking a hydrophobic protein in either the major or minor groove of DNA was found to create a force which opens the groove and starts to bend the DNA away from the protein. Examination of the docking of the ETS1 and SRY proteins to their operators revealed that the bending of DNA away from these proteins made the phosphate-phosphate contribution to the electrostatic energy of binding significantly more favorable, confirming the importance of the above force in actual protein-DNA complexes. However, this effect was overcome if a positively charged model protein was docked instead, and only a relatively small amount of charge was enough to bend the DNA back toward the protein if it was deeply buried. Finally, due to the very effective solvent screening predicted by the continuum model, interactions between phosphate groups across the grooves of DNA were found to be small in magnitude; hence, asymmetric neutralization of these phosphates was found to cause only a very small and local force on the DNA structure.

**Acknowledgments**

Neither science nor education are solitary endeavors, and I would like to thank a few of the many people made this work possible. Discussions with Zak Hendsch about just about everything were always enlightening, and many of his ideas are embedded throughout this thesis. Susan Brighton was uniquely encouraging to me, as were my parents and my friends, Kelly Conway, Stacey Eckman, Rick Farrer, Dave Kayes, Bill Kobertz, Rose Koch, Paige Mahaney, Jim Masserano, Ted Trautman, Greg Wakeham, Dave Wang, Marc Wefers and Doug Whittington. I also thank Bruce for constantly challenging me and always demanding excellence. Finally, I wish to thank M. Karplus for making the CHARMM program available, B. Honig for the DelPhi program, D. L. Beveridge for the Dials and Windows program, and M. Carson for the Ribbons program. This work was supported by grants from the NIH (GM-47678), NSF pre-doctoral fellowship program, and an MIT Science Partnership Award.

**References**

F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, (1977). The protein data bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

O'M Bockris and A. K. N. Reddy, (1973). *Modern Electrochemistry*. Plenum, New York.

B. Brooks and M. Karplus, (1983). Harmonic dynamics of proteins: normal mode and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA* **80**, 6571–6575.

B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.

B. R. Brooks, D. Janezic and M. Karplus, (1995). Harmonic analysis of large systems. I. Methodology. *J. Comput. Chem.* **16**, 1522–1542.

C. R. Calladine and H. R. Drew, (1992). *Understanding DNA*, Academic Press, London.

M. Carson, (1987). Ribbon models of macromolecules. *J. Mol. Graphics* **5**, 103–106.

M. Carson and C. E. Bugg, (1986). Algorithm for ribbons models of proteins. *J. Mol. Graphics* **4**, 121–122.

M. L. Connolly, (1983a). Analytical molecular surface calculation. *J. Appl. Cryst.* **16**, 548–558.

M. L. Connolly, (1983b). Solvent-accessible surfaces of proteins and nucleic acids. *Science* **221**, 709–713.

M. E. Davis and J. A. McCammon, (1989). Calculating electrostatic forces from grid-calculated potentials. *J. Comput. Chem.* **11**, 401–409.

M. E. Davis and J. A. McCammon, (1990). Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* **90**, 509–521.

M. E. Davis and J. A. McCammon, (1991). Dielectric boundary smoothing in finite-difference solutions to the Poisson equation – an approach to improve accuracy and convergence. *J. Comput. Chem.* **12**, 909-912.

R. E. Dickerson, M. L. Kopka and P. Pjura, (1983). A stochastic model for helix bending in B-DNA. *J. Biomol. Struct. Dyn.* **1**, 755–771.

R. E. Dickerson, D. Goodsell and M. L. Kopka, (1996). MPD and DNA bending in crystals and in solution. *J. Mol. Biol.* **256**, 108–125.

A. H. Elcock and J. A. McCammon, (1996). The low dielectric interior of proteins is sufficient to cause major structural changes in DNA on association. *J. Am. Chem. Soc.* **118**, 3787–3788.

M. A. El Hassan and C. R. Calladine, (1996). Propeller-twisting of base-pairs and the conformational mobility of dinucleotide steps in DNA. *J. Mol. Biol.* **259**, 95–103.

R. A. Friedman and B. Honig, (1992). The electrostatic contribution to DNA base-stacking interactions. *Biopolymers* **32**, 145–159.

J. Gao, M. Mammen and G. M. Whitesides, (1996). Evaluating electrostatic contributions to binding with the use of protein charge ladders. *Science* **272**, 535–537.

M. K. Gilson and B. H. Honig, (1987). Calculation of electrostatic potentials in an enzyme active site. *Nature* **330**, 84–86.

M. K. Gilson and B. Honig, (1988). Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins: Struc. Funct. Genet.* **4**, 7–18.

M. K. Gilson, K. A. Sharp and B. H. Honig, (1988). Calculating the electrostatic potential of molecules in solution: methods and error assessment. *J. Comput. Chem.* **9**, 327–335.

M. K. Gilson, M. E. Davis, B. A. Luty and J. A. McCammon, (1993). Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. *J. Phys. Chem.* **97**, 3591–3600.

A. A. Gorin, V. B. Zhurkin and W. K. Olson, (1995). B-DNA twisting correlates with base-pair morphology. *J. Mol. Biol.* **247**, 34–48.

F. Han, W. Watt, D. J. Duchamp, L. Callahan, F. J. Kézdy and K. Agarwal, (1990). Molecular structure of deoxycytidyl-3'-methylphosphonate ($R_p$) 5'-

deoxyguanidine, d(Cp($CH_3$)G). A neutral dinucleotide with Watson-Crick base pairing and a right handed helical twist. *Nuc. Acids Res.* **18**, 2759–2767.

R. E. Harrington, (1992). DNA curving and bending in protein-DNA recognition. *Mol. Microbiol.* **6**, 2549–2555.

F. H. Hausheer, U. C. Singh, T. C. Palmer and J. D. Saxe, (1990). Dynamic properties and electrostatic potential surface of neutral DNA heteropolymers. *J. Am. Chem. Soc.* **112**, 9468–9474.

B. Honig and A. Nicholls, (1995). Classical electrostatics in biology and chemistry. *Science* **268**, 1144–1149.

D. Janezic and B. R. Brooks, (1995a). Harmonic analysis of large systems. II. Comparison of different protein models. *J. Comput. Chem.* **16**, 1543–1553.

D. Janezic, R. M. Venable and B. R. Brooks, (1995b). Harmonic analysis of large systems. III. Comparison with molecular dynamics. *J. Comput. Chem.* **16**, 1554–1566.

B. Jayaram, K. A. Sharp and B. Honig, (1989). The electrostatic potential of B-DNA. *Biopolymers* **28**, 975–993.

J. G. Kirkwood, (1934). Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Phys.* **2**, 351–361.

I. Klapper, R. Hagstrom, R. Fine, K. Sharp and B. Honig, (1986). Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: effects of ionic strength and amino-acid modification. *Proteins: Struct. Funct. Genet.* **1**, 47–59.

R. Lavery and H. Sklenar, (1988). The definition of generalized helicoid parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **6**, 63–91.

B. A. Luty, M. E. Davis and J. A. McCammon, (1992). Solving the finite-difference non-linear Poisson–Boltzmann equation. *J. Comput. Chem.* **13**, 1114–1118.

A. D. MacKerell, Jr., J. Wiórkiewicz-Kuczera and M. Karplus, (1995). An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* **117**, 11946–11975.

G. Manning, (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11**, 179–246.

G. Manning, K. Ebralidse, A. D. Mirzabekov and A. Rich, (1989). An estimate of the extent of folding of nucleosomal DNA by laterally asymmetric neutralization of phosphate groups. *J. Biomol. Struct. Dyn.* **6**, 877–889.

P. S. Miller and P. O. P. Ts'o, (1987). A new approach to chemotherapy based on molecular biology and nucleic acid chemistry – Matagen (masking tape for gene expression). *Anti-cancer Drug Design* **2**, 117–128.

A. D. Mirzabekov and A. Rich, (1979). Asymmetric lateral distribution of unshielded phosphate groups in nucleosomal DNA and its role in bending. *Proc. Natl. Acad. Sci. USA* **76**, 1118–1121.

V. Mohan, M. E. Davis, J. A. McCammon and B. M. Pettitt, (1992). Continuum model calculations of solvation free energies: accurate evaluation of electrostatic contributions. *J. Phys. Chem.* **96**, 6428–6431.

G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery and H. Sklenar, (1989). Conformational and helicoid analysis of 30ps of molecular dynamics on the d(CGCGAATTCGCG) double helix: "curves," dials and windows. *J. Biomol. Struct. Dyn.* **6**, 669–699.

F. M. Richards, (1977). Areas, volumes, packing and protein structure. *Ann. Rev. Biophys. Bioeng.* **6**, 151–176.

K. Sharp, (1991). Incorporating solvent and ion screening into molecular dynamics using the finite-difference Poisson-Boltzmann method. *J. Comput. Chem.* **12**, 454–468.

K. A. Sharp and B. Honig, (1990a). Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 301–332.

K. A. Sharp and B. Honig, (1990b). Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* **94**, 7684–7692.

K. A. Sharp, B. Honig and S. C. Harvey, (1990). Electric potential of transfer RNAs: codon–anticodon recognition. *Biochemistry* **29**, 340–346.

J. K. Strauss and L. J. Maher III, (1994). DNA bending by asymmetric neutralization of phosphates. *Science* **266**, 1829–1834.

M. Suzuki and N. Yagi, (1995). Stereochemical basis of DNA bending by transcription factors. *Nuc. Acids Res.* **23**, 2083–2091.

T. Szabó, D. Noréus, R. Norrestam and J. Stawinski, (1993). Molecular and crystal structure of *S*p-thymidin-3'-yl 4-thiothymidin-5'-yl methylphosphonate. *Nuc. Acids Res.* **21**, 3921–3926.

A. A. Travers, (1995). Reading the minor groove. *Nature Struct. Biol.* **2**, 615–618.

P. C. van der Vliet and C. P. Verrijzer, (1993). Bending of DNA by transcription factors. *BioEssays* **15**, 25–32.

J. Warwicker, (1994). Improved continuum electrostatic modeling in proteins, with comparison to experiment. *J. Mol. Biol.* **236**, 887–903.

J. Warwicker and H. C. Watson, (1982). Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.* **157**, 671–679.

M. H. Werner, G. M. Clore, C. L. Fisher, R. J. Fisher, L. Trinh, J. Shiloach and A. M. Gronenborn, (1995a). The solution structure of the human ETS1-DNA complex reveals a novel mode of binding and true side-chain intercalation. *Cell* **83**, 761–770.

M. H. Werner, J. R. Ruth, A. M. Gronenborn and G. M. Clore, (1995b). Molecular basis of human 46X,Y sex reversal revealed from the 3-dimensional solution structure of the human SRY-DNA complex. *Cell* **81**, 705–714.

M. H. Werner, A. M. Gronenborn and G. M. Clore, (1996). Intercalation, DNA kinking, and control of transcription. *Science* **271**, 778–784.

E. B. Wilson, Jr., J. C. Decius, and P. C. Cross, (1980). *Molecular Vibrations*. Dover, New York.

M. A. Young, G. Ravishanker, D. L. Beveridge and H. M. Berman, (1995). Analysis of local helix bending in crystal structures of DNA oligonucleotides and DNA-protein complexes. *Biophysical J.* **68**, 2454–2468.

M. Zacharias, B. A. Luty, M. E. Davis and J. A. McCammon, (1992). Poisson-Boltzmann analysis of the lambda repressor-operator interaction. *Biophysical J.* **63**, 1280–1285.

Table 1. 14GC Normal Modes

| mode # | frequency ($cm^{-1}$) | symmetry[1] | description[2,3] |
|---|---|---|---|
| 1-6 | < 0.01 | — | overall translation and rotation |
| 7 | 1.43 | sym | large bend in xz plane towards minor groove (–x axis), opening the major groove in the middle. |
| 8 | 1.52 | anti | two-hinge bend in yz plane, with the hinges about 1/4 and 3/4 of the way down in the helix, opening the minor groove at one terminus and the major groove at the other |
| 9 | 2.29 | sym | termini tip inward toward the z axis and twist, opening central major groove slightly |
| 10 | 3.13 | sym | untwisting that opens major groove at center and elongates helix |
| 11 | 3.23 | anti | termini rock, closing the minor grooves near them |
| 12 | 3.96 | sym | rocking of backbone, closing the minor and opening major groove near center, and closing the minor grooves near termini |
| 13 | 5.29 | anti | twist, which opens major groove near one terminus and closes it near the other |
| 14 | 5.77 | sym | skewing of whole helix in yz plane |
| 15 | 6.38 | anti | terminal riboses rock, opening minor groove of one terminus and closing it near the other |
| 16 | 6.54 | sym | contraction along helical axis that closes whole minor groove at once |
| 17 | 6.84 | anti | two phosphates (A7 and B21) rock |
| 18 | 6.86 | sym | closing of major and minor grooves near center and opening of minor grooves near termini |
| 19 | 7.10 | anti | riboses near termini rock and middle skews in yz plane |
| 20 | 7.24 | sym | riboses near termini rock |

[1] "Symmetric" modes preserve the pseudo-two-fold symmetry axis running through the central major and minor grooves, while "anti-symmetric" modes break this.
[2] For convenience, let the z-axis be the helical axis and the positive x-axis point towards the central major groove. Note that the direction in the xy plane of the major and minor grooves rotates as one moves along the helix.
[3] Since the motion represented by each mode is harmonic about the minimum structure, only the positive (forward) direction along each mode is described.

**Table 2. Dominant electrostatic interactions in 14GC**

The total energy and components (in kcal/mol; see Eq. 6) of the interactions between nucleotide 7A and the rest of the DNA are shown. All nucleotides were divided into three "groups": the bases (Gua: 15 atoms, total charge=0.0; Cyt: 12 atoms, total charge 0.0), the ribose rings (13 atoms, total charge = 0.2) and the phosphates (5 atoms, total charge –1.2). The interactions whose total energy was greater in magnitude than 0.15 kcal/mol are listed, with a positive number signifying a repulsion. The numbering for the nucleotides is shown in Fig. 2.

| interaction | coulombic energy | solvation energy | total energy (kcal/mol) |
|---|---|---|---|
| gua7A — cyt8B | –5.228 | 1.196 | –4.032 |
| gua7A — phos7A | –2.032 | 0.774 | –1.258 |
| ribo7A — phos8A | –6.256 | 5.115 | –1.141 |
| phos7A — ribo6A | –6.453 | 5.276 | –1.177 |
| phos7A — phos6A | 18.013 | –16.985 | 1.027 |
| gua7A — gua7B | 2.093 | –1.081 | 1.011 |
| phos7A — phos8A | 16.937 | –16.025 | 0.912 |
| gua7A — ribo7A | –0.682 | –0.066 | –0.749 |
| gua7A — ribo8A | 0.523 | –0.094 | 0.429 |
| phos7A — cyt6A | 1.462 | –1.051 | 0.411 |
| gua7A — gua9B | –1.033 | 0.674 | –0.359 |
| phos7A — ribo7A | –3.640 | 3.969 | 0.329 |
| phos7A — phos12B | 9.472 | –9.260 | 0.212 |
| gua7A — phos9B | –1.293 | 1.085 | –0.208 |
| phos7A — gua9B | –1.304 | 1.098 | –0.206 |
| phos7A — cyt8A | –1.898 | 1.718 | –0.180 |
| phos7A — phos9A | 9.532 | –9.360 | 0.172 |
| phos7A — phos5A | 9.504 | –9.335 | 0.168 |

**Table 3. Phosphate-phosphate interactions in 14GC.**

(a) The Coulomb's law ($\varepsilon$=4) interaction between pairs of phosphates in the DNA. Note each phosphate has a total charge of –1.2 in this parameter set.
(b) The total interaction energies at 0.145 M ionic strength for pairs of phosphates, calculated with continuum electrostatic methods according to Eq. 6.
(c) Total interaction energies at zero ionic strength, calculated with continuum electrostatic methods according to Eq. 6.
The molecule is essentially two-fold symmetric, so the interaction energy of 4A–5A virtually equals that of 4B–5B, and 4A–5B virtually equals 5A–4B. The numbering for the nucleotides is shown in Fig. 2.

a)

| | 4A | 5A | 6A | 7A | 8A | 9A | 10A | 11A | 12A |
|---|---|---|---|---|---|---|---|---|---|
| 4A | — | | | | | | | | |
| 5A | 18.19 | — | | | | | | | |
| 6A | 9.08 | 17.00 | — | | | | | | |
| 7A | 6.56 | 9.50 | 18.01 | — | | | | | |
| 8A | 5.20 | 6.54 | 9.02 | 16.94 | — | | | | |
| 9A | 4.62 | 5.46 | 6.55 | 9.53 | 17.88 | — | | | |
| 10A | 4.31 | 4.78 | 5.18 | 6.53 | 8.97 | 16.91 | — | | |
| 11A | 4.05 | 4.42 | 4.59 | 5.45 | 6.54 | 9.64 | 17.85 | — | |
| 12A | 3.99 | 4.29 | 4.26 | 4.76 | 5.16 | 6.57 | 8.96 | 16.89 | — |
| 4B | 4.26 | | | | | | | | |
| 5B | 5.24 | 6.51 | | | | | | | |
| 6B | 5.92 | 6.81 | 6.26 | | | | | | |
| 7B | 6.83 | 7.31 | 6.36 | 6.45 | | | | | |
| 8B | 6.24 | 6.32 | 5.74 | 6.09 | 6.38 | | | | |
| 9B | 6.38 | 6.46 | 6.13 | 6.70 | 7.49 | 8.55 | | | |
| 10B | 5.70 | 6.04 | 6.37 | 7.44 | 9.41 | 9.93 | 9.37 | | |
| 11B | 6.16 | 6.66 | 7.40 | 8.33 | 9.58 | 7.99 | 6.62 | 5.06 | |
| 12B | 6.41 | 7.37 | 9.13 | 9.47 | 8.94 | 6.52 | 5.17 | 4.20 | 3.64 |

**Table 3 continued**

b)

| | 4A | 5A | 6A | 7A | 8A | 9A | 10A | 11A | 12A |
|---|---|---|---|---|---|---|---|---|---|
| **4A** | — | | | | | | | | |
| **5A** | 1.03 | — | | | | | | | |
| **6A** | 0.14 | 0.94 | — | | | | | | |
| **7A** | 0.05 | 0.17 | 1.03 | — | | | | | |
| **8A** | 0.02 | 0.05 | 0.14 | 0.91 | — | | | | |
| **9A** | 0.01 | 0.02 | 0.05 | 0.17 | 1.02 | — | | | |
| **10A** | 0.01 | 0.01 | 0.02 | 0.05 | 0.14 | 0.91 | — | | |
| **11A** | 0.01 | 0.01 | 0.01 | 0.02 | 0.05 | 0.18 | 1.01 | — | |
| **12A** | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.05 | 0.14 | 0.89 | — |
| **4B** | 0.02 | | | | | | | | |
| **5B** | 0.04 | 0.09 | | | | | | | |
| **6B** | 0.06 | 0.09 | 0.06 | | | | | | |
| **7B** | 0.10 | 0.12 | 0.06 | 0.05 | | | | | |
| **8B** | 0.06 | 0.06 | 0.03 | 0.03 | 0.04 | | | | |
| **9B** | 0.06 | 0.05 | 0.03 | 0.04 | 0.08 | 0.14 | | | |
| **10B** | 0.03 | 0.03 | 0.04 | 0.08 | 0.20 | 0.24 | 0.21 | | |
| **11B** | 0.03 | 0.04 | 0.08 | 0.13 | 0.22 | 0.13 | 0.07 | 0.02 | |
| **12B** | 0.04 | 0.08 | 0.19 | 0.21 | 0.19 | 0.07 | 0.03 | 0.01 | 0.00 |

c)

| | 4A | 5A | 6A | 7A | 8A | 9A | 10A | 11A | 12A |
|---|---|---|---|---|---|---|---|---|---|
| **4A** | — | | | | | | | | |
| **5A** | 1.48 | — | | | | | | | |
| **6A** | 0.46 | 1.32 | — | | | | | | |
| **7A** | 0.30 | 0.49 | 1.45 | — | | | | | |
| **8A** | 0.21 | 0.28 | 0.46 | 1.33 | — | | | | |
| **9A** | 0.18 | 0.22 | 0.30 | 0.50 | 1.44 | — | | | |
| **10A** | 0.17 | 0.19 | 0.21 | 0.28 | 0.46 | 1.32 | — | | |
| **11A** | 0.17 | 0.18 | 0.18 | 0.22 | 0.30 | 0.50 | 1.42 | — | |
| **12A** | 0.19 | 0.19 | 0.17 | 0.18 | 0.21 | 0.28 | 0.45 | 1.31 | — |
| **4B** | 0.22 | | | | | | | | |
| **5B** | 0.29 | 0.40 | | | | | | | |
| **6B** | 0.34 | 0.41 | 0.33 | | | | | | |
| **7B** | 0.41 | 0.43 | 0.30 | 0.28 | | | | | |
| **8B** | 0.33 | 0.30 | 0.23 | 0.23 | 0.25 | | | | |
| **9B** | 0.31 | 0.28 | 0.23 | 0.25 | 0.32 | 0.41 | | | |
| **10B** | 0.23 | 0.23 | 0.25 | 0.32 | 0.52 | 0.58 | 0.57 | | |
| **11B** | 0.23 | 0.25 | 0.32 | 0.40 | 0.55 | 0.43 | 0.35 | 0.24 | |
| **12B** | 0.26 | 0.32 | 0.50 | 0.55 | 0.53 | 0.34 | 0.25 | 0.18 | 0.15 |

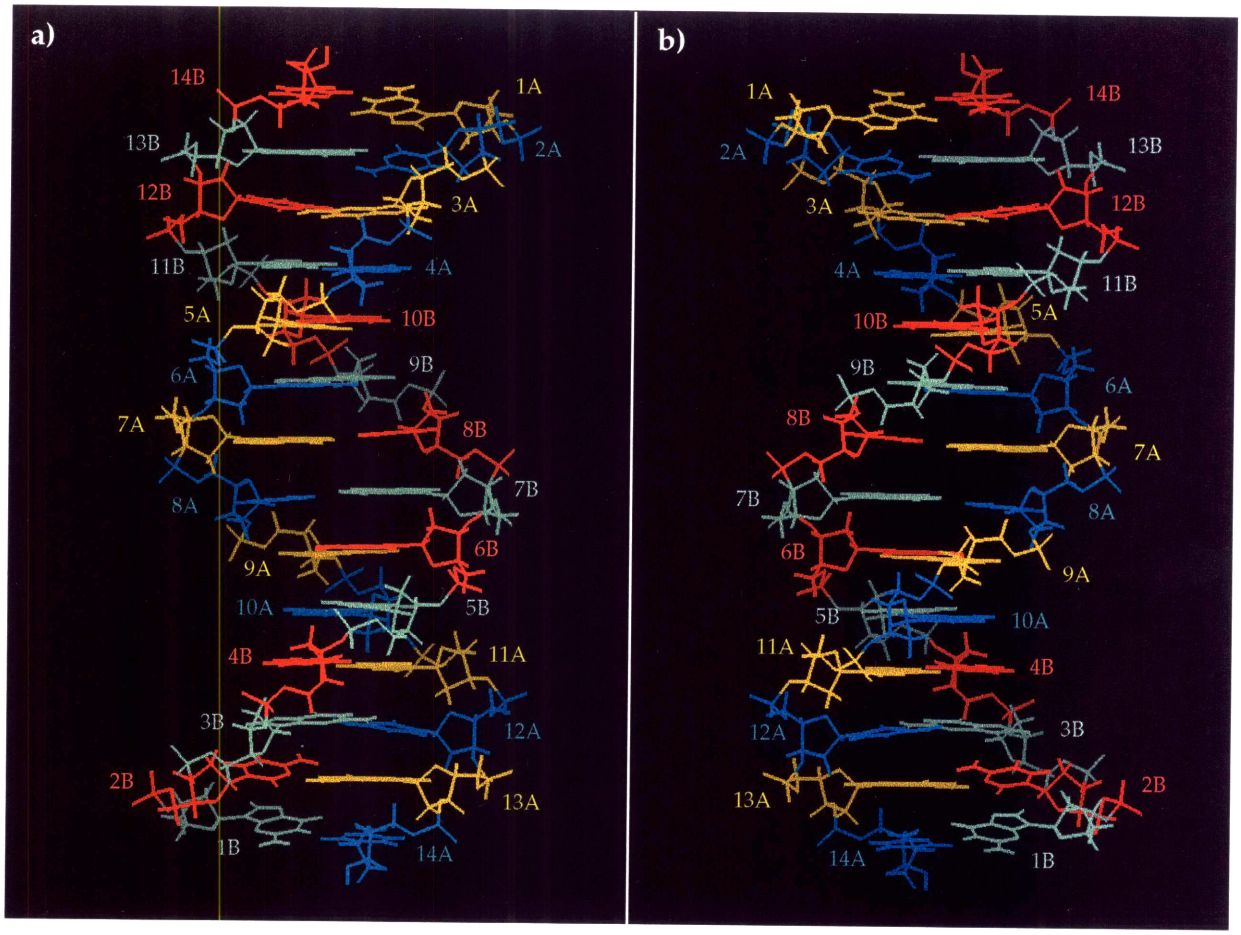


**Figure 2. The 14GC minimum structure,** viewed down (a) the central major groove and (b) the central minor groove (a 180 degree rotation about the helical axis). Colors show the numbering of the nucleotides.

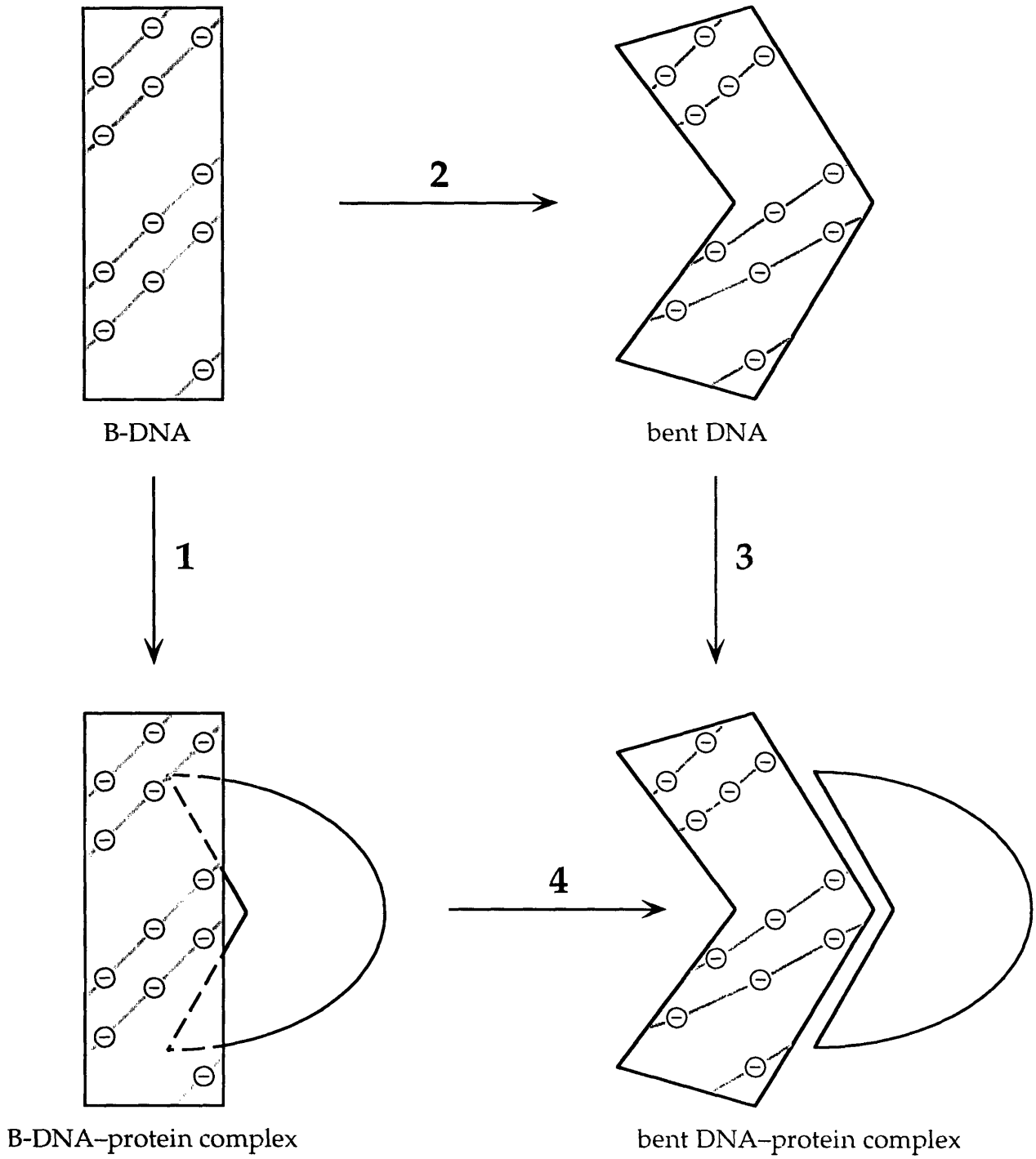


**Figure 1. DNA bending effects on protein binding.** Free energy changes (1) and (3) represent the binding the hydrophobic protein to B-DNA and bent-DNA, respectively. Changes (2) and (4) represent bending DNA free in solution and when bound to the protein.

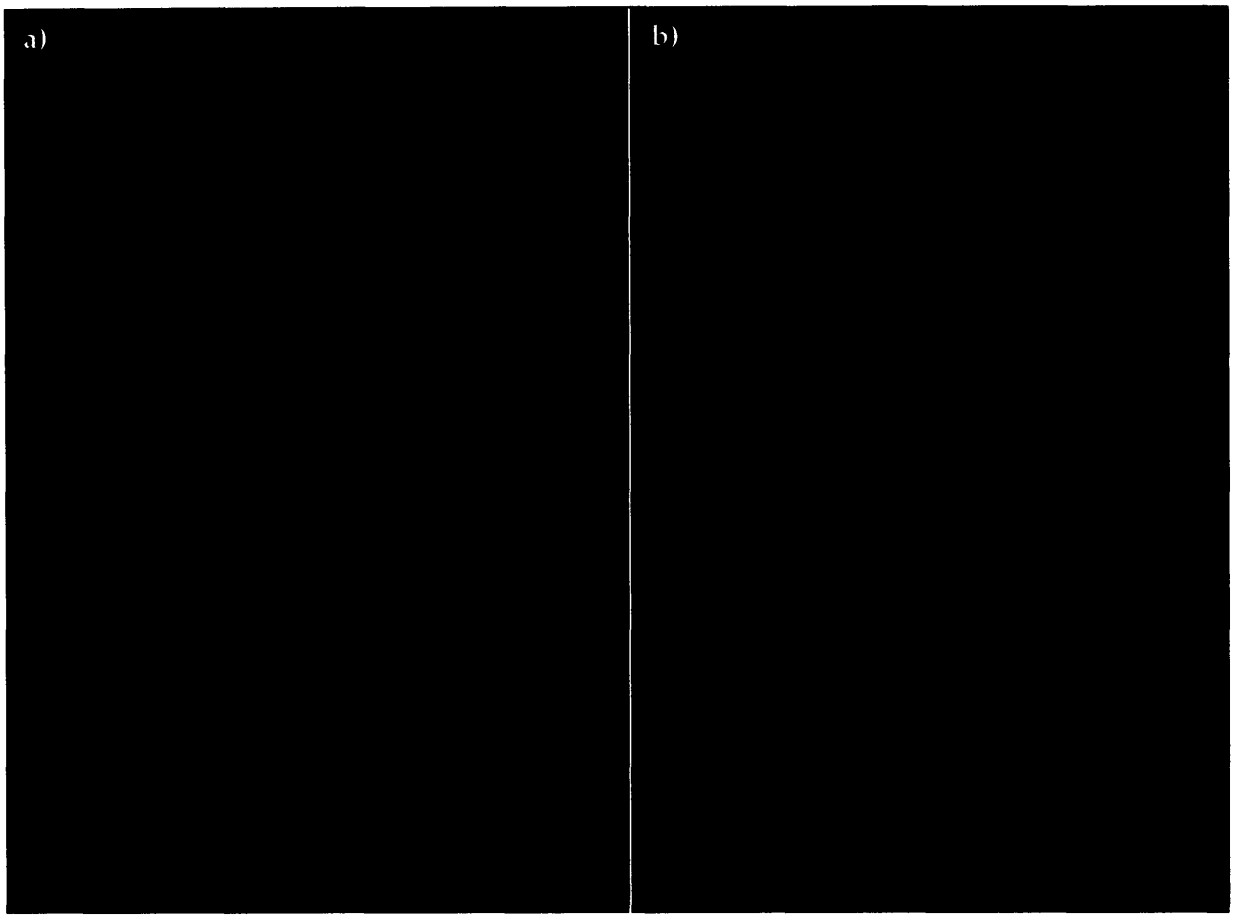


**Figure 2. The 14GC minimum structure,** viewed down (a) the central major groove and (b) the central minor groove (a 180 degree rotation about the helical axis). Colors show the numbering of the nucleotides.

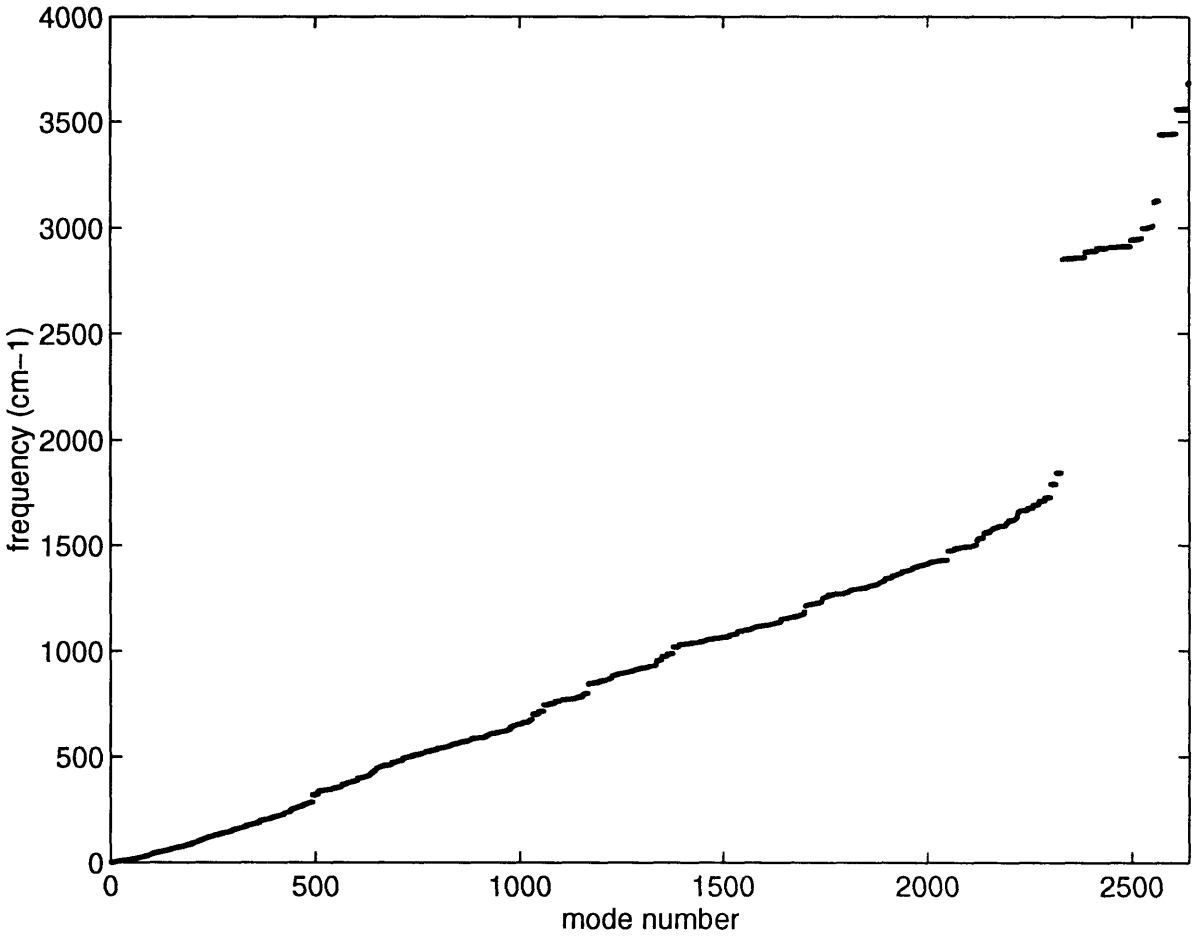


**Figure 3. Normal mode frequencies of 14GC.**

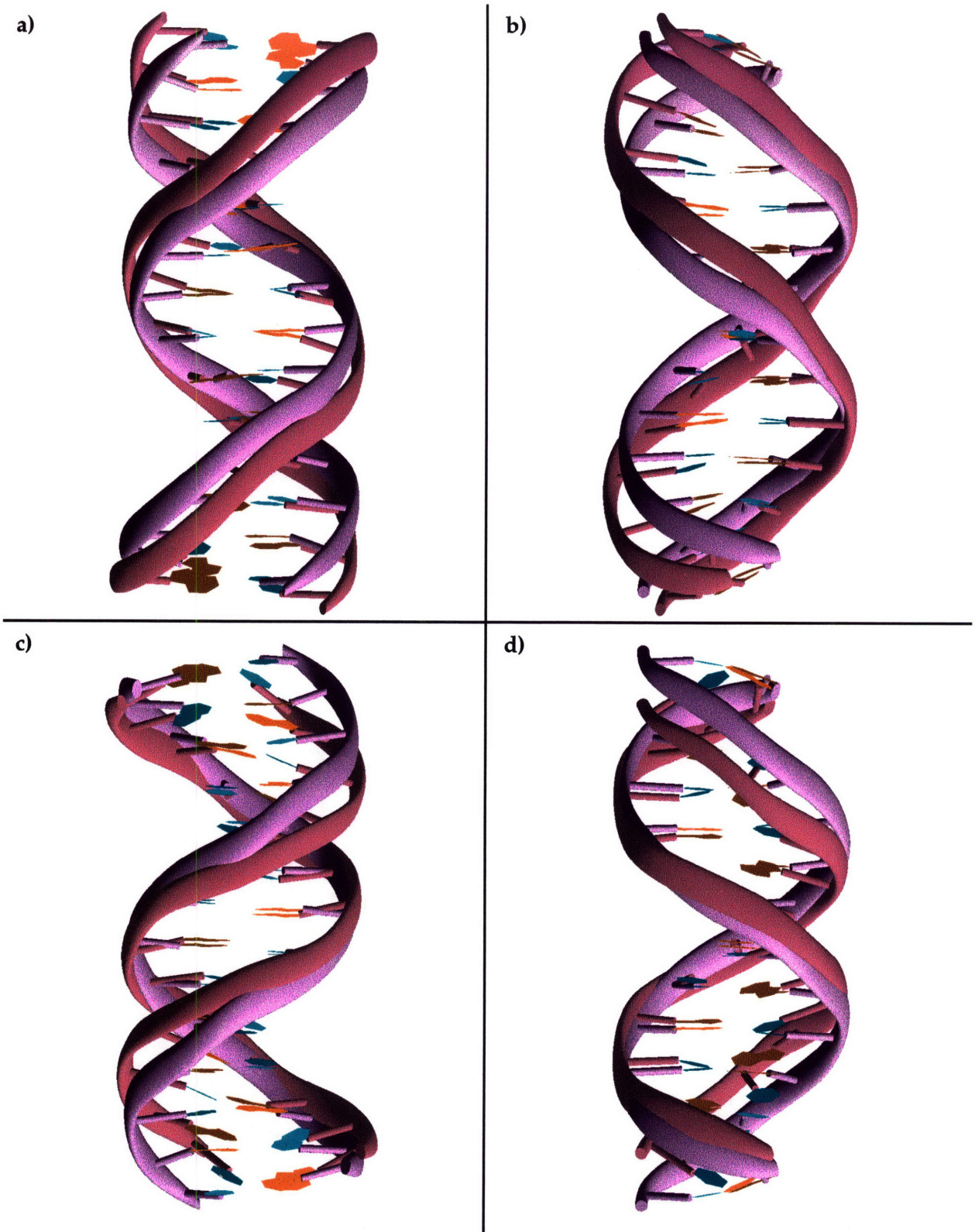


**Figure 4. Two normal modes of 14GC.** Mode 10, which untwists and elongates the helix is shown from two views in (a) and (b), while mode 16, which closes the whole minor groove and shrinks the helix is shown in (c) and (d). For each picture, the unmoved DNA structure is purple, while the structure after a large step (the atom moving farthest moves 5 Angstroms) in the positive direction along the normal mode is red.

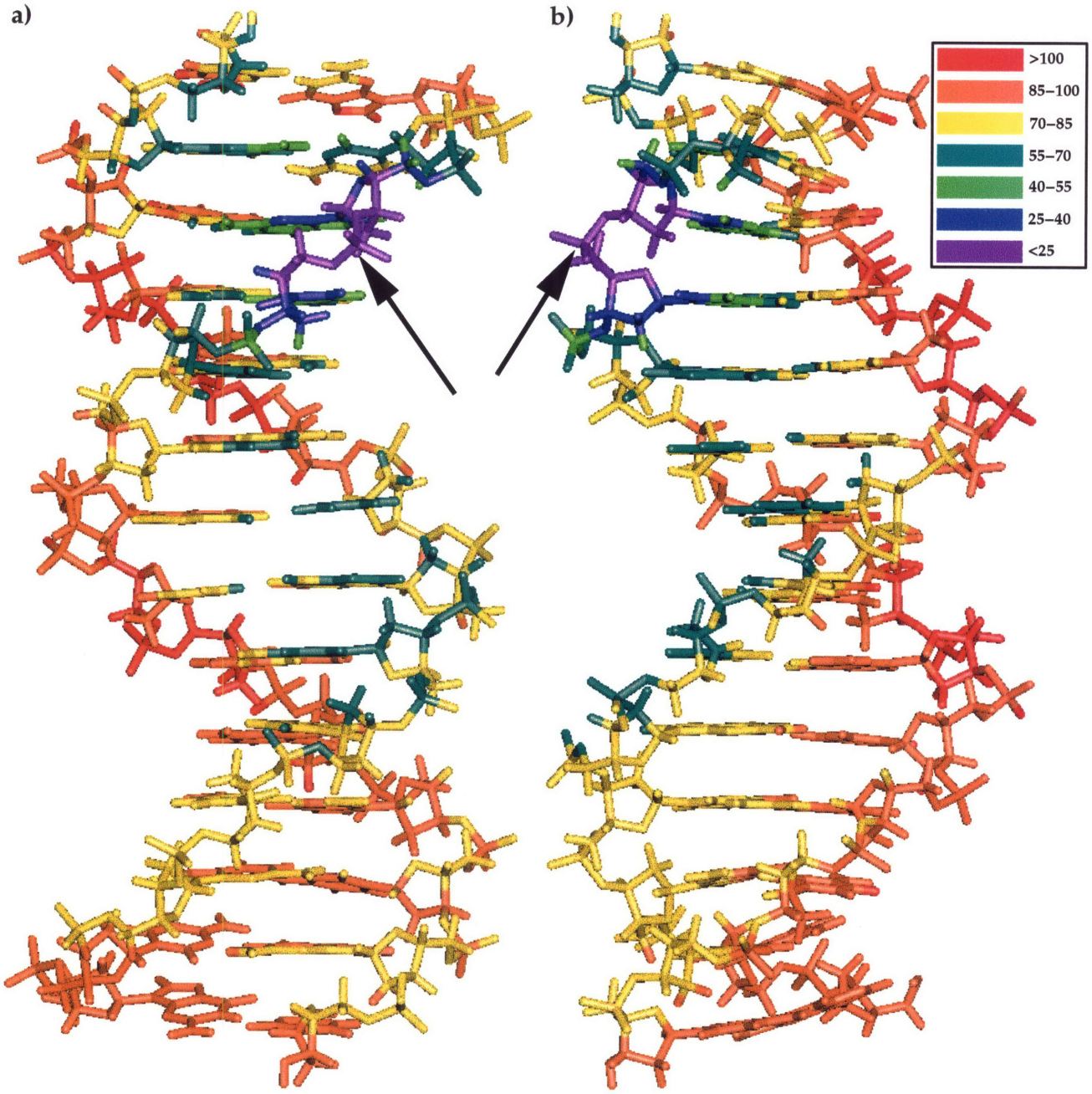


**Figure 5. Effective dielectric constants in 14GC.** Two views (a) and (b) show the effective dielectric between phosphate group 4A (highlighted with the arrow) and the rest of the atoms in the 14GC structure, color–coded according to the legend.

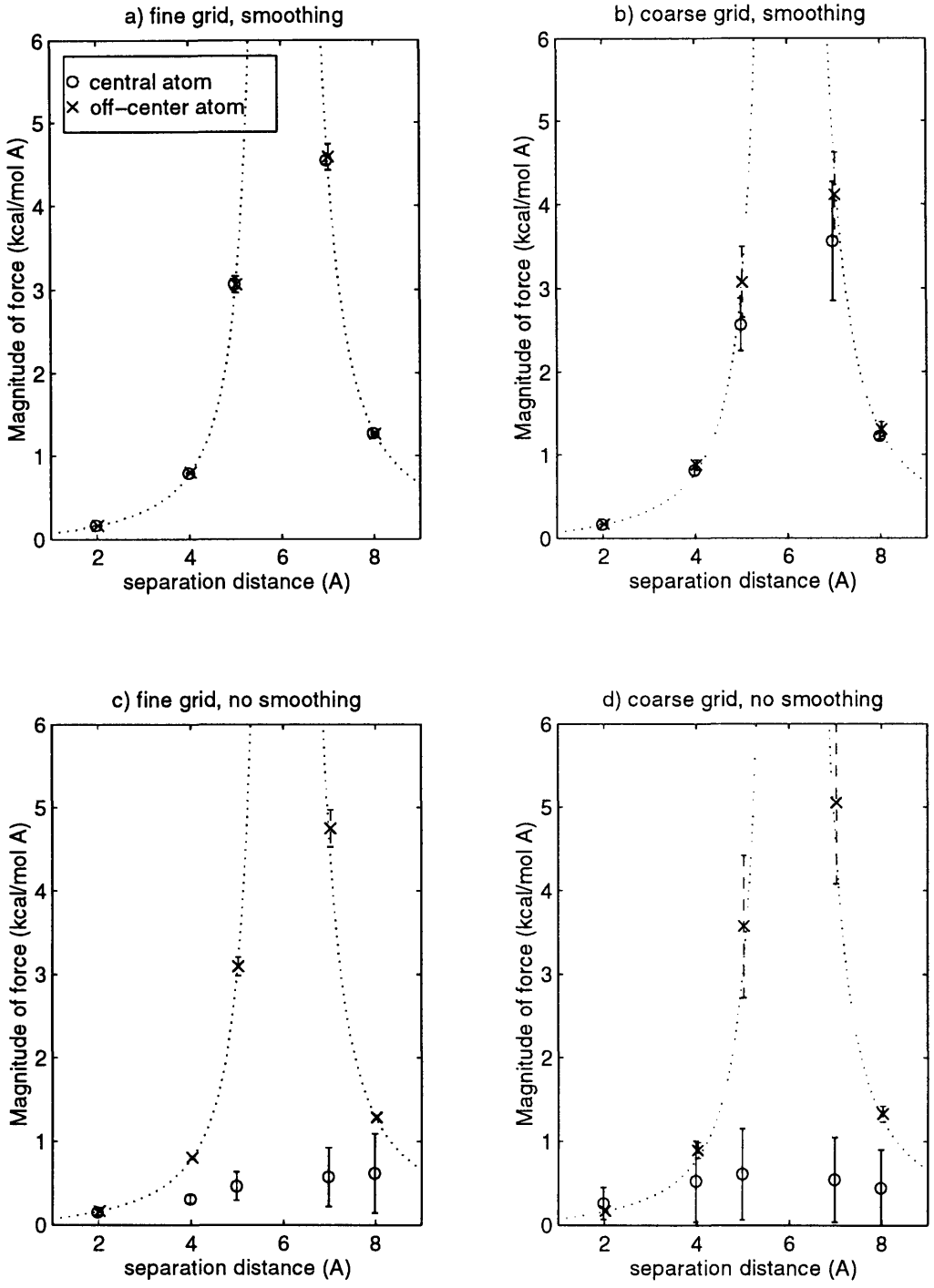


**Figure 6. Comparison of DelPhi finite–difference forces with analytical solution.** The dotted line shows the magnitude of the analytical force from Eq. 10. The circles (slightly to the left of the x–values) show the magnitude of the force on the central atom calculated by DelPhi with a finite difference step size of 0.001 Angstroms, while the "x"s (slightly right) are for the off–center atom. Error bars are twice the standard deviation of the mean of 11 rotations on the grid. The coarse grid had 1.2 grids/A (50A cube) and the fine grid had 4.9 grids/A (12A cube). See text for a complete description of the system and a discussion of smoothing.

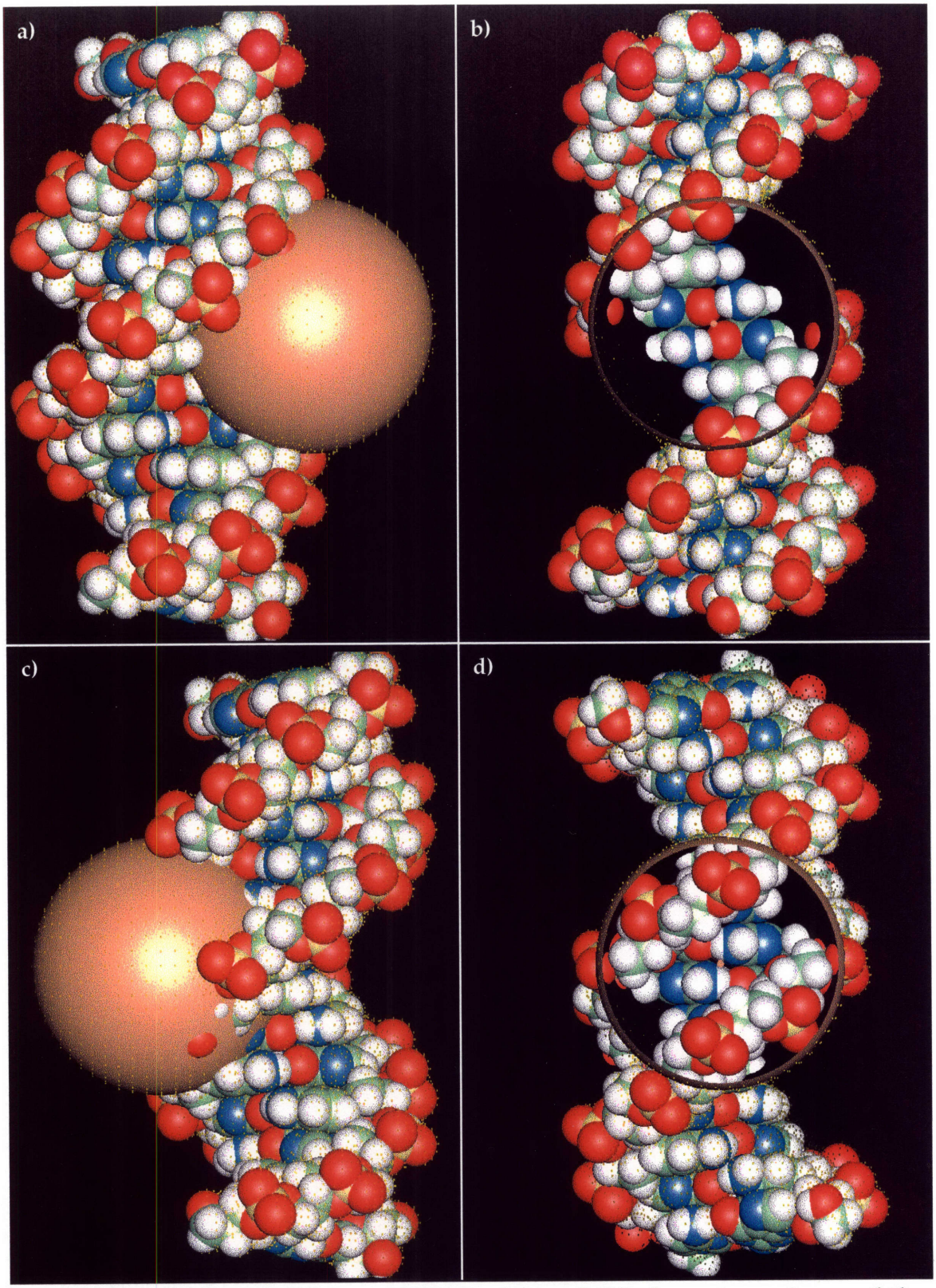


**Figure 7. Protein dockings to DNA.** (a) and (b) show two views of the model protein, a 10 A radius hydrophobic sphere, docked to the 14GC structure in the major groove while (c) and (d) show the minor groove protein. The spheres are cut away in (b) and (d) to reveal the atoms of 14GC *inside* them.

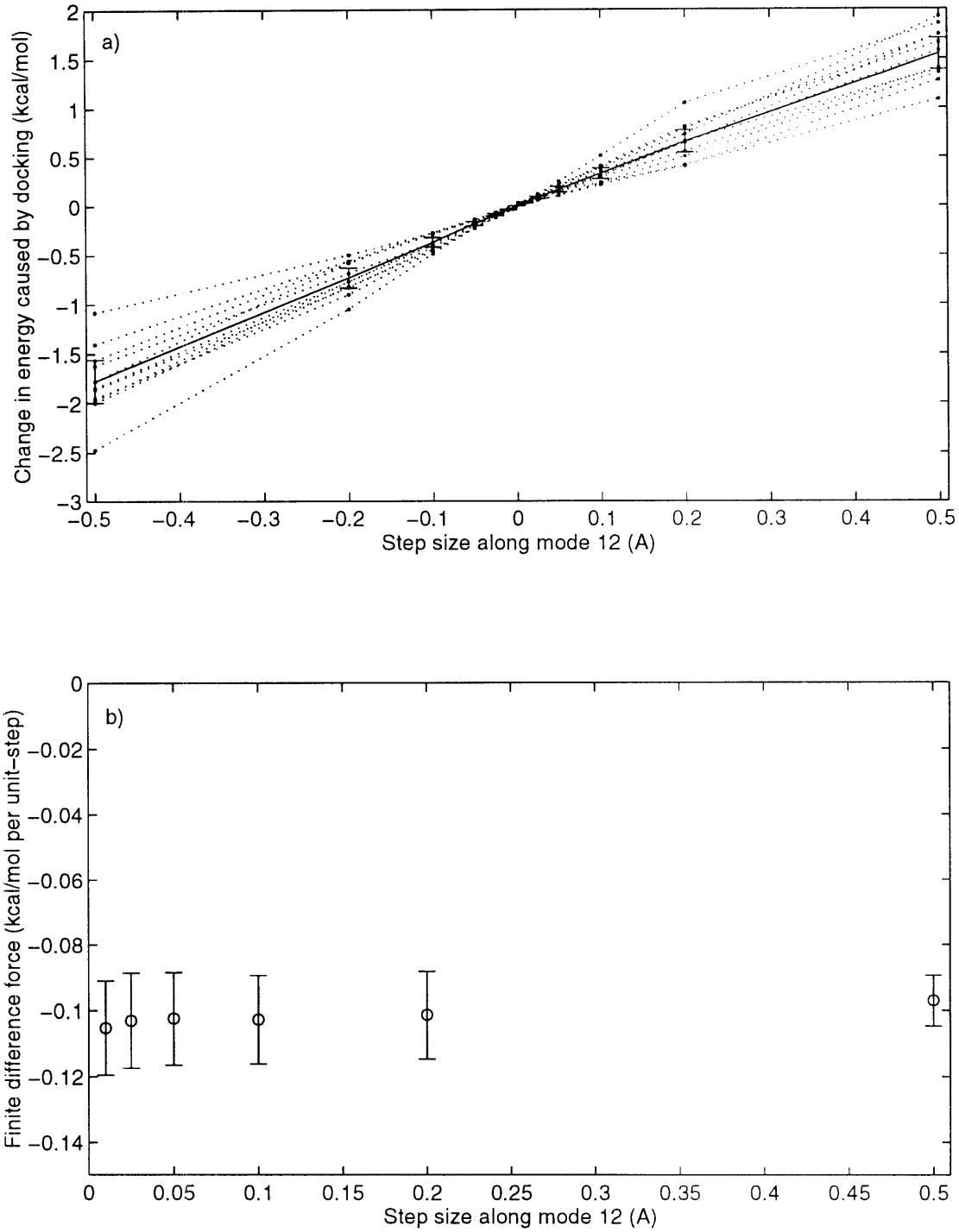


**Figure 8. Force of minor groove docking along mode 12.** (a) The energy change caused by the docking of the model protein is plotted for various step sizes along mode 12. The data from each of 11 rotations is shown with dotted lines, while the average and twice the standard deviation of the mean of the slopes of these calculations is shown with solid lines. The x–values are the maximum distance any single atom moved, and the y–values are shifted such that the unmoved energy for each rotation is zero. (b) The force along mode 12 for various step sizes. The negative of each energy change from part *a* divided by the step size gives the force according to Eq. 5. See text for details.

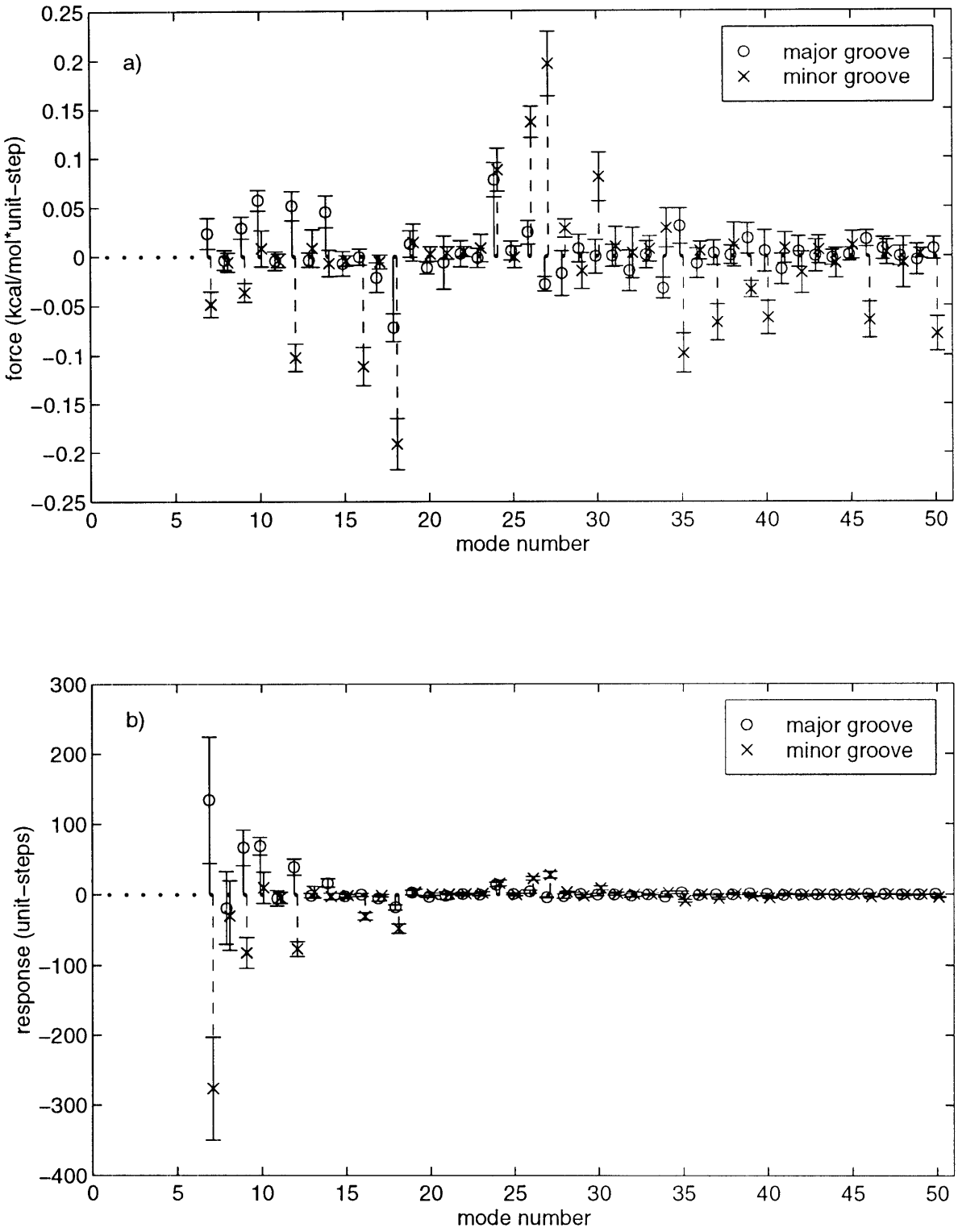


**Figure 9. Force and response to protein docking along the low-frequency normal modes.** (a) The average force along the first 50 modes caused by introducing the low-dielectric sphere into the major (circle, slightly left of x-value) and minor ("x", slightly right) grooves. (b) The response to this force. Error bars are twice the standard deviation of the mean of 11 rotations. See text for details.

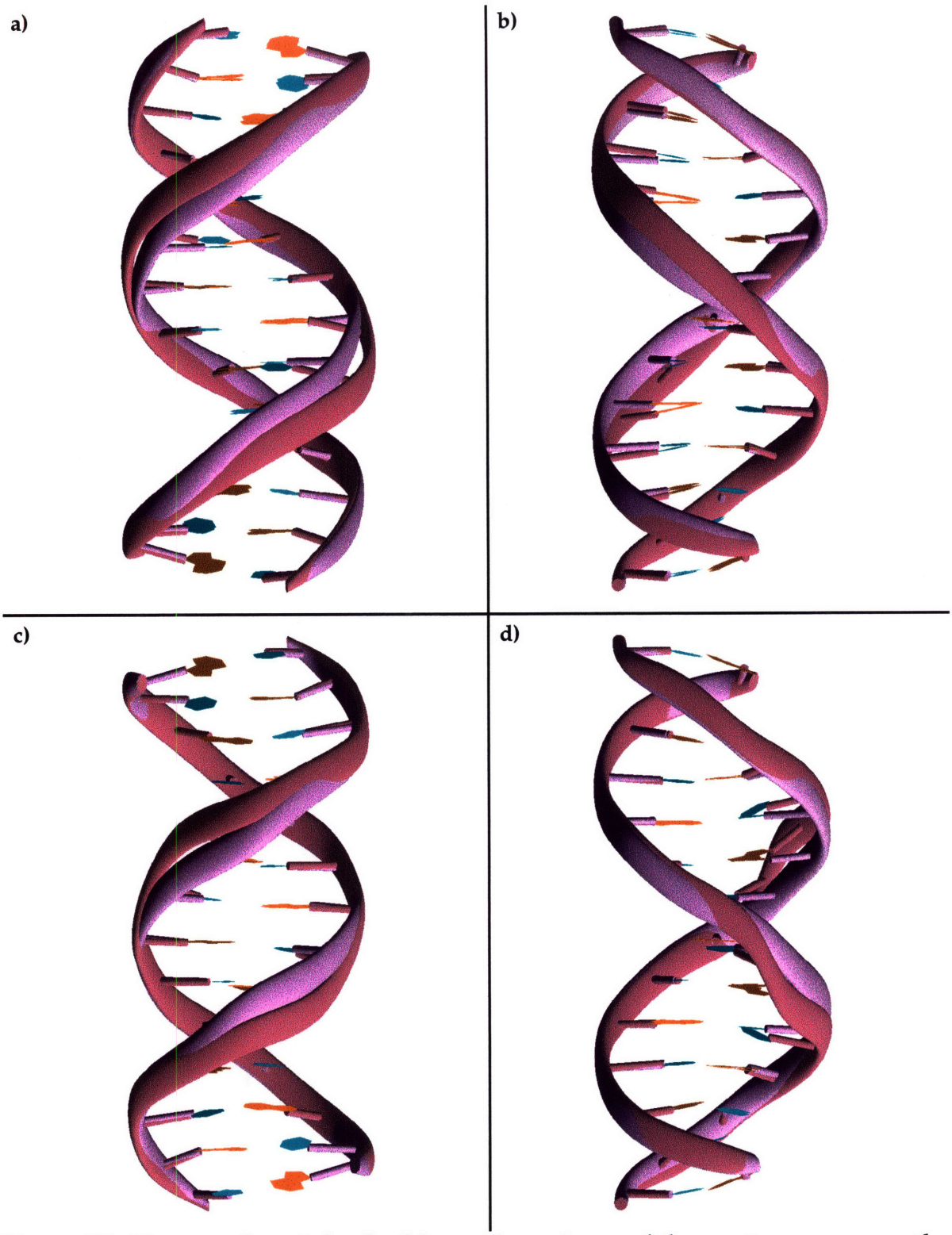


**Figure 10. Forces of protein docking.** Two views of the vector sum over the lowest 50 modes of the force for the major groove docking are shown in (a) and (b), and the minor groove docking is in (c) and (d). For each picture, the unmoved DNA is in purple, while the structure after a large step (the atom moving farthest moves 5A) is taken the direction of the appropriate force is in red.

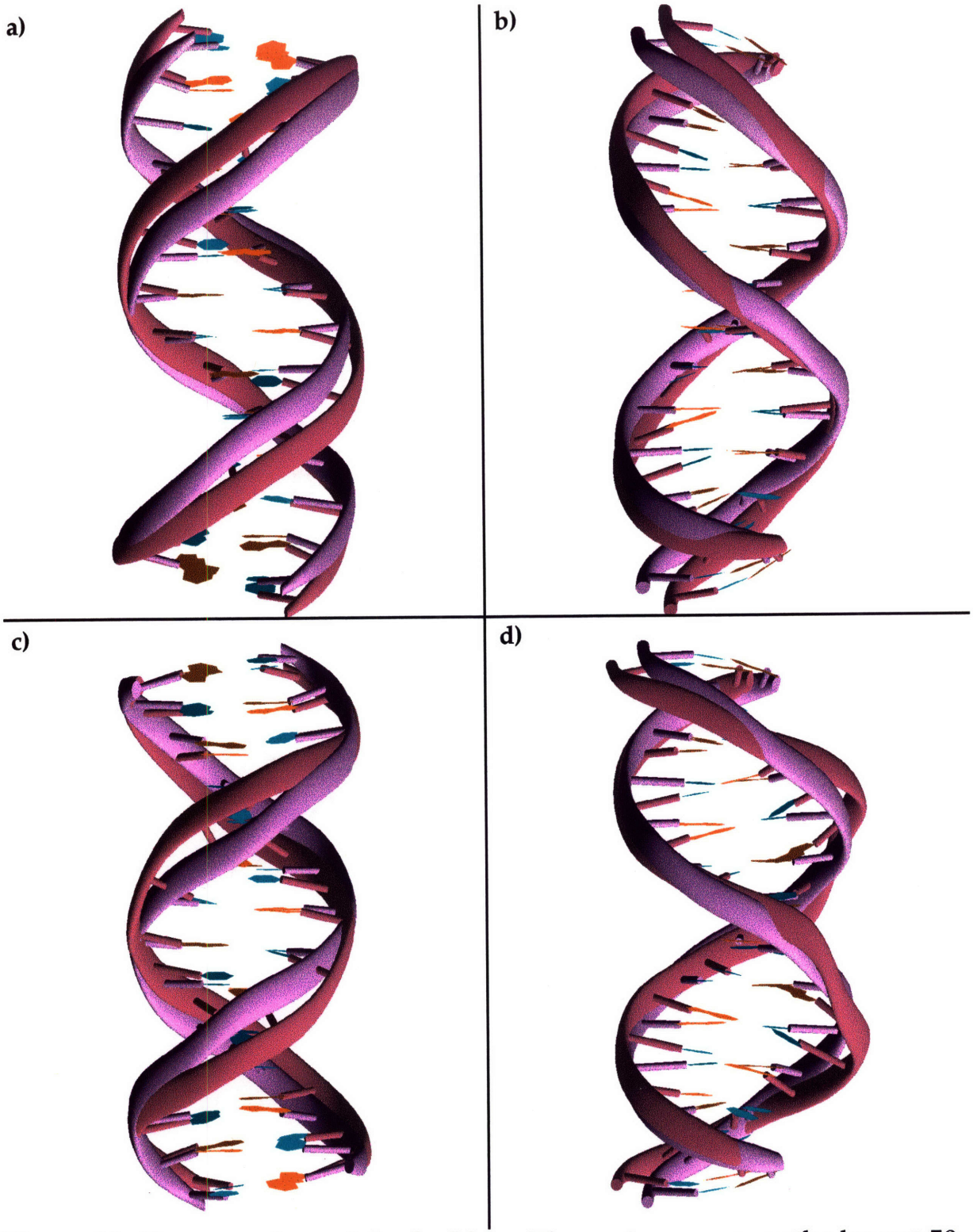


**Figure 11. Response to protein docking.** The vector sum over the lowest 50 modes of the response to the major groove docking is shown in (a) and (b), and the minor groove docking is in (c) and (d). For each picture, the unmoved DNA structure is shown in purple, and the structure after a large step (atom moving farthest moves 5A) is taken in the direction of the appropriate response is in red. See text for definition of the response.

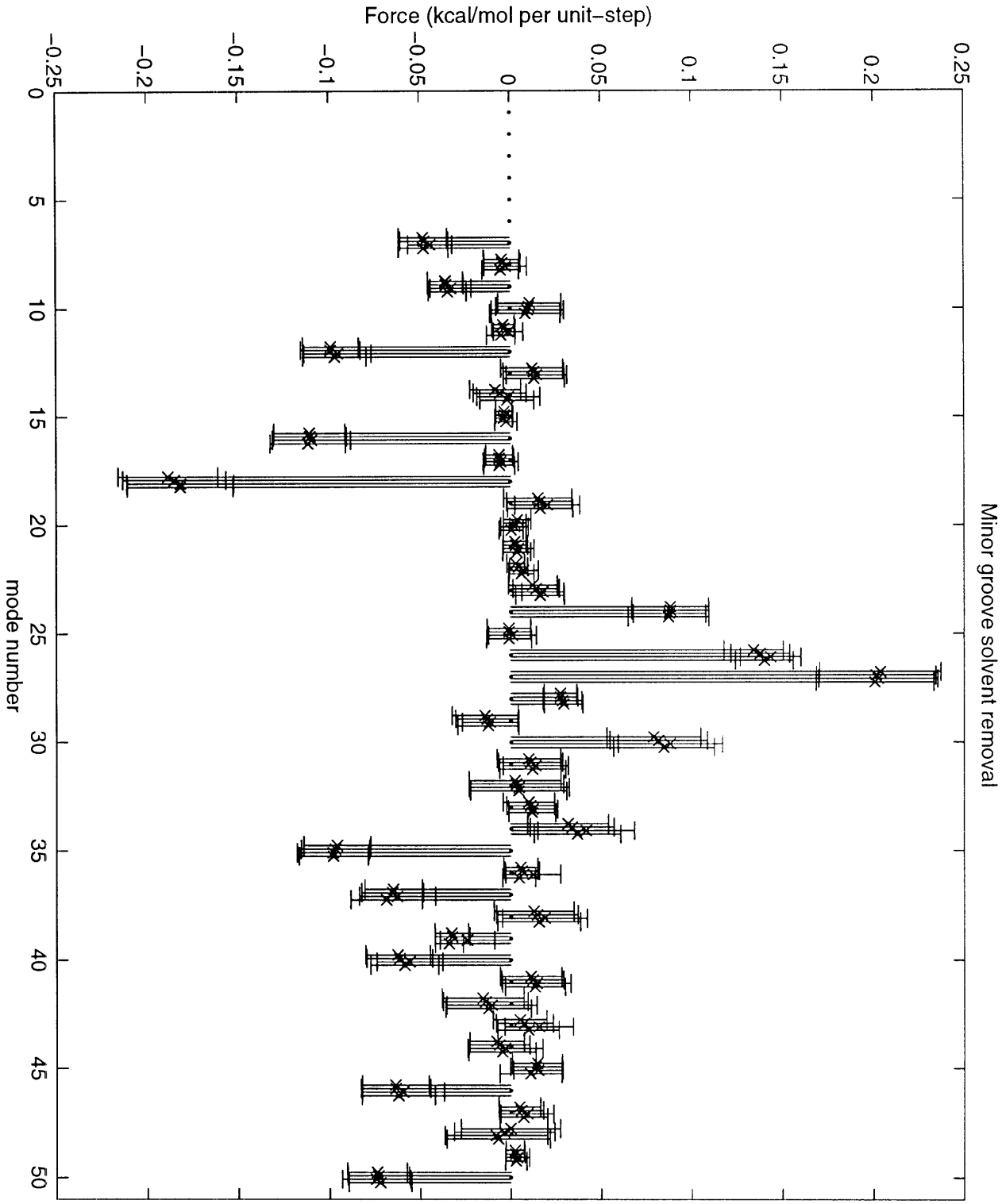


**Figure 12. Salt effects on the force of the minor groove protein docking.** The projection of this force along the lowest 50 modes was calculated at zero ionic strength (far left value for each mode), at 0.145M by solving the linear Poisson–Boltzmann equation (second from the left), and at 0.145M by solving the non–linear equation and computing only the first term in Eq. 7 (second from the right) and all terms from this equation (far right) for the total energy.

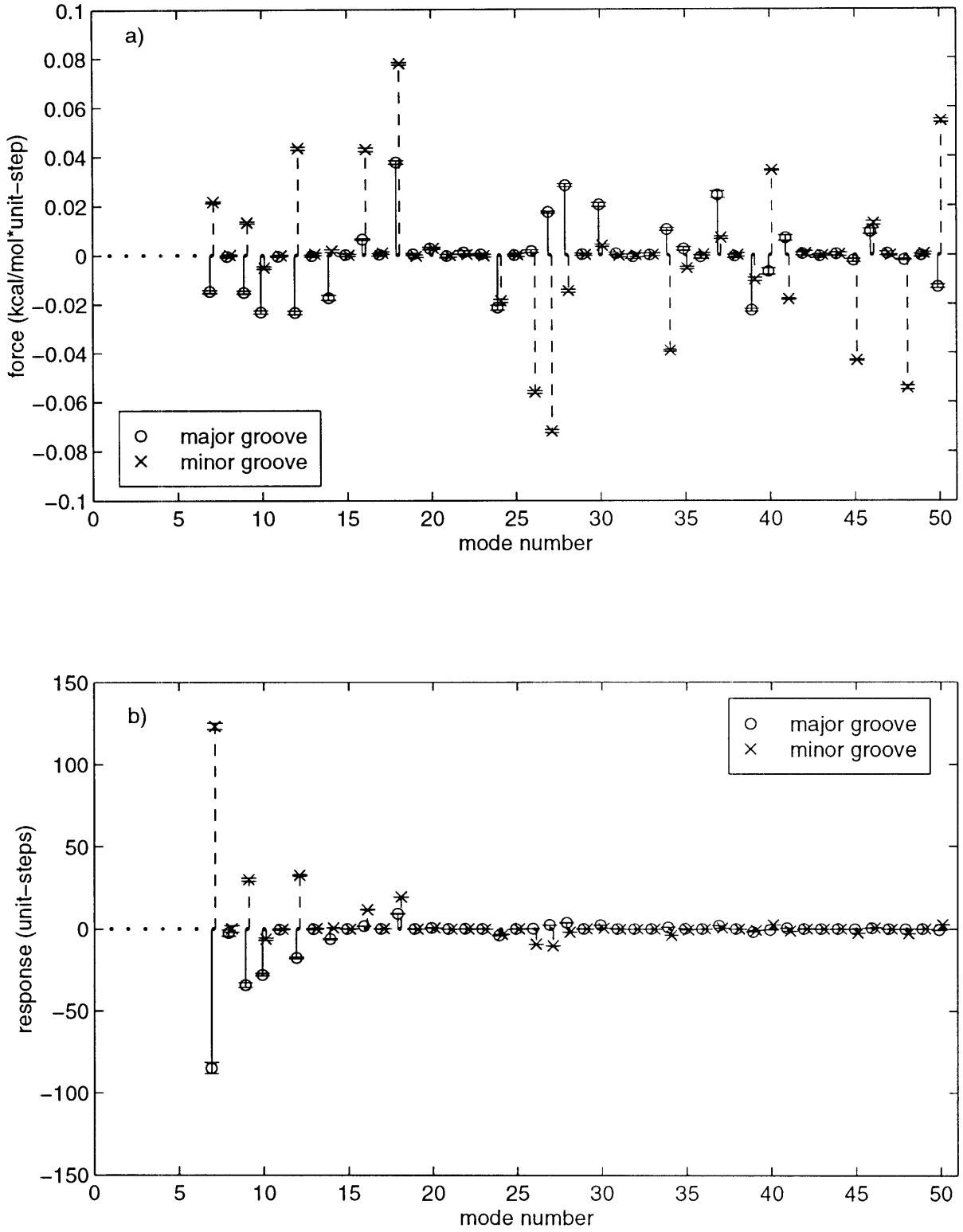


**Figure 13. Force and response of charging the docked protein.** (a) The force projected along the lowest 50 modes caused by adding a charge of magnitude 1*e* to the center of the major groove protein (circles, slightly left of the x–value) and the minor groove protein ("x"s, slightly right). (b) The response to that force. Error bars are twice the standard deviation of the mean of 11 calculations.

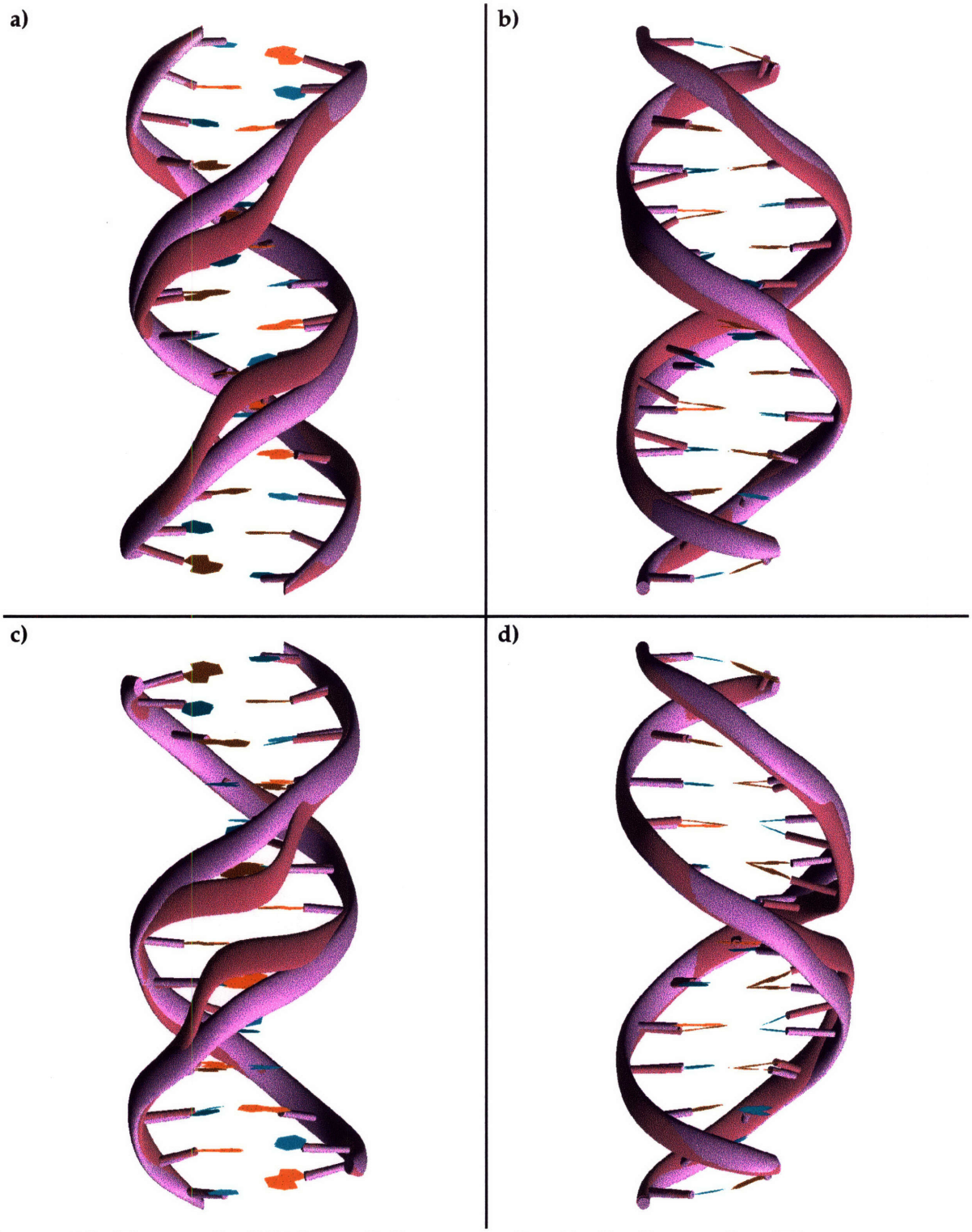


**Figure 14. Force of addition of charge to the docked protein.** The vector sum over the lowest 50 modes of the force for adding a charge of magnitude 1*e* to the center of the major groove protein is shown in (a) and (b), and the force the minor groove is in (c) and (d). For each figure, the unmoved DNA structure is in purple, and the structure after moving a large step (atom moving farthest moves 5Å) in the direction of the appropriate force is in red.

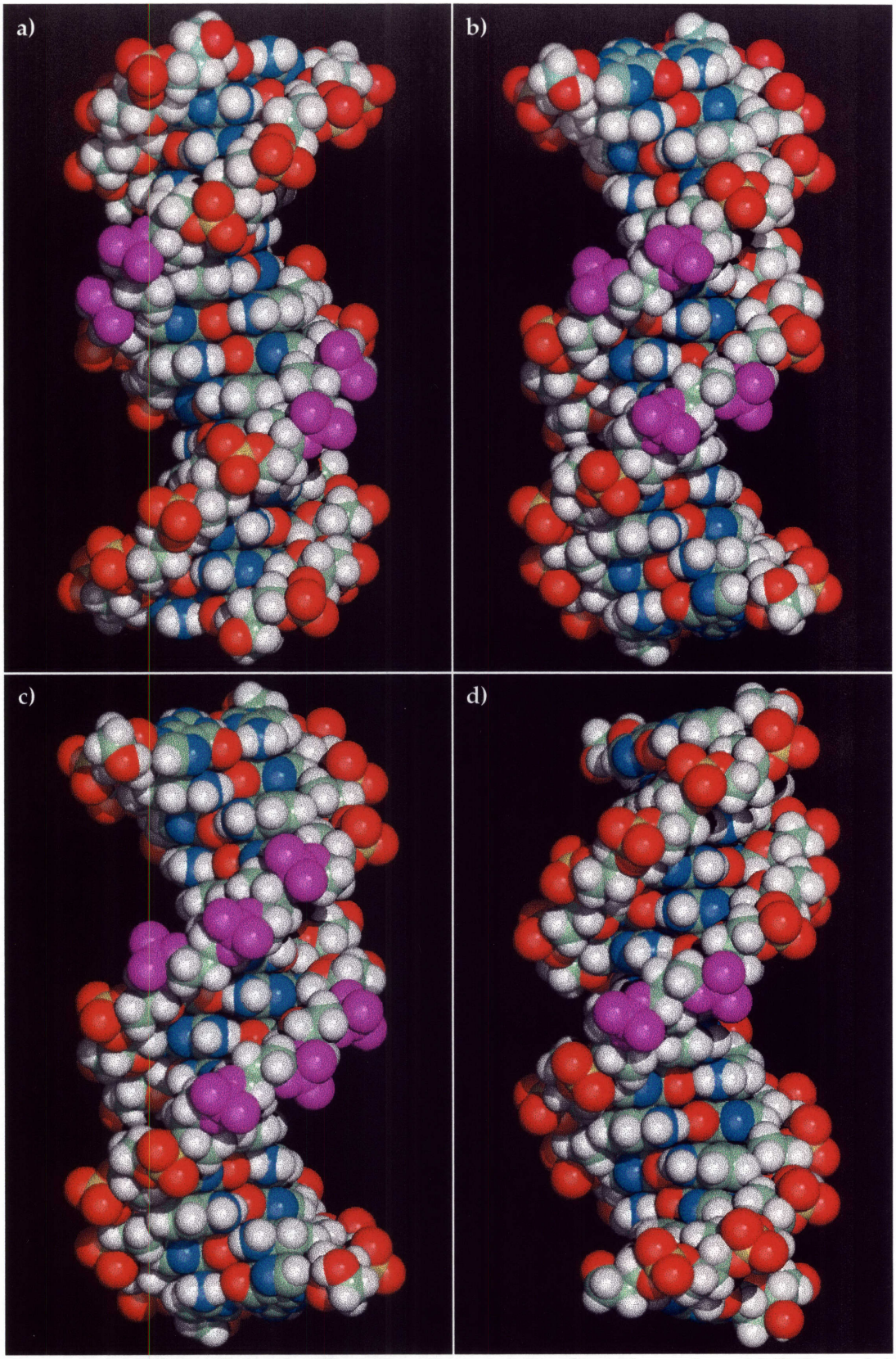


**Figure 15. Phosphates chosen for neutralization in 14GC.** (a) four phosphates across the major groove. (b) four phosphates across the minor groove. (c) six phosphates across the minor groove. (d) two phosphates on the same strand.

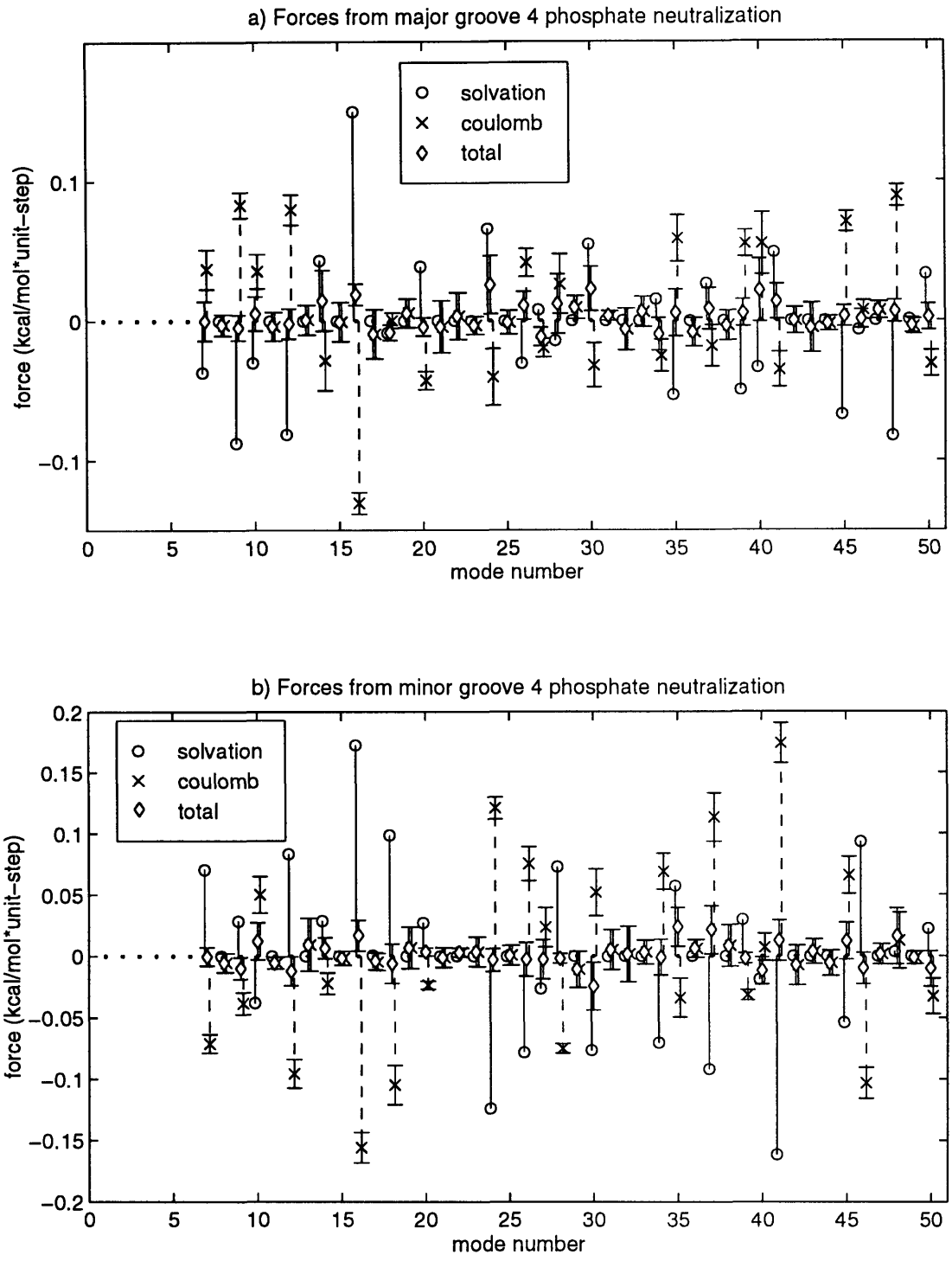


**Figure 16. Forces of phosphate neutralizations.** Each plot shows the average coulombic (circles, slightly left of the x–value), solvation ("x"s, slightly right), and total (diamonds) force caused by phosphate neutralization along the lowest 50 normal modes. The four different neutralizations (a), (b), (c), and (d) are shown in Fig. 15. Error bars for the solvation and total forces are twice the standard deviation of the mean of 11 calculations, and the projection of the coulombic force onto the normal modes is exact and has no error.

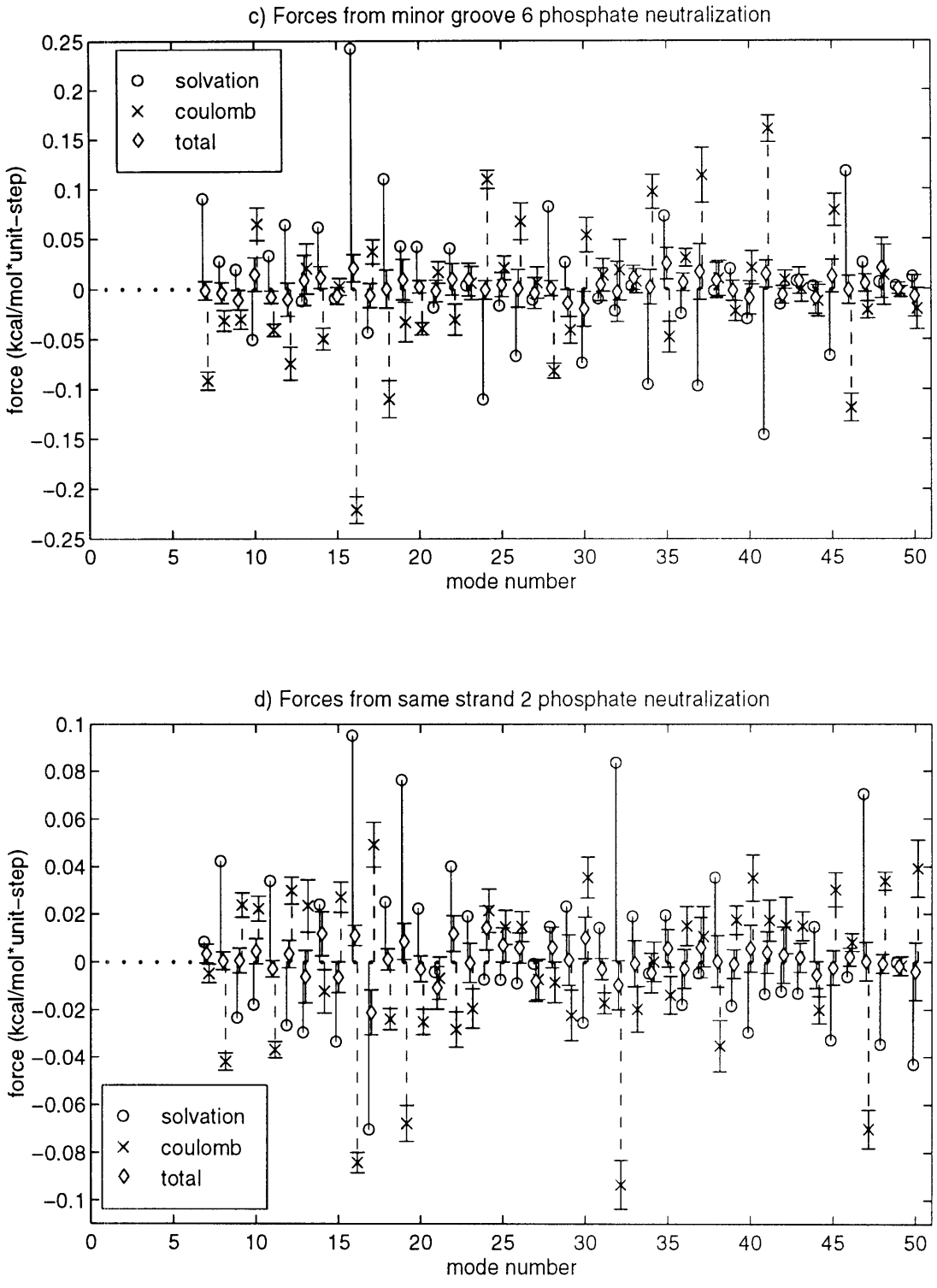


**Figure 16c and 16d.** For legend see part *a*.

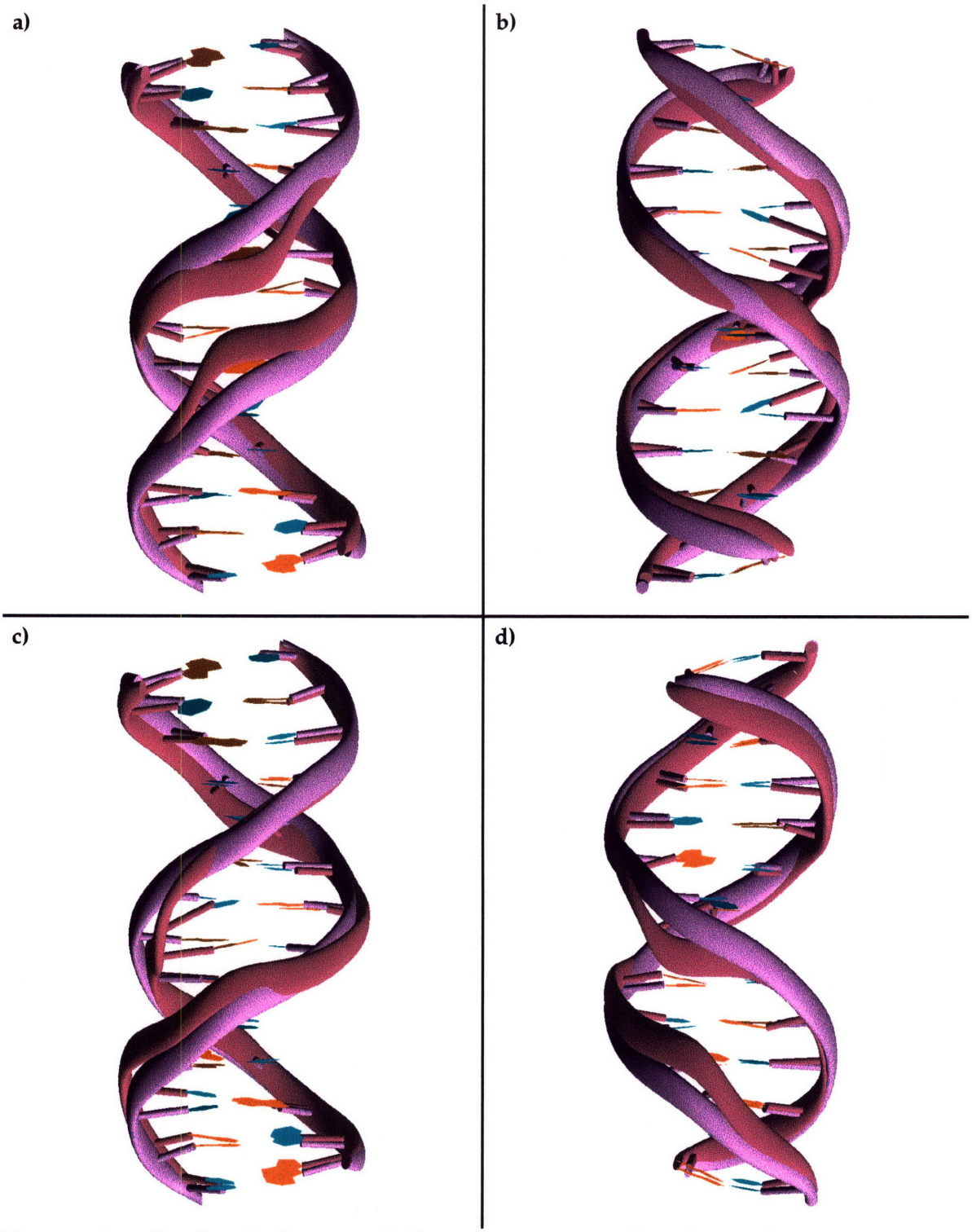


**Figure 17. Coulomb forces of phosphate neutralization.** The vector sum over the lowest 50 modes of the coulombic component of the force created by neutralization of the four minor–groove phosphates is shown in (a) and (b), and the force of the same–strand two phosphate neutralization is in (c) and (d). For each figure, the unmoved DNA structure is shown in purple, and the structure after a large step (atom moving farthest moves 5Å) is taken in the direction of the appropriate force is in red.

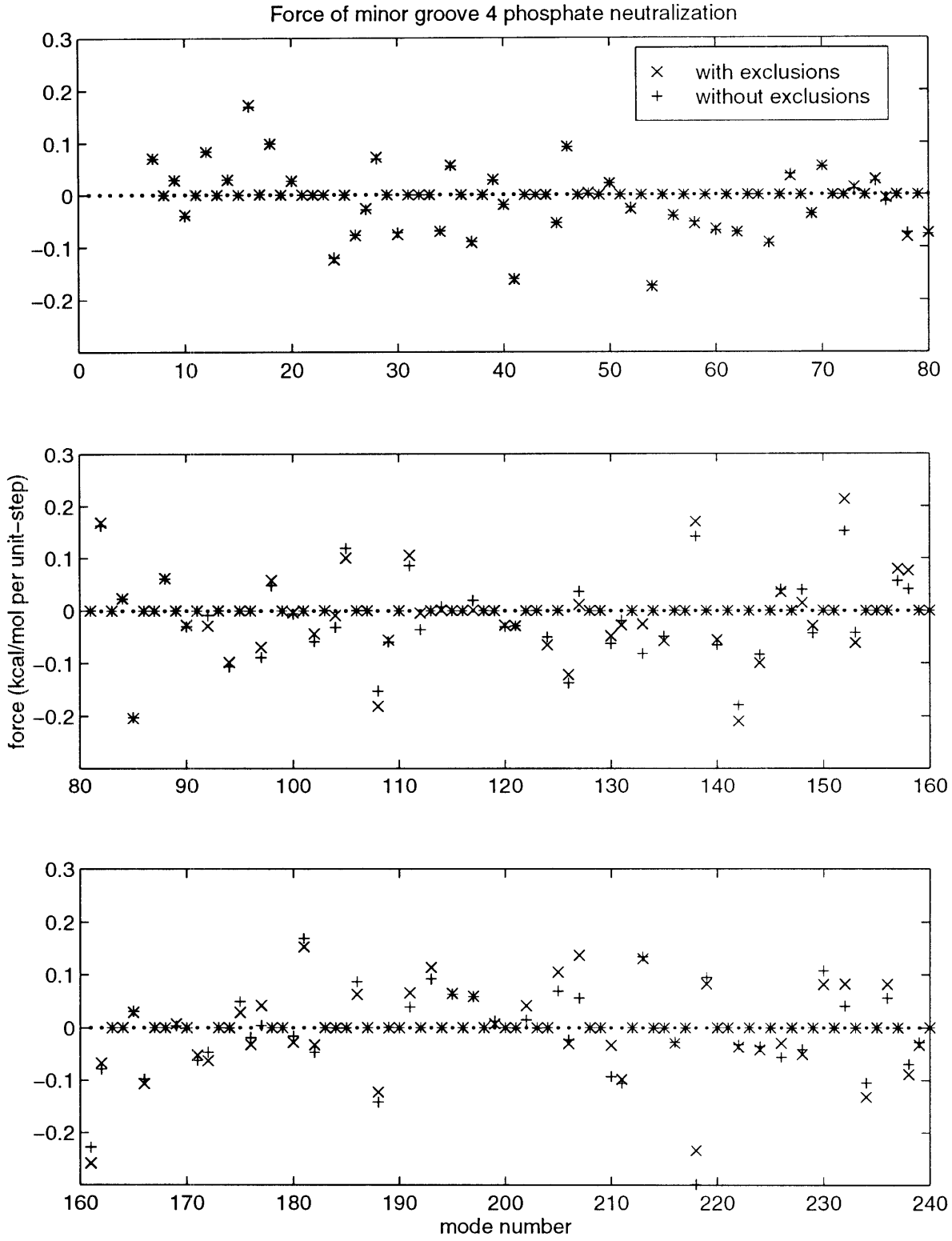


**Figure 18. Coulombic forces along the lowest normal modes.** The coulombic force of neutralizing the four minor groove phosphates is projected onto the normal modes. The projection when 1–2 and 1–3 nonbonded interactions are excluded from the force is shown with 'x', and the projection with all interactions included is shown with '+'.