

---

**A molecular mechanical model to predict the helix twist angles of B-DNA**

---

Chang-Shung Tung<sup>1</sup> and Stephen C. Harvey<sup>2+</sup>

---

<sup>1</sup>Department of Physics and <sup>2</sup>Department of Biochemistry, University of Alabama, Birmingham, AL 35294, USA

---

Received 15 December 1983; Accepted 2 March 1984

---

**ABSTRACT**

We present here a model for the prediction of helix twist angles in B-DNA, a model composed of a collection of torsional springs. Statistically averaged conformational energy calculations show that, for a specified basepair step, the basepair-basepair conformational energy is quadratically dependent on the helix twist angle, so the calculations provide the spring parameters for the basepair-basepair interactions. Torsional springs can also be used to model the effects of the backbone on the helix twist, and the parameters for those springs are derived by fitting the model to experimental data. The model predicts a macroscopic torsional stiffness and a longitudinal compressibility (Young's modulus) which are both in good agreement with experiment. One biological consequence of the model is examined, the sequence specificity of the Eco RI restriction endonuclease, and it is shown that the discriminatory power of the enzyme receives a substantial contribution from the energetic cost of torsional deformations of the DNA when wrong sequences are forced into the enzyme binding site.

**INTRODUCTION**

While it is not yet known how proteins recognize specific DNA sequences, it is clear that the DNA double helix is not perfectly regular with a fixed helix rotation from one step to the next. The single crystal x-ray analyses of both B-DNA (1-3) and A-DNA (4-7) have revealed large variations in the helix parameters. Furthermore, the twist angle of the helix has been shown to depend on sequence when the molecule is in solution, both by the effects on supercoiling parameters when short segments of known sequence are inserted into closed circular DNA (8-10) and by the nuclease digestion patterns of DNA adsorbed on surfaces (11,12). These variations in helix geometry may play an important role in the recognition of specific sequences by proteins.

We have recently begun to examine the effects of sequence on the geometry of the DNA double helix. Earlier models (13-16) had suggested that the helix twist angle of each basepair step would depend on the sequence, and we hoped that conformational energy calculations would provide a rigorous model, with atomic resolution, that could correctly predict helix twist angles of spe-

cified sequences. Here we describe our progress in developing that model.

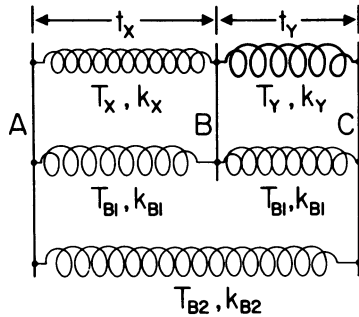
Two earlier models had been put forward for predicting the helix twist angles of a given DNA sequence. The model of Kabsch *et al.* (13) is entirely empirical, and it is summarized in a single table that gives the twist angle of each of the possible basepair steps. The second model is a physical one, based on a set of rules originally proposed by Calladine (14) after an examination of the base stacking patterns and stereochemical clashes in the DNA double helix. Calladine's rules are so simple and his arguments so convincing that Dickerson abandoned an earlier model of his own (2,15) to develop a quantitative model (3,16) based on those rules. It is important to note that the patterns of base stacking and propeller twisting predicted by Calladine's rules have been shown to occur both in crystals (3,16) and in solution (17).

Our model, like the Calladine/Dickerson model (14,16), is a physical one, but it is considerably more detailed than theirs, since we use conformational energy calculations to determine the parameters of the model. As a consequence, we can quantitatively predict twist angles, whereas the Calladine/Dickerson model will only predict the relative amplitudes of helix twist angles, not the absolute amplitudes. Furthermore, Calladine and Dickerson make a distinction only between purines and pyrimidines, but not between adenine and guanine (or between cytosine and thymine), while the atomic resolution of conformational energy calculations allows us to examine those subtle differences.

The remainder of this paper is divided into four parts. The first gives a general description of our model of helix twisting, explaining how we have separated the effects of interactions of successive basepairs from the contributions due to the backbone. The second section treats basepair-basepair interactions, while the third treats backbone contributions. The final section is the Discussion, in which we describe the strengths and shortcomings of our model in detail. We show that the model can account for the torsional stiffness and the Young's modulus of B-DNA, and we argue that the model is biologically relevant, because it provides a possible explanation for some of the sequence specificity of DNA-binding proteins.

### DESCRIPTION OF THE MODEL

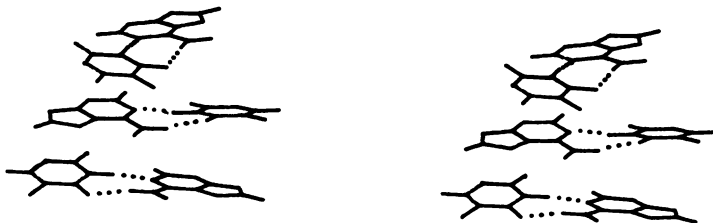
We have chosen to explore a model for helix twisting that separates the effects of basepair-basepair interactions from those arising from deformations of the backbone. For the former, we have treated individual bases as



**Fig. 1:** Torsional spring model for helix twist angles of a trimer. Note that distances in this figure correspond to angles, not to lengths. The trimer's sequence is ABC, so that step 1 is the AB step and step 2 is the step BC.  $t_x$  and  $t_y$  are the equilibrium twist angles for the two steps under the influence of the five torsional springs. The steric interactions of basepairs A and B produce a force field which is modeled by a single torsional spring with an equilibrium angle of  $T_x$  and a torsional constant  $k_x$ , and a similar situation between basepairs B and C gives  $T_y$  and  $k_y$ . Backbone deformations for the twisting of a single step are modeled by a spring characterized by  $T_{B1}$  and  $k_{B1}$ . Propagation effects from one step into the other are produced by a spring connecting basepairs A and C, with an equilibrium twist angle and torque constant  $T_{B2}$  and  $k_{B2}$ , respectively.

rigid bodies and calculated the variation in conformational energy as we varied interbasepair distances, helix twist angles, basepair roll angles, and the propeller twist angles. Those calculations, described in more detail in the next section, gave good qualitative agreement with Calladine's rules (14) for the effects of sequence on all of the basepair orientational parameters. Furthermore, when we plotted the dependence of energy on helix twist angle for all ten possible dimers, the curves turned out to be nearly quadratic, so it is plausible to model the torque between successive basepairs as they twist relative to one another by a single torsional spring per basepair step.

To model the effect of backbone stiffness in modulating the twist angle of the basepair step, we assume that over the small range of angles that is typical of variations in twist angle, the backbone deformations within a basepair step are also well modeled by a single torsional spring. While the equilibrium twist angle and the torque constant for the basepair-basepair interactions can be determined from the conformational energy calculations described above, the same parameters for the backbone ( $T_{B1}$  and  $k_{B1}$  in Fig. 1) are treated as parameters which must be determined when the model is fitted to data. If the model were left with only two springs per step (the upper four springs in Fig. 1), successive steps would be independent of one



**Fig. 2:** Stereoscopic view of one conformation of the trimer d(ATA)·d(TAT). All of the hydrogens are shown except those of the thymine methyl groups (see text for explanation). No atoms of the ribose-phosphate backbone are included. Note the nonzero roll and propeller twist angles.

another, and our model would be comparable to that of Kabsch *et al* (13). It has been shown, however, that a deviation from the average helix twist angle in one step is compensated by propagating fractional deviations of the opposite sign into neighboring steps (3,14). To guarantee such propagation, a torsional spring spanning two basepair steps is introduced with an equilibrium twist angle  $T_{B2}$  and a torsional constant of  $k_{B2}$  (Fig. 1). The resulting trimer is the simplest unit in which these propagation effects can be examined. We have left  $k_{B2}$  as a third backbone parameter to be determined by fitting the model to data, but we have chosen to assume that this second backbone spring attempts to maintain the total helical twist of two successive steps at a value twice that of a single step, by setting  $T_{B2} = 2T_{B1}$ . This choice was made simply to reduce the number of free parameters.

#### INTERACTIONS BETWEEN SUCCESSIVE BASEPAIRS

In order to describe the local conformations of basepairs within the double helix, we use a set of helix parameters similar to that of Dickerson and coworkers (3,16). For simplicity, we have kept the basepair tilt angle fixed at  $0^\circ$  and not allowed basepair sliding, so that the only parameters which are varied are interbasepair separation, helix twist angle, basepair roll angle, and propeller twist angle. The effects of this reduction in the degrees of freedom are considered in the Discussion.

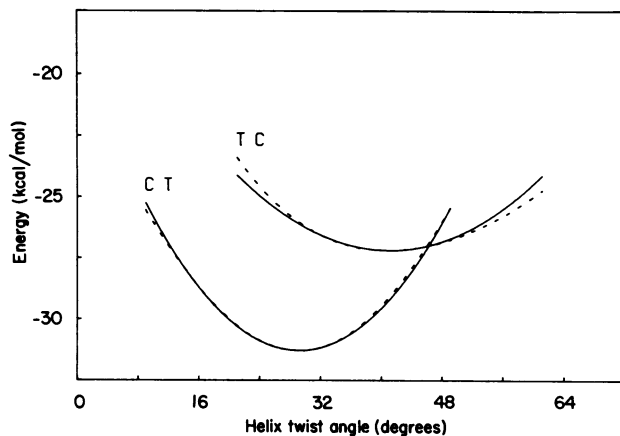
The detailed atomic model for the trimer, which permits the calculation of the basepair-basepair parameters ( $T_X$ ,  $k_X$ ,  $T_Y$ , and  $k_Y$  in Fig. 1), is shown for the trimer d(ATA)·d(TAT) in Fig. 2. As the conformational parameters described in the previous paragraph are varied, we can calculate the nonbonded energies if we have a suitable set of parameters for the van der Waals and electrostatic interactions. For the former, we have chosen those

proposed by Levitt (18), with slight modification. We have included explicit hydrogen atoms except on the thymine methyl group, where a single extended atom eliminates the steric complications that would arise from having to consider methyl rotations. These parameters have slightly larger van der Waals radii than those used in our previous studies (19,20) and they are particularly appropriate for static calculations like these (18). For electrostatic energies, we have used the partial atomic charges proposed by Miller (21) and a distance-dependent dielectric constant to mimic solvent effects,  $D(r) = r$ ,  $r$  being the interatomic distance measured in Å (19,20,22).

As a first approximation for the examination of how helix parameters depend on sequence, we began with trimers which were either alternating copolymers, such as d(ATA)·d(TAT) or strict homopolymers, such as d(AAA)·d(TTT). We chose a standard geometry for the ends of the trimer (23,24), setting the total helix twist angle between the first and third basepairs at 72°, the total base separation between these basepairs at 6.8Å, the roll angles of both end basepairs at 0°, and the propeller twist of the end basepairs at 17°. The total nonbonded energy was then calculated as the conformation of the middle basepair was varied, and the expectation values for angles and energies were calculated by averaging over all conformations, using the appropriate Boltzmann factors for weighting (25).

The results of the trimer calculations were satisfying on two accounts: they agreed qualitatively with Calladine's rules (14,16), and the conformational energies showed a nearly quadratic dependence on helix twist angle for the middle basepair, suggesting that torsional forces within the double helix might be simply modeled by torsional springs. Let us examine each of these points in more detail.

Calladine's rules (14,16) predict that the steric clashes between neighboring purines on opposite strands may be relieved by appropriate changes in helix twist angle, basepair roll angle, and propeller twist angle. As for helix twist, Calladine's rules predict that a pyrimidine-purine step will have a smaller twist angle ( $t_g$ ) than will a purine-pyrimidine step, and our trimer calculations showed that in the alternating copolymers the former have an average  $t_g$  of 13°, while the latter have an average  $t_g$  of 42°, exactly as predicted. (Note that the absence of any backbone modulating effects at this stage of the calculations allows very large deviations in helix twist.) Regarding roll angles, the trimer calculations showed changes of roll angle in the alternating copolymers of 3°-5°, positive for pyrimidine-purine steps (open to the minor groove) and negative for purine-pyrimidine



**Fig. 3:** Conformational energy for d(CTC)·d(GAG) as a function of helix twist angle, partitioned into the CT and TC steps. The dashed lines represent values from the conformational energy calculations, while the solid lines are the best fit quadratic curves, from which the equilibrium twist angles and the torsional spring constants can be determined. The best fit curve for a given basepair step was determined by finding the minimum energy and the twist angle corresponding to that energy; those values were held fixed while the value of the torsional spring constant,  $k$ , was varied to find the minimum root mean square difference between the data and the fitted curve over a range of  $20^\circ$  on either side of the minimum. Note that the equilibrium twist angle of the CT step is smaller than that of the TC step, a subtlety observed by Kabsch (13) but not predicted by Calladine's rules (14).

steps (open to the major groove), while for the homopolymers, the changes in roll angle were less than  $1^\circ$ , all in agreement with Calladine's predictions. Finally, the propeller twists ranged from  $9.4^\circ$  to  $18.5^\circ$  (agreeing with values from crystallography (16) and from solution studies (17,24)), with the propeller twist of a given basepair always being larger when it was the central basepair in a homopolymer than when it was in an alternating copolymer. This difference, which averaged  $4.0^\circ$ , is also in agreement with Calladine's rules. The agreement between the trimer calculations and the rules proposed by Calladine is encouraging and suggests that conformational energy calculations should prove useful in examining the interplay of the various mechanisms for releasing steric stress.

The total conformational energy for the trimers was observed to vary approximately quadratically with the helix twist angle of the mobile central basepair. When the roll angle and propeller twist angle were held fixed at their minimum energy values and the energy was partitioned into two parts (interactions between the central basepair and the bottom basepair; interactions between the central basepair and the upper basepair), a quadratic

dependence was also observed for the dependence of energy on helix twist angle of each step (Fig. 3). A series of plots (one for each basepair step) was generated, and for each one the best fit quadratic curve provides the equilibrium twist angle and the torsional constant for the basepair interactions of that step. These parameters are summarized in Table 1A.

#### CONTRIBUTIONS FROM THE BACKBONE

The simplest model in which the torsional interactions of the basepairs are modulated by backbone contributions would consist of two parallel torsional springs per basepair step, with successive steps connected in series; a trimer for that model would contain the upper four springs in Fig. 1. Adding the fifth spring, spanning the two steps, gives the simplest trimer model in which the backbone also plays a role in propagating torsional stresses from one step into the other. The equilibrium backbone twist angles and torsional constants can be determined by using the base-base parameters (Table 1A) and fitting the model to experimental data.

A given set of spring parameters for the model in Fig. 1 will produce equilibrium helix twist angles  $t_X$  and  $t_Y$ , but it remains to be determined how  $t_X$ , for instance, depends on the intrinsic twist angle of the step AB and on the propagation effect from the step BC. If there were no propagation, with each step independent (13), we would have

$$t_X = t_{AB} \quad (1)$$

In the case of full propagation (14,16), however, we would have

$$t_X = t_{AB} - 0.5(t_{BC} - \bar{t}) \quad (2)$$

where  $\bar{t}$  is the mean helix twist angle of the ten different basepair steps. We can treat the most general case by introducing the propagation parameter,  $\alpha$  ( $0 < \alpha < 0.5$ ), to represent the extent to which a deviation from the mean helix twist angle in a given step is offset by compensating deviations of opposite sign in the two neighboring steps. With this parameter, equations (1) and (2) are just special subcases of the general case, namely

$$t_X = t_{AB} - \alpha(t_{BC} - \bar{t}) \quad (3)$$

Our model thus has five undetermined parameters, the four backbone spring parameters (Fig. 1) and the propagation parameter. This number is reduced to four by our assumption that  $T_{B2} = 2T_{B1}$ . The model has been fitted to experiment by a grid search in which the four parameters were varied over wide ranges, finding those for which the root mean square deviation between the calculated and observed helix twist angles is a minimum.  $t_X$  and  $t_Y$  were

Table 1 - Spring Parameters

1A - Spring parameters for basepair-basepair steps (from conformational energy calculations)										
Step:	TT	TA	AT	TC	CT	TG	GT	CG	GC	CC
T <sub>1</sub> (degrees)	42.5	16.8	39.5	41.3	28.8	9.5	43.8	12.5	42.0	29.8
k <sub>1</sub> (kcal·mol <sup>-1</sup> ·deg <sup>-2</sup> )	0.0175	0.0266	0.0151	0.0159	0.0307	0.0226	0.0178	0.0222	0.0185	0.0390
1B - Spring parameters for backbone (from optimum fit to experimental data)										
	T <sub>B1</sub>	k <sub>B1</sub>	k <sub>B2</sub>	$\bar{t}$	$\alpha$					
	(degrees)	(kcal·mol <sup>-1</sup> ·deg <sup>-2</sup> )	(kcal·mol <sup>-1</sup> ·deg <sup>-2</sup> )	(degrees)	(degrees)					
Best fit to solution data <sup>1,3</sup>	35.5	0.021	0.016	34.0	0.35					
Best fit to crystal structure <sup>3</sup>	36.0	0.041	0.056	35.1	0.50					

Table 2 - Parameters for Prediction of Helix Twist Angles with Equation 4

	$\Delta t_{TT}$	$\Delta t_{TA}$	$\Delta t_{AT}$	$\Delta t_{TC}$	$\Delta t_{CT}$	$\Delta t_{TG}$	$\Delta t_{GT}$	$\Delta t_{CG}$	$\Delta t_{GC}$	$\Delta t_{CC}$	$\bar{t}$	$\alpha$
Solution data:	4.9°	-6.3°	4.2°	4.5°	-0.9°	-8.8°	5.9°	-7.4°	5.1°	-1.2°	34.0°	0.35
Crystal data:	3.3°	-3.2°	1.6°	2.4°	-0.1°	-4.1°	2.3°	-3.4°	2.1°	-0.9°	35.1°	0.50



calculated (Fig. 1) for each set of spring parameter values, and the ten basepair step deviations  $\Delta t_i$  ( $i = 1,10$ ) were then determined by noting that if the basepair step AB in Fig. 1 corresponds to the index  $i$  and if step BC corresponds to the index  $j$ , then equation 3 gives

$$t_X = \bar{t} + \Delta t_i - \alpha \Delta t_j$$

Similarly,

$$t_Y = \bar{t} + \Delta t_j - \alpha \Delta t_i$$

where  $t_i$  is the intrinsic twist angle of step  $i$ , and  $\Delta t_i = t_i - \bar{t}$ . The ten deviations and the value of  $\alpha$  allow the prediction of the helix twist angles for a DNA of specified sequence: step  $n$  will have a twist angle given by

$$t(n) = \bar{t} + \Delta t(n) - \alpha \Delta t(n-1) - \alpha \Delta t(n+1). \quad (4)$$

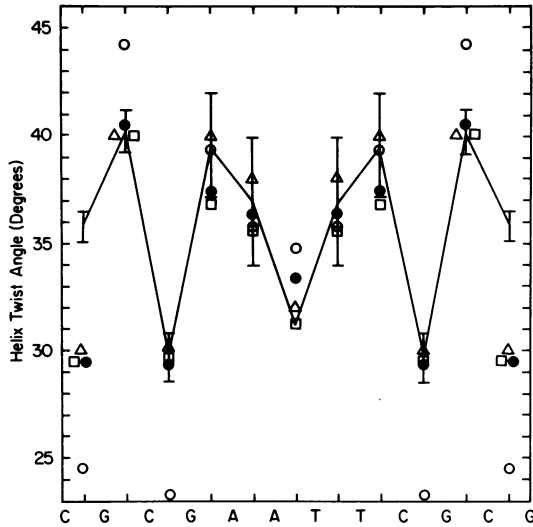
Note that if  $\alpha = 0$ , equation (4) reduces to Kabsch's model (13), for which there is no propagation and the helix twist angle of each step is independent of all others. When  $\alpha = 0.5$ , equation (4) is identical to the Calladine/Dickerson model (14,16).

We have fitted two sets of data, producing the two sets of backbone parameters given in Table 1B and the two sets of basepair helix twist angle deviations given in Table 2. The first set of data is the same set which Kabsch et al. (13) used, while the crystal structure (3,16) of the Dickerson dodecamer d(CGCGAATT(Br<sup>5</sup>C)GCG) provided the second set.

## DISCUSSION

This remarkably simple physical model gives a surprisingly good fit to the solution data. With only four parameters, we get a root mean square deviation of  $0.48^\circ$  from the data of Table I of reference 13, while the Kabsch model (13), with ten parameters, provides an r.m.s. deviation of  $0.37^\circ$  (we have used the same weighting factors,  $1/\sigma_{\theta}^2$ ). The difference in these fits is not statistically significant, so the fact that our model has fewer parameters and the fact that our parameters represent definite physical quantities would support the utility of our model. The best fit value for  $\alpha$ , 0.35, indicates that models with no propagation (13) and full propagation (14) are both oversimplifications. Dickerson himself pointed this out (16), although he confined his investigations to the case of full propagation.

When we plot the predicted twist angles for the Dickerson dodecanucleotide using our best fit parameters for the solution data (row 1 of Table 2), we come up with a fit which is qualitatively correct but which is quantitatively not very good (open circles of Fig. 4). If, however, we fit our model



**Fig. 4:** Comparison of the observed helix twist angles for the dodecamer CGCGAATTTCGCG with those predicted by several models. The experimental data are those labeled MPD7 of Table M1 of reference 3: the solid line connects the end-for-end average values, while the error bars show the ranges. (For example, the angles for the steps G4-A5 and T8-C9 are 37.1° and 41.9°, respectively, giving an end-for-end average of 39.5° for those steps.) The model values are the Calladine/Dickerson model (reference 3, indicated by triangles); the Kabsch model (reference 13, indicated by squares); our model optimized to fit the solution data of reference 13 (indicated by open circles); and our model optimized to fit the middle nine steps of this dodecamer (indicated by filled circles).

directly to the dodecanucleotide data (row 2 of Table 2), the predictions provide a good fit (closed circles of Fig. 4). The principal deviations are in the terminal CG steps (where none of the three models fits the data, probably because of end effects) and in the central AT step (where our homopolymers and alternating copolymers are poor models for the AATT sequence). It is interesting to note that the best value for the propagation factor in this case is  $\alpha = 0.5$ , the value which Dickerson used.

There are several advantages to this model. It combines the best feature of the Kabsch model (both are quantitatively predictive, where the model of Calladine and Dickerson predicts only relative step sizes) with that of the Calladine/Dickerson model (both are physical models, where Kabsch's is an empirical model). It goes somewhat beyond the latter model, however, because it treats the basepairs in atomic detail, distinguishing between an adenine and a guanine, for instance. (Calladine and Dickerson only recognize

differences between purines and pyrimidines.) It would be a straightforward extension of the model to treat modified bases, shedding light on the subtle structural changes produced by such modifications. Also in contrast with the previous models (13-16), the parameters which remain to be adjusted in fitting the model to data represent simple physical quantities. Consequently they provide other predictions for testing the model - such as a value for the torsional elastic modulus, discussed below. One final advantage is that the model treats the propagation effects of steric clashes along the helix in detail, exploring the region between the model with no propagation (13) and that with full propagation (14,16).

The model has one weakness. While it was to be hoped that a single set of parameters would suffice for the prediction of helix twist angles, the results of the dodecanucleotide calculations (Fig. 4) clearly show that this is not the case. We can attribute the differences in the sets of predicted twist angles and propagation factors (Table 2) to differences in environment (solution vs crystal), but we cannot yet explain how helix twist parameters depend on environmental factors. For a given environment, however, the model makes unambiguous predictions of helix twist parameters.

The model provides two predictions about the gross physical properties of B-DNA. First, the torsional elastic modulus can be calculated by applying the parameters of Table 1 to the model in Fig. 1. With an average torsional constant of about  $0.02 \text{ kcal}\cdot\text{mole}^{-1}\cdot\text{degree}^{-2}$  for the basepair steps, and using the backbone torsional constants for the solution data, we obtain a value of  $5.7 \times 10^{-19} \text{ erg}\cdot\text{cm}$  for the total torsional constant. (The corresponding value if we use the backbone torsional constant for the x-ray data is  $13 \times 10^{-19} \text{ erg}\cdot\text{cm}$ .) This compares very favorably with the values of  $0.4 - 4.0 \times 10^{-19} \text{ erg}\cdot\text{cm}$  from fluorescence depolarization (26-29),  $0.6 - 1.1 \times 10^{-19} \text{ erg}\cdot\text{cm}$  from supercoiling studies (30-32) (which may be low by a factor of four; ref. 33),  $0.4 - 14.0 \times 10^{-19} \text{ erg}\cdot\text{cm}$  from triplet anisotropy decay (34),  $3.1 \times 10^{-19} \text{ erg}\cdot\text{cm}$  from electron paramagnetic resonance measurements (35), and  $2.4 \times 10^{-19} \text{ erg}\cdot\text{cm}$  from studies on the cyclization probability (36). Note that the relative stiffnesses of our torsional springs implies roughly equal contributions to the torsional stiffness from the basepairs rotating relative to one another and from backbone contributions.

Second, the longitudinal elastic modulus (Young's modulus) can be estimated for the stack of basepairs, since we allowed interbasepair separation to vary in the conformational energy calculations. Our model value, which does not include the backbone effect, is  $13 \times 10^{10} \text{ dyne}\cdot\text{cm}^{-2}$ , a factor of

4-200 times higher than the values obtained from the acoustic velocity ( $3.4 \times 10^{10}$  dyne $\cdot$ cm $^{-2}$ ; ref. 37), from the persistence length ( $0.5 \times 10^{10}$  dyne $\cdot$ cm $^{-2}$ ; ref. 38), and from triplet anisotropy decay ( $0.07 - 0.29 \times 10^{10}$  dyne $\cdot$ cm $^{-2}$ ; ref. 34). This difference and the high value of the torsional constant in the previous paragraph probably arise because we have constrained the end basepairs of the trimers to fixed geometries and because we have not allowed the central basepair to tilt and slide to relieve the steric stresses; the additional degrees of freedom should produce a softer trimer and a lower value for both the Young's modulus and the torsional constant. Even at this stage of the model, however, our results suggest that the core of the DNA double helix is responsible for much of the longitudinal stiffness of the molecule, in contrast with the belief that the stiffness arises primarily from long-range electrostatic interactions in the backbone (39).

Our results indicate that conformational energy calculations can provide information on the sequence dependence of helix parameters. Earlier calculations by Kollman *et al.* (40) examined the helix twist of  $dA_{12}\cdot dT_{12}$  and  $d(CGCGAATTCGCG)_2$  using an all atom model and no adjustable parameters. Although there was a hint of a smaller helix repeat for  $dA_{12}\cdot dT_{12}$ , no definitive difference in helix repeats was reported, presumably because of the difficulties in accurately representing the phosphate repulsions. We believe our approach circumvents this difficulty in a straightforward manner. Kollman's observation (40,41) that there is only a weak coupling between the displacements of the bases and the displacements of the backbone supports our separation of base-base and backbone effects.

As a test of the biological relevance of the model, we have asked the following question: if a protein were designed to bind to a specified DNA sequence by recognizing the helix twist angles of that sequence, how much energy would be required to deform the DNA helix if a different sequence were forced into the binding site? While this approach ignores protein flexibility, and while it does not consider the probable role of other forces between the protein and the DNA, it should provide an indication of whether or not the dependence of helix twist on sequence is important in the recognition process. As an example, we chose the Eco RI restriction endonuclease (42,43). This enzyme recognizes the hexanucleotide sequence  $d(GAATTC)$ , so we used the parameters of Table 2 to determine the energetic cost of deforming a helix of the wrong sequence,  $d(CAATTG)$ , so that its twist angles match those of the correct sequence. The resulting energy difference,  $8.5 \text{ kcal}\cdot\text{mol}^{-1}$ , is surprisingly large. Even if we reduce this number by a factor of two, to

compensate for the fact that our calculated total torsional stiffness is somewhat higher than the experimental value (see above), we still obtain an energy difference that would reduce the binding constant of the wrong sequence by over three orders of magnitude relative to the correct sequence. The size of this effect suggests the importance of helix twist angle in the recognition of specific DNA sequences by proteins.

The ability to fit the available data on the helix twist angles of B-DNA with this simple model, the fact that it predicts macroscopic elasticity parameters in reasonable agreement with experiment, and the possible explanation for the sequence specificity of DNA-binding proteins are all very encouraging. The present model can be used to predict the helix twist angles for any sequence by using equation (4) and the parameters of Table 2. We hope to refine it and use it to examine the subtle effects of sequence differences and base modifications on helix parameters, thereby establishing a method for helping to investigate how these factors influence protein-DNA interactions.

#### ACKNOWLEDGMENTS

We thank Dr. Gayle Knapp for suggesting the Eco RI endonuclease calculation. This work was supported by grant PCM-8118827 from the National Science Foundation.

+To whom correspondence should be addressed

#### REFERENCES

1. Wing, R. et al. (1980) *Nature* 287, 755-758.
2. Dickerson, R.E. and Drew, H.R. (1981) *J. Mol. Biol.* 149, 761-786.
3. Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) *J. Biol. Chem.* 257, 14686-14707.
4. Shakked, Z. et al. (1981) *Proc. Roy. Soc. Ser. B* 213, 479-487.
5. Wang, A.H.J., Fujii, S., van Boom, J.H. and Rich, A. (1982) *Proc. Natl. Acad. Sci. USA* 79, 3968-3972.
6. Conner, B.N. et al. (1982) *Nature* 295, 294-299.
7. Shakked, Z. et al. (1983) *J. Mol. Biol.* 166, 183-201.
8. Wang, J.C. (1979) *Proc. Natl. Acad. Sci. USA* 76, 200-203.
9. Peck, L.J. and Wang, J.C. (1981) *Nature* 292, 375-378.
10. Strauss, F., Gaillard, C. and Prunell, A. (1981) *Eur. J. Biochem.* 118, 215-222.
11. Rhodes, D. and Klug, A. (1981) *Nature* 292, 378-380.
12. Behe, M., Zimmerman, S. and Felsenfeld, G. (1981) *Nature* 293, 233-235.
13. Kabsch, W., Sander, C. and Trifonov, E.N. (1982) *Nucl. Acids Res.* 10, 1097-1104.
14. Calladine, C.R. (1982) *J. Mol. Biol.* 161, 343-352.
15. Dickerson, R.E. and Drew, H.R. (1981) *Proc. Natl. Acad. Sci. USA* 78, 7318-7322.

16. Dickerson, R.E. (1983) *J. Mol. Biol.* 166, 419-441.
17. Patel, D.J., Kozlowski, S.A. and Bhatt, R. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3908-3912.
18. Levitt, M. (1983) *Gold Spring Harbor Symp. Quant. Biol.* 47, 251-262.
19. Harvey, S.C. and McCammon, J.A. (1981) *Nature* 294, 286-287.
20. Prabhakaran, M., Harvey, S.C., Mao, B. and McCammon, J.A. (1983) *J. Biomolec. Struct. Dyns.* 1, 357-369.
21. Miller, K.J. (1979) *Biopolymers* 18, 959-980.
22. McCammon, J.A., Wolynes, P.G. and Karplus, M. (1979) *Biochemistry* 18, 927-942.
23. Dickerson, R.E. *et al.* (1982) *Science* 216, 475-485.
24. Wu, H.M., Dattagupta, N. and Crothers, D.M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6808-6811.
25. Reif, F. (1965) *Fundamentals of Statistical and Thermal Physics* McGraw-Hill, New York) Chap. 6.
26. LeBret, M. (1978) *Biopolymers* 17, 1939-1955.
27. Thomas, J.C., Allison, S.A., Appellof, C.J. and Schurr, J.M. (1980) *Biophys. Chem.* 12, 177-188.
28. Barkley, M.D. and Zimm, B.H. (1979) *J. Chem. Phys.* 70, 2991-3007.
29. Millar, D.P., Robbins, R.J. and Zewail, A.H. (1981) *J. Chem. Phys.* 74, 4200-4201.
30. Depew, R.E. and Wang, J.C. (1975) *Proc. Natl. Acad. Sci. USA* 72, 4275-4279.
31. Pulleyblank, D.E. *et al.* (1975) *Proc. Natl. Acad. Sci. USA* 72, 4280-4284.
32. Vologodskii, A.V. *et al.* (1979) *Nature* 280, 294-298.
33. Strogatz, S. (1982) *J. Chem. Phys.* 77, 580-581.
34. Hogan, M., LeGrange, J. and Austin, B. (1983) *Nature* 304, 752-754.
35. Robinson, B.H. *et al.* (1980) *J. Mol. Biol.* 139, 19-44.
36. Shore, D. and Baldwin, R.L. (1983) *J. Mol. Biol.* 170, 957-981.
37. Putnam, B.F., Prohofsky, E.W. and Van Zandt, L.L. (1982) *Biopolymers* 21, 885-894.
38. Cohen, G. and Eisenberg, H. (1966) *Biopolymers* 4, 429-440.
39. Mei, W.N., Kohli, M., Prohofsky, E.W. and Van Zandt, L.L. (1981) *Biopolymers* 20, 833-852.
40. Kollman, P., Keepers, J.W. and Weiner, P. (1982) *Biopolymers* 21, 2345-2376.
41. Keepers, J.W. and James, T.L. (1982) *J. Amer. Chem. Soc.* 104, 929-939.
42. Modrich, P. (1982) *CRC Crit. Revs. Biochem.* 13, 287-323.
43. Wells, R.D., Klein, R.D. and Singleton, C.K. (1980) in *The Enzymes*, Third Ed., Vol. 14, Part A (P.D. Boyer, Ed., Academic Press, New York), 157-191.