Trefoil Knotting Revealed by Molecular Dynamics Simulations of Supercoiled DNA

Tamar Schlick and Wilma K. Olson

Computer simulations of the supercoiling of DNA, largely limited to stochastic search techniques, can offer important information to complement analytical models and experimental data. Through association of an energy function, minimum-energy supercoiled conformations, fluctuations about these states, and interconversions among forms may be sought. In theory, the observation of such large-scale conformational changes is possible, but modeling and numerical considerations limit the picture obtained in practice. A new computational approach is reported that combines an idealized elastic energy model, a compact B-spline representation of circular duplex DNA, and deterministic minimization and molecular dynamics algorithms. A trefoil knotting result, made possible by a large time-step dynamics scheme, is described. The simulated strand passage supports and details a supercoiled-directed knotting mechanism. This process may be associated with collective bending and twisting motions involved in supercoiling propagation and interwound branching. The results also demonstrate the potential effectiveness of the Langevin/implicit-Euler dynamics scheme for studying biomolecular folding and reactions over biologically interesting time scales.

Although the hereditary information is encoded in the primary sequence of the genetic material, much of the regulatory information of DNA is hidden in its topology and geometry in the cell. Through supercoiling—the higher organizational form of DNA involving nonplanar bending and twisting of the global double helix itself—central processes such as replication and transcription, which involve helical winding, strand separation, and movement along the DNA, are intricately regulated. The steric and topological problems associated with supercoiling can be studied with the tools of biochemistry, enzymology, differential geometry, topology, and mechanics. These techniques are effective for analyzing the energetics and dynamical aspects of DNA supercoiling, the interactions of DNA with other macromolecules, and the interconversions among topological isomers (*1*, *2*). Reactions involving knotting and catenation, in particular, are associated with management of important conformational problems for the compactly organized cellular DNA. Such problems include the unlinking of replication and recombination products and the untangling of knots. The topological changes induced by such strand break and passage activities alter the DNA geometry and thereby regulate biological function (*3*–*10*).

Duplex strand passages are typically mediated by type II topoisomerases (*3*–*6*). Although molecular aspects remain largely unknown, information on these reactions can be obtained through analysis of product families of DNA knots and catenanes. Slithering, tracking, and random-collision models have been proposed to explain synaptic and strand passage recombination events that lead to characteristic knot and catenane products (*7*, *8*). Structural information on DNA that is experimentally limited or ambiguous can also be deduced because product topology and complexity reflect directly the substrate DNA geometry as well as reaction intermediates (*7*, *8*). Although such mechanistic information is accumulating, many details are missing in these key regulatory processes. Most reactions are studied in vitro, and little information is available on the structure of topoisomerase enzymes themselves (*4*). However, the intracellular state of DNA is a crucial factor in directing strand passage events (*8*, *9*). The "supercoiled-directed knotting" model (*8*) suggests that supercoiling provides the right geometry and energy to permit strand passage. Products are affected by the degree of DNA supercoiling as well as the local structural details. In fact, if the DNA is sufficiently supercoiled, duplex strand passage may not require ATP (adenosine-5′-triphosphate) in the presence of excess T4 DNA topoisomerase (*6*). The large topoisomerase population may contribute by increasing protein binding and the degree of DNA condensation, thereby enhancing knotting probability.

Our recent simulation study of supercoiled DNA (*11*) combines an idealized elastic energy model (*12*), a compact B-spline representation of closed circular-duplex DNA curves (*13*), and two new minimization (*14*) and molecular dynamics (MD) algorithms (*15*, *16*). The energy E includes components for bending (E_B), twisting (E_T), contour length (E_L), and attractive as well as repulsive van der Waals interactions (E_V) (*17*). It is formulated in terms of the DNA curve $\mathbf{r}(u) \in \mathcal{R}^3$ (see Fig. 1 caption) so that the topological invariant Lk (linking number), describing the number of times the DNA strands intertwine around one another, is the sum of the two geometric quantities, Tw (twist) and Wr (writhing number) (*18*). The critical variable ΔLk, representing the imposed linking number difference as measured from a relaxed circular state Lk_0 (*11*), is included by assuming an isotropic rod model (*19*). The term E_V is important for eliminating close contacts that might allow unrealistically high values of intrachain nodes to compensate for topological strain. The attractive part, in particular, may stabilize the association of two DNA fragments and promote the formation of supercoiled or supercoiled-like structures, such as revealed by the novel cross-packing of duplexes in several recent crystal structures (*20*). An attractive potential between segments along the DNA may also reflect an excess topoisomerase population that effectively decreases the overall polymer dimension and increases knotting frequency (*21*). This attraction may account for the observed knotting of supercoiled DNA by T4 DNA topoisomerase in the absence of ATP (*6*).

Until now, most simulations of supercoiled DNA have been limited to simple curve representations (for example, polygonal line) and stochastic Metropolis-biased Monte Carlo techniques (*22*, *23*) because of computational complexity (both model size and elastic energy formulation in terms of the independent Cartesian variables). Our spline construction represents complex smooth curves in terms of a small number of control points $\{\mathbf{x}_i\}$, ensures chain closure and second-order continuity (important for minimization) and, through the curve's local dependency on the $\{\mathbf{x}_i\}$, reduces considerably the number of nonzero derivative elements (*11*). The Langevin/implicit-Euler scheme for MD permits the use of large time steps (Δt) typically prohibited in macromolecular simulations (*24*). Numerical stability over larger time steps is achieved through an implicit discretization formulation but requires solution of a minimization subproblem at every step (*25*). A frequency-discriminating damping is introduced through the scheme's two damping forces: frictional (physical) and intrinsic (numerical). By choosing a "cutoff frequency" ω_c and setting the frictional damping constant $\gamma = (\omega_c)^2\ \Delta t$, we establish a regime where modes with vibrational frequencies $\omega >> \omega_c$ are effectively frozen, while modes $\omega << \omega_c$ are fully activated (*15*, *16*).

The combination of compact modeling

T. Schlick, Chemistry Department and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012.
W. K. Olson, Department of Chemistry, Rutgers, State University of New Jersey, Wright-Rieman Laboratories, New Brunswick, NJ 08903.

and large time-step dynamics captures here a trefoil knotting of supercoiled DNA. Although "accidental" strand passages may occur in stochastic search techniques because of a large perturbation at a given step, our focus here is identifying thermodynamically driven events dependent on the supercoiled DNA topology and geometry. With dynamics at small time steps, van der Waals repulsion would prohibit strand crossing. For the present combination of ribbon-like curve and elastic model, we use Δt = 100 femtoseconds (fs) and $\omega_c = 10^{10}$ s^{-1} [$\sim 10^{-3}(k_B T/\hbar)$ around room temperature T, where k_B is Boltzmann's constant and $\hbar$ is Planck's constant divided by 2π]. For a 1000-bp model with a bending-to-torsional rigidity constant $\rho = A/C = 1$ (lowest in the physically feasible range) and $|\Delta Lk| = 6$, the closed circular state is highly unstable (*11*).

Selected frames capturing the knotting dynamics, started from a closed circular DNA duplex, are shown in Fig. 1. Figures 2 and 3 sketch the knotting mechanism and illustrate the variation of E_T, E_B, E_V, Wr, and the principal moments of the radius of gyration, $\{I_1, I_2, I_3\}$, as a function of iteration. The resulting knot is a trefoil (easily recognized from the isotopic structures in Fig. 4), the simplest and most abundant knot in many strand passage reactions (*8*). The knot sign is positive (see Fig. 2 caption) as ΔLk is chosen positive. By symmetry, the resulting interwound handedness is simply reversed for negative ΔLk, and an analogous knotting mechanism produces a (−) trefoil product. Both mirror images are found in vivo, though the (−) trefoil knot

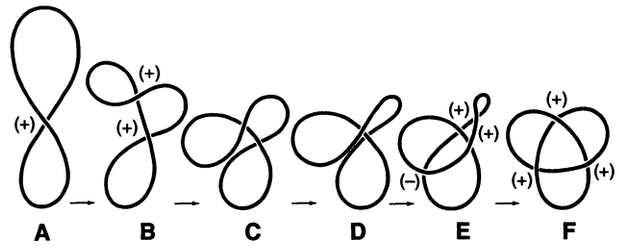

Fig. 2. Sketch of inferred trefoil knotting mechanism. Supercoiling propagation and branching may lead to close contacts between two segments of the duplex DNA chain and to a DNA geometry that may favor knotting: (**A**) A figure-8–like structure results from a first superhelical fold in response to linking number strain. (**B**) A second superhelical fold begins to form and divides the chain into three regions. (**C**) Collective bending and twisting induce closer intrachain contacts as one folded region moves closer to the other. (**D**) Twisting of the middle region follows. (**E**) Strain results from two nearby nodes of opposite signs. (Node sign is determined for each crossing by assigning an arbitrary orientation along the DNA axis and counting a crossing negative or positive, depending on the direction of rotation required to align an arrow of the top segment with one of the bottom segment: negative for clockwise and positive for counterclockwise rotations of less than 180°.) (**F**) A direct strand passage, through introduction of a transient break and subsequent rejoining of the top broken segment on the opposite side, inverts a (−) node to a (+) node to form the (+) trefoil knot. A similar mechanism can be applied to a negatively supercoiled DNA substrate to produce a (−) trefoil knot.

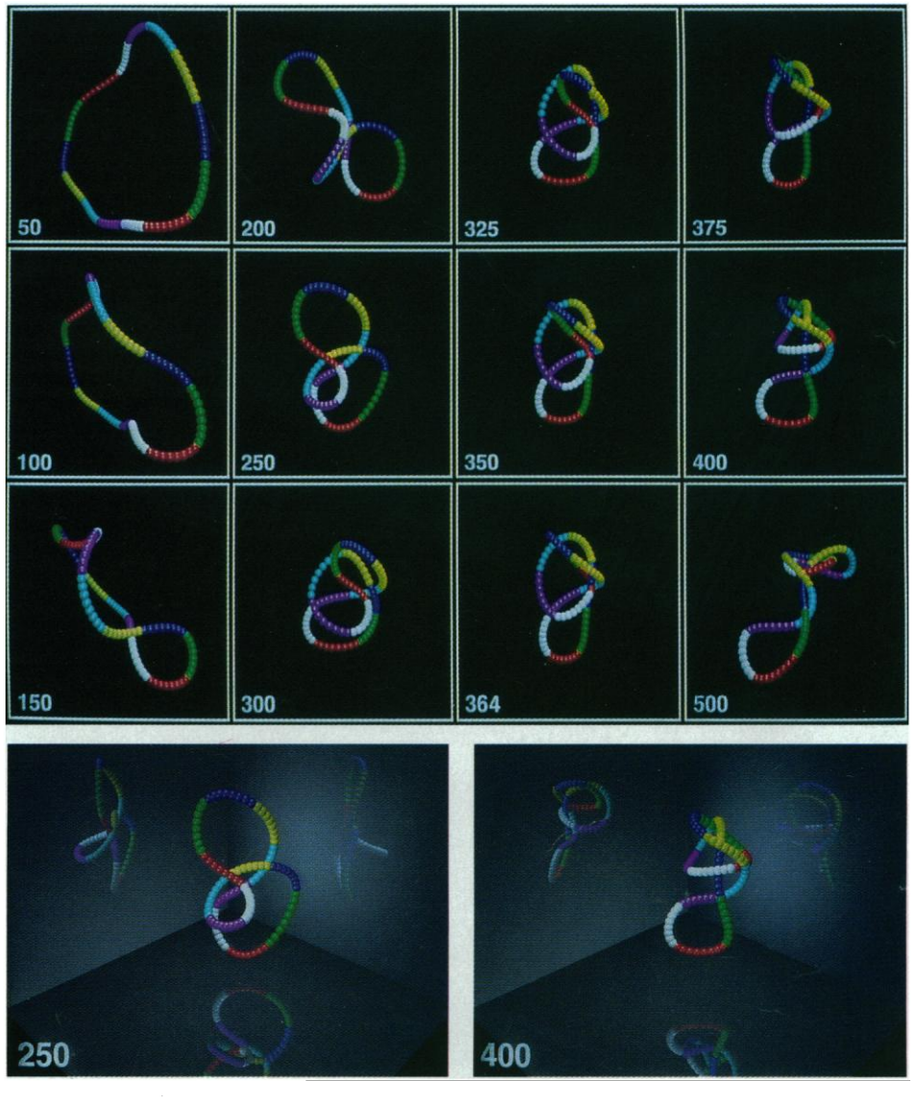

Fig. 1. (**Top**) Selected frames from a molecular dynamics simulation capturing the trefoil knotting reaction. (**Bottom**) Expanded views at frames 250 and 400. The parameters documented in (*17, 19*) were used with a simulation temperature of T = 400 K and a gradual modification of ΔLk, as described in the text, to mimic the topological change during strand passage. Different colors correspond to the B-spline segments in our ribbonlike duplex curve $\mathbf{r}(u)$, represented compactly by piecewise parametric cubic splines. The curve $\mathbf{r}(u) = \{\mathbf{r}_1(u), \mathbf{r}_2(u), \ldots, \mathbf{r}_N(u)\}$, $0 \le u \le 1$, is constructed by N controlling points $\mathbf{x}_i \in \mathcal{R}^3$. These points are used to define each curve segment locally by the linear combination: $\mathbf{r}_i(u) = \Sigma_{j=1}^{4} \mathbf{x}_{j+i-2} F_j(u)$, where indices are defined modulo N. The four basis functions are: $F_1(u) = (1/6)(1-u)^3$; $F_3(u) = (1/6)[1 + 3u(1 + u - u^2)]$; $F_2(u) = F_3(1-u)$; and $F_4(u) = F_1(1-u)$. These cubic polynomials ensure parametric continuity of position, slope, and curvature at the junction points. The spheres represent points on the curve at regular intervals of u. Frame numbers correspond to iteration numbers. Each time step corresponds to 100 fs, but a direct correlation between the simulated and physical times is difficult: the model is macroscopic, and unit masses are assigned to the independent variables $\{\mathbf{x}_i\}$. This assignment allows sufficient flexibility of the curve and succeeds at capturing global deformations. Thus a time frame three orders of magnitude greater than nanoseconds is suggested (*11*).

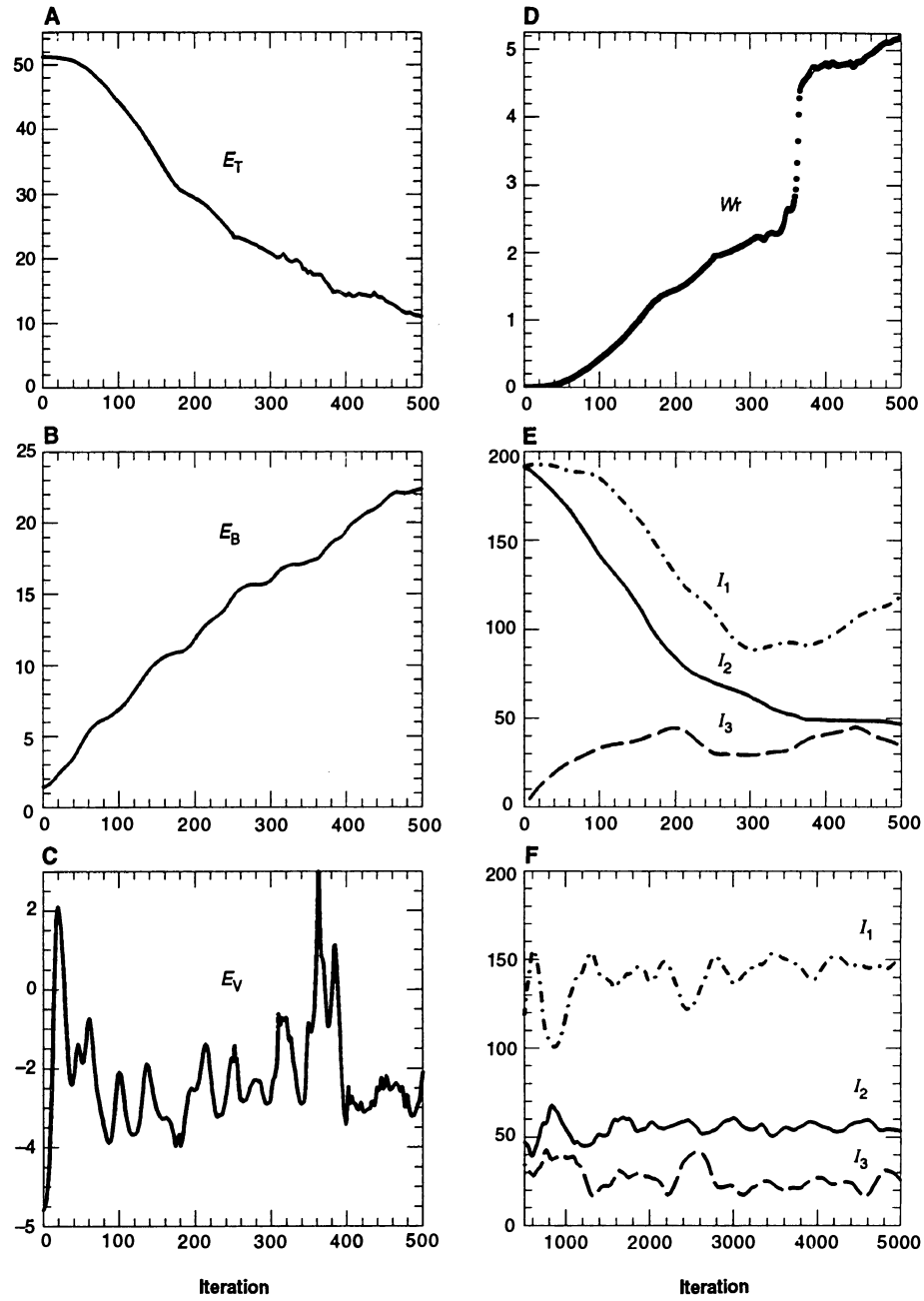

Fig. 3. Energetic and geometric variations during the knotting simulation. Energy components in kilocalories per mole are given for: (**A**) twisting (E_T); (**B**) bending (E_B); and (**C**) van der Waals (E_V) terms. Also given are (**D**) the writhing number (Wr) and (**E** and **F**) the principal moments of the radius of gyration (I_1, I_2, I_3) in angstroms, all as a function of the MD iteration. The moments characterize the overall shape of the B-spline DNA curve. At each time step, they are computed from the parametric curve vector $\mathbf{r}(u)$ as follows. The tensor **R**, containing the elements of the average second moments of the displacement vectors with respect to the center of gravity, is computed through summations of outer products: $\mathbf{R} = (1/L)\,\Sigma_{i,u}\{(\mathbf{r}_i(u) - \langle\mathbf{r}(u)\rangle)(\mathbf{r}_i(u) - \langle\mathbf{r}(u)\rangle)^T\}du$, where the center of gravity vector $\langle\mathbf{r}(u)\rangle = (1/L)\,\Sigma_{i,u}\mathbf{r}_i(u)du$. The quantity L denotes the total contour length of the chain, and the summation symbols $\{i,u\}$ extend over all piecewise segments of the curve **r** and all mesh points considered in the unit interval over u. The nine elements of the matrix **R** describe the best fitting ellipsoid to the B-spline curve. To determine the principal axes of this ellipsoid, we perform the similarity transformation $\mathbf{Q}^{-1}\mathbf{RQ} = \mathbf{D}$, where **Q** is the diagonalizing eigenvector matrix relating the principal axis system to the original coordinate frame of the chain. The eigenvalues d_i of **D** (i = 1, 2, and 3) describe rough ellipsoidal bound areas for the closed chain. The lengths of the principal axes of the ellipsoid are given as the root-mean-square displacement lengths $l_i = \sqrt{d_i}$. If the chain is folded in an extended interwound trajectory, $I_1 >> I_2, I_3$. If the chain is folded in an extended interwound trajectory, $I_1 >> I_2, I_3$. If the chain is circular, $I_1 = I_2$ and $I_3 = 0$.

is more often a recombination product of negatively supercoiled DNA.

Figure 1 reveals how kinks leading to superhelical folds appear quickly in the circular curve (frames 50 and 100) and lead into a figure-8–like configuration as E_B increases beyond a critical value (*11*). A triply looped structure emerges at frame 200 as a second superhelical fold (frame 150) is propagated. The contacts generated by such looped forms have been suggested as a key feature of a trefoil knotting mechanism (*10*). It is possible that the supercoiling dynamics leading to such a form may be associated with branching in interwound supercoils (*23, 26*). Closer contacts within the chain produce another pronounced superhelical fold, accompanied by twisting of the middle region (frames 250 and 300 and Fig. 2). The overall more compact shape can be noted in Fig. 3E: the root-mean-square displacement lengths, I_1, I_2, and I_3, of the ellipsoidal axes that best fit the overall DNA shape are very close. Further twisting and bending gradually lead to critical contacts (after frame 300) involving two very close figure-8–like folded regions strained by two nearby nodes of opposite signs (Fig. 2). In fact, Liu *et al.* have suggested a trefoil knotting mechanism in which two helices separated by a twist, at the top and bottom of a folded figure-8 structure, pass through each other (*6*). A protein wrapping could be imagined to direct or stabilize this folding geometry.

Significantly, we find that this preknotting state corresponds to a minimum in E_V (Fig. 3C). In addition to the crossover site, several regions of close contact along the DNA are noted. Moreover, the orientation is crossed between the chains before passage. Unfavorable nonbonded interactions between duplex segments are thereby limited, and the segments surrounding the "crossing" chain may help force the knotting. Such crossed orientations of duplex DNA segments may be important to any interaction involving closely packed chains, as in supercoiled DNA. A close relation between crossed-oriented helices and nodes of superhelical DNA has been suggested from unusual crystal lattices of DNA (*20*). Here, if the strands are properly aligned, a mechanism involving a concerted sequence of side-by-side recombination steps, rather than strand passage, might also produce the same knotted product.

The strand "passage" can be observed in frames 300 to 400. From a critical value around three, Wr increases rapidly during five time intervals in steps 360 to 365, resulting in an overall change of two (Fig. 3D). We mimic the topological change by modifying ΔLk by ΔWr when $\Delta Wr > 0.04$ between two consecutive MD steps. A total ΔLk change of two is absorbed over 15 MD iterations. Accompanying a steady increase in E_B is thus a gradual decrease in E_T (Fig. 3, A and B) and conservation of the sum $Tw + Wr$ at every

step. During passage, concerted bending and twisting motions can be observed for a top portion of the supercoiled DNA (Figs. 1 and 2). A "sign inversion" mechanism (5) passes one duplex segment through the other and inverts the handedness at one node (Fig. 2). Although at no time step does E_V reach a relatively high value, large fluctuations and attainment of a maximum are evident throughout passage. Subsequently, the knotted geometry is apparently quite stable and E_V fluctuates much less. A gradual motion separates the structure into a "knotted" and "interwound" region, resulting in an overall elongated interwound shape (end of Fig. 3, E and F, where notably $I_1 >> I_2 \approx I_3$).

Once knotting occurs, smaller overall variations in geometry and energy are evident. Time steps 500 to 5000 (Fig. 3F) reveal collective bending motions of the loops in the knotted region that are accompanied by winding and unwinding of the interwound region. A doubly wrapped-loop structure "W" soon emerges (Fig. 4A) that is stabilized by a low bending energy and favorable van der Waals contacts. This form suggests a favorable geometry for protein contact. Indeed, Wang proposed a similar coiled geometry for a DNA ring wrapped around gyrase (10). Cozzarelli has further suggested that supercoiling winding and unwinding may direct topoisomerase binding and unbinding (5). Here, collective bending and twisting gradually open the two loops of W to opposite directions and direct its interconversion to an open trefoil-knotted geometry "O" (Fig. 4B). A moderate barrier in E_V is involved, and thus protein intervention may be a crucial factor. Both W and O were found to be energy minima, with energies of 18.0 and 15.8 kcal/mol, respectively; W has higher elastic and overall potential energy than O, although its van der Waals component is lower (see Fig. 4 caption).

Interestingly, we identified similar knotted supercoiled DNA forms as energy minima in an MD simulation with ΔLk fixed at

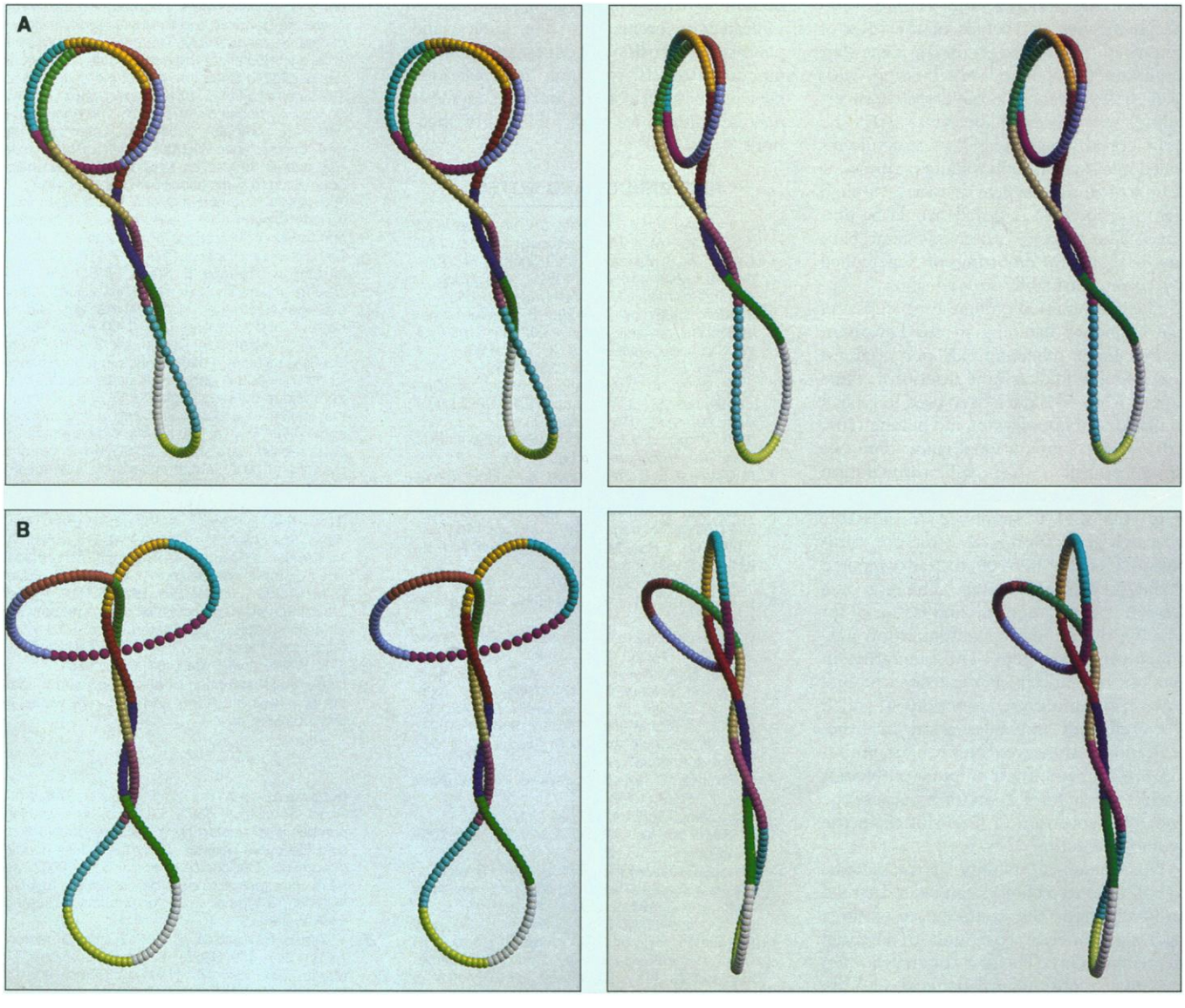

Fig. 4. Energy-minimized trefoil-knotted DNA structures. Through collective bending of the loops in the "knotted" region of the chain (frame 500 of Fig. 1) and winding and unwinding of the "interwound" end, (**A**) a "wrapped" or double-looped structure W forms. The structure W is an energy minimum, with Wr = 5.52, E_B = 16.12 kcal/mol, E_T = 8.52 kcal/mol, and E_V = −6.62 kcal/mol. A combination of collective bending and twisting produces (**B**) an "open" trefoil knotted geometry O, which is also an energy minimum and has Wr = 6.12, E_B = 16.14 kcal/mol, E_T = 4.84 kcal/mol, and E_V = −5.17 kcal/mol. Two stereo pair views are shown for each structure.

6 (8.4 and 9.1 kcal/mol for W and O, respectively). Lower values of Wr were achieved through fewer superhelical turns in the interwound region, and a smaller difference in Wr in this pair (0.1) in relation to the pair of Fig. 4 (0.6) produces the overall energy difference. In both cases, the crossed-oriented strands in the preknotting state appear to stimulate the knotting strongly.

Overall, these energetic and dynamical details strengthen the importance of the DNA supercoiling geometry and energy in directing strand passage events. They suggest that collective bending and twisting involved in supercoiling propagation and interwound branching may produce naturally looped helical regions that accommodate topoisomerase binding and activity. Although mechanistic details of the roles of proteins in strand passages remain somewhat mysterious (*3*), noncovalent binding, with strength dependence on the degree of supercoiling, was suggested for gyrase (*5*). The triply looped supercoiled DNA structure revealed here may arise in folding pathways at physiological superhelical densities; through sign inversion, DNA trefoils would be produced. The suggested thermodynamic path may account for knotting of supercoiled DNA from torsionally stressed states.

The combination of large time steps and high-frequency damping in our Langevin/implicit-Euler scheme holds potential for biomolecular folding and reactions. Time steps of 10 to 40 fs have been used in typical simulations of liquid water and butane (*16*), with overall computational gain. Here, we gain a factor of 10 over explicit integration (*11*, *16*). The minimization subproblem at every step (*25*) is simplified considerably because a good initial value for the minimum is available from quantities computed at the previous time step. The use of a Newton minimization method is crucial for reducing the number of minimization iterations per time step. The similarity in composition and Hessian sparsity structure between the minimized functions Φ and E allows efficient implementation of structure-tailored truncated Newton minimization (*14*). The multiple-minima problem is also less severe for Φ than for E because the eigenvalue spectrum of Φ is shifted in the favorable direction (*25*).

The numerical stability of the scheme over all time-step choices can be understood intuitively from the composition of Φ: a quadratic "kinetic" term and a potential energy term (*25*). The first dominates when Δt is small (continuous dynamics), and the second dominates when Δt is large, so low-energy forms in conformation space are essentially followed. However, even if we assume that the minimization subproblem can always be solved, how realistic would the physical results be with very large time moves and damping in conformation space? Clearly, vibrational modes in biomolecules are intricately coupled, and certain small, high-frequency fluctuations may facilitate collective motions. It may nonetheless be possible to focus on specific low-frequency and relatively slow conformational changes by choosing appropriate time steps and cutoff frequencies and then to resolve the motion of interest on finer time and spatial scales. Particularly intriguing examples of such concerted conformational transitions involve hinge-bending proteins such as triosephosphate isomerase and lactoferrin (*27*). Rigid-body–like motions, observed in helical pivoting and lobe projections over active sites, have recently received heightened attention because of their wide-range applications to interdomain interactions. The kinetic and potential composition of our target dynamics function suggests a natural "interpolation" between molecular statics and dynamics that may be suitable for such folding investigations.

REFERENCES AND NOTES

1. J. Vinograd and J. Lebowitz, *J. Gen. Physiol.* **49**, 103 (1965); M. Le Bret, *Biopolymers* **18**, 1709 (1979); W. R. Bauer, F. H. C. Crick, J. H. White, *Sci. Am.* **243**, 118 (July 1980); J. C. Wang, in *Cyclic Polymers*, J. A. Semlyen, Ed. (Elsevier, London, 1986), pp. 225–260; S. A. Wasserman and N. R. Cozzarelli, *Science* **232**, 951 (1986); D. W. Sumners, *Math. Intell.* **3**, 71 (1990); C. J. Benham, *Phys. Rev. A* **39**, 2852 (1989).
2. J. H. White, in *Mathematical Methods for DNA Sequences*, M. S. Waterman, Ed. (CRC Press, Boca Raton, FL, 1989), chap. 9.
3. J. C. Wang, *J. Biol. Chem.* **266**, 6659 (1991); *Annu. Rev. Biochem.* **54**, 665 (1985).
4. N. Osheroff, E. L. Zechiedrich, K. C. Gale, *BioEssays* **13**, 269 (1991); D. B. Wigley, G. J. Davies, E. J. Dodson, A. Maxwell, G. Dodson, *Nature* **351**, 624 (1991); S. Krueger *et al.*, *J. Mol. Biol.* **211**, 211 (1990); D. C. Rau, M. Gellert, F. Thoma, A. Maxwell, *ibid.* **193**, 555 (1987).
5. N. R. Cozzarelli, *Science* **207**, 953 (1980).
6. L. F. Liu, C.-C. Liu, B. M. Alberts, *Cell* **19**, 697 (1980).
7. H. W. Benjamin and N. R. Cozzarelli, *Proc. Robert A. Welch Found. Conf. Chem. Res.* **29**, 107 (1986); S. A. Wasserman, J. M. Dungan, N. R. Cozzarelli, *Science* **229**, 171 (1985); S. J. Spengler, A. Stasiak, N. R. Cozzarelli, *Cell* **42**, 325 (1985); M. Adrian *et al.*, *EMBO J.* **9**, 4551 (1990).
8. S. A. Wasserman and N. R. Cozzarelli, *J. Biol. Chem.* **266**, 20567 (1991).
9. K. Shishido, S. Ishii, N. Komiyama, *Nucleic Acids Res.* **17**, 9749 (1989).
10. J. C. Wang, *Sci. Am.* **247**, 94 (July 1982).
11. T. Schlick and W. K. Olson, *J. Mol. Biol.* **223**, 1089 (1992).
12. Although elasticity theory [A. E. Love, *Treatise on the Mathematical Theory of Elasticity* (Dover, New York, ed. 4, 1944), chaps. 18 and 19; L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, Oxford, 1975), chap. 2] clearly provides an idealized representation of DNA structure and flexibility, the elastic approximation is valid as long as the allowed local bending and twisting deformations are limited to small angular fluctuations, which are consistent with the average end-to-end extension and twisting dynamics observed in naturally occurring chains of mixed sequence.
13. M.-H. Hao and W. K. Olson, *Biopolymers* **28**, 873 (1989); R. H. Bartels, J. C. Beatty, B. A. Barsky, *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling* (Morgan Kaufmann, Los Altos, CA, 1987).
14. T. Schlick and M. Overton, *J. Comput. Chem.* **8**, 1025 (1987); T. Schlick and A. Fogelson, *ACM Trans. Math. Software* **18**, 46 (1992).
15. C. S. Peskin and T. Schlick, *Commun. Pure Appl. Math.* **42**, 1001 (1989); T. Schlick and C. S. Peskin, *ibid.*, p. 1141.
16. T. Schlick, S. Figueroa, M. Mezei, *J. Chem. Phys.* **94**, 2118 (1991); A. Nyberg and T. Schlick, *ibid.* **95**, 4986 (1991); *Chem. Phys. Lett.*, in press.
17. The energy is given by $E_B + E_T + E_L + E_V$:

$$E = \frac{A}{2}\oint \kappa^2(s)ds + \frac{C}{2}\oint (\omega - \omega_0)^2\, ds + K(L - L_0)^2 + \sum_{i,j}\left[\frac{-a}{d_{ij}^6} + \frac{b}{d_{ij}^{12}}\right] \quad (1)$$

where s denotes the arc length, κ the curvature, ω the twist density, and L_0 the contour length of the DNA chain; a and b are attractive and repulsive nonbonded parameters, respectively, and A and C are the bending and torsional-rigidity stiffness constants, respectively. The contour length term E_L is a computational device, and its contribution has negligible value throughout the calculations. The summation in E_V extends over pairs of curve points $i < j$ whose distance is d_{ij}. The quantities L and κ are computed according to standard integral formulas involving the vector of the curve, $\mathbf{r}(u)$, and its derivatives, $\mathbf{r}'(u)$ and $\mathbf{r}''(u)$, by analogous discrete summations over curve points (*11*). The elastic parameters A and C are taken from various experimental measurements of bending and torsional fluctuations [P. J. Hagerman, *Annu. Rev. Biophys. Biophys. Chem.* **17**, 265 (1988); G. Muzard, B. Thévény, B. Révet, *EMBO J.* **9**, 1289 (1990)]. Inherently, the measured values reflect average sequence and ionic-medium effects. Here we use the values $L_0 = 3400$ Å, $\rho = A/C = 1$, $K = 0.1$ kcal/(mol Å^2), and $A = 2 \times 10^{-19}$ erg cm. This value of A corresponds to a persistence length $\ell \approx 500$ Å and a root-mean-square (rms) per residue bending angle $\langle\theta_B^2\rangle^{1/2} \approx 5°$ at 300 K; the choice of ρ corresponds to a rms twisting angle $\langle\theta_t^2\rangle^{1/2} = 7°$ at 300 K (*11*). We parameterize E_V to reproduce a shallow energy minimum at a distance of 30 Å so that acceptable interactions occur at distances greater than 25 Å, the estimated helix diameter of a hydrated B-DNA in solution [D. C. Rau, B. Lee, V. A. Parsegian, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2621 (1984); F. Livolant, A. M. Levelut, J. Doucet, J. P. Benoit, *Nature* **339**, 724 (1989); M. Mandelkern, J. G. Elias, D. Eden, D. M. Crothers, *J. Mol. Biol.* **152**, 153 (1981)]. For a minimum of −0.05 kcal/mol, the corresponding attraction-repulsion parameters are $a = 7.3 \times 10^7$ (kcal/mol)Å^6 and $b = 2.7 \times 10^{16}$ (kcal/mol)Å^{12}.
18. J. H. White, *Am. J. Math.* **91**, 693 (1969).
19. Under the assumption of an isotopic rod model, the local twist is uniform, and E_T can be written in terms of ΔTw, as

$$E_T = \frac{2\pi^2 C}{L_0}(\Delta Lk - Wr)^2 \quad (2)$$

because $\Delta Lk = \Delta Tw + Wr$ [F. B. Fuller, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 815 (1971)]. The writhing number is evaluated by a discrete double sum over the curve points (*11*) by the Gauss double integral (*2*). Our choice of $\Delta Lk = 6$ for 1000 base pairs corresponds to superhelical density $\Delta Lk/Lk_0 = 0.063$, a typical value for naturally occurring DNA systems.
20. Y. Timsit, E. Westhof, R. P. P. Fuchs, D. Moras, *Nature* **341**, 459 (1989); Y. Timsit, E. Vilbois, D. Moras, *ibid.* **354**, 167 (1991); Y. Timsit and D. Moras, *J. Mol. Biol.* **221**, 919 (1991).
21. M. D. Frank-Kamenetskii and A. V. Vologodskii, *Sov. Phys. Usp.* **24**, 679 (1981).
22. W. K. Olson and P. Zhang, in *Methods in Enzymology, Molecular Design and Modeling: Concepts and Applications*, J. J. Langone, Ed. (Academic Press, Orlando, FL, 1991), vol. 203, pp.

403–432; R. K. Z. Tan and S. C. Harvey, in *Theoretical Biochemistry and Molecular Biophysics*, D. L. Beveridge and R. Lavery, Eds. (Adenine, Guilderland, NY, 1991), vol. 1, pp. 125–137.
23. K. V. Klenin, A. V. Vologodskii, V. V. Anshelevich, A. M. Dykhne, M. D. Frank-Kamenetskii, *J. Mol. Biol.* **217**, 413 (1991).
24. The time-step restriction in typical explicit MD schemes is severely constrained by the most rapid vibrational modes [J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge Univ. Press, Cambridge, 1987)]. This generally limits Δt to the femtosecond range and the trajectory to the picosecond range. Two basic approaches to this problem have emerged. In the first, the fast vibrational modes are effectively frozen by a constrained MD formulation [W. F. van Gunsteren, *Mol. Phys.* **40**, 1015 (1980)]; although this technique increases Δt by a small factor (2 or 5), each time step still requires more work. In the second, the motion is separated into multiple time scales for the slow and fast force components [for example, M. E. Tuckerman, B. J. Berne, A. Rossi, *J. Chem. Phys.* **94**, 1465 (1991); O. Teleman and B. Jönsson, *J. Comput. Chem.* **7**, 58 (1986)]; special systems where such a separation is natural (such as a liquid containing heavy and light particles) are most suitable though this technique is now more generally applied.
25. The resulting discretization of the Langevin equation by the implicit-Euler scheme [G. Dahlquist and Å. Björck, *Numerical Methods* (Prentice-Hall, Englewood Cliffs, NJ, 1974)] produces the following pair of differential equations:

$$M[(\mathbf{v}^{n+1} - \mathbf{v}^n)/\Delta t] = -\mathbf{g}_E(\mathbf{x}^{n+1}) - \gamma M\mathbf{v}^{n+1} + \mathbf{r}^{n+1} \quad (3a)$$

$$(\mathbf{x}^{n+1} - \mathbf{x}^n)/\Delta t = \mathbf{v}^{n+1} \quad (3b)$$

where the random force r is a stationary Gaussian process with mean and covariance matrix given by

$$\langle \mathbf{r}^n \rangle = 0, \langle \mathbf{r}^n (\mathbf{r}^m)^T \rangle = 2\gamma k_B T M(\delta_{nm}/\Delta t) \quad (3c)$$

The superscripts n represent vector values at time $(n\Delta t)$, x and v are the collective position and velocity vectors, respectively; M is the diagonal mass matrix; γ the damping constant; $\mathbf{g}_E$ the gradient vector of the potential energy E, and δ the Kronecker delta function. To calculate $\mathbf{x}^{n+1}$ from $\mathbf{x}^n$ and $\mathbf{v}^n$, we formulate a minimization subproblem (*15, 16*) for the "dynamics function" $\Phi(\mathbf{x})$, where

$$\Phi(\mathbf{x}) = \frac{1}{2}(1 + \gamma\Delta t)(\mathbf{x} - \mathbf{x}_0^n)^T M(\mathbf{x} - \mathbf{x}_0^n) + (\Delta t)^2 E(\mathbf{x}) \quad (4a)$$

$$\mathbf{x}_0^n = \mathbf{x}^n + [\Delta t/(1 + \gamma\Delta t)][\mathbf{v}^n + \Delta t M^{-1}\mathbf{r}^{n+1}] \quad (4b)$$

This minimization is efficiently performed with the truncated Newton variant for potential energy functions (*11, 14*). After a minimum $\mathbf{x}^{n+1}$ for Φ is calculated, the new velocity vector, $\mathbf{v}^{n+1}$, is calculated directly by Eq. 3b.
26. M. Le Bret, *Biopolymers* **23**, 1835 (1984).
27. D. Joseph, G. A. Petsko, M. Karplus, *Science* **249**, 1425 (1990); H. R. Faber and B. W. Matthews, *Nature* **348**, 263 (1990); B. F. Anderson, H. M. Baker, G. E. Norris, S. V. Rumball, E. N. Baker, *ibid.* **344**, 784 (1990).
28. We thank H. M. Berman, S. Broyde, S. Lifson, M. L. Overton, J. K. Percus, and C. S. Peskin for their inspiration and contributions; C. J. Benham, D. L. Beveridge, N. R. Cozzarelli, S. C. Harvey, J. Hermans, B. M. Pettitt, D. W. Sumners, N. Seeman, I. Tobias, and the reviewers for their valuable comments; and E. Friedman, J. Greenberg, K. Fallon, A. De Falco, and J. Westbrook for their dedicated molecular graphics work. Generous support was provided by the National Science Foundation (CHE-9002146 and Presidential Young Investigator Award ASC-9157582), U.S. Public Health Service (GM 34809), Searle Scholars Program/The Chicago Community Trust, Faculty of Arts and Sciences at New York University, AAUW Educational Foundation, New York State Science and Technology Foundation, and Whitaker Foundation. Computational support was provided by the San Diego Supercomputer Center, the Academic Computing Facility at New York University, and the Rutgers University Center for Computational Chemistry.

17 March 1992; accepted 5 June 1992

A Mutation in the POU-Homeodomain of Pit-1 Responsible for Combined Pituitary Hormone Deficiency

Sally Radovick,* Michelle Nations, Yuefen Du, LaVonne A. Berg, Bruce D. Weintraub, Fredric E. Wondisford†‡

Pit-1 is a pituitary-specific transcription factor responsible for pituitary development and hormone expression in mammals. Mutations in the gene encoding Pit-1 have been found in two dwarf mouse strains displaying hypoplasia of growth hormone, prolactin, and thyroid-stimulating, hormone-secreting cell types in the anterior pituitary. A point mutation in this gene was identified on one allele in a patient with combined pituitary hormone deficiency. Mutant Pit-1 binds DNA normally but acts as a dominant inhibitor of Pit-1 action in the pituitary.

Pit-1–growth hormone factor–1 (Pit-1) is a member of the POU family of transcription factors that regulate mammalian development (*1*). Pit-1 contains two protein domains, termed POU-specific and POU-homeo, which are both necessary for high-affinity DNA binding on the genes encoding the growth hormone (GH) and prolactin (Prl) (*2, 3*). Pit-1 is also important for regulation of the genes encoding Prl and the thyrotropin β subunit (TSH-β) by thyrotropin-releasing hormone (TRH) and cyclic adenosine 3′,5′-monophosphate (*4–7*). Because mutations in the gene encoding Pit-1 (*Pit-1*) have been found in dwarf mice strains (*8*) that lack GH, Prl, and TSH-secreting cell types, we wanted to determine if mutations within this gene resulted in a similar deficiency of pituitary hormones in humans.

The patient (W.T.R.) studied in this report was previously documented to have a deficiency of GH, Prl, and TSH, which was manifest as severe mental retardation and short stature (*9*). Both GH and TSH were undetectable in the plasma before or after provocative stimuli; Prl in the plasma was low (3 μg per liter) and did not respond to stimulation by chlorpromazine or TRH. Baseline gonadotropin amounts, however, were normal [luteinizing hormone (LH), 15 IU per liter, and follicle-stimulating hormone, 7 IU per liter] and responded appropriately to stimulation by gonadotropin-releasing hormone. Serum cortisol levels were normal (8 a.m., 378 nmol per liter; and 4 p.m., 190 nmol per liter); steroid precursors increased appropriately after metyrapone. The patient's mother (E.R.) is of normal stature and has normal pituitary hormone levels. Unfortunately, the remaining family members were inaccessible for study; but, according to historical reports, they all have normal stature, which suggests that this patient may represent a sporadic case of combined pituitary hormone deficiency (CPHD).

The polymerase chain reaction (PCR) was used to amplify genomic DNA fragments from human *Pit-1* (*10*). The POU-specific and POU-homeodomains were amplified separately by means of specific oligonucleotides (*11*). The Pit-1 domains amplified in this study correspond to exons 4 through 6 of the mouse *Pit-1* gene (*8*) and contain most of the POU-specific and the entire POU-homeo domains. As a control in these experiments, genomic DNA from three normal patients was also included. We found no abnormalities in the POU-specific domain after sequencing at least seven independent clones from two separate PCR amplifications from either the patient or normal controls. A C to T mutation of codon 271, however, was found in approximately one-half of independent clones we sequenced from the patient (three of seven cloned fragments); the

S. Radovick and Y. Du, Division of Endocrinology, Department of Pediatrics, Rainbow Babies and Childrens Hospital and Case Western Reserve University Medical School, Cleveland, OH 44106.
M. Nations, L. A. Berg, F. E. Wondisford, Division of Endocrinology and Hypertension, Department of Medicine, University Hospitals of Cleveland and Case Western Reserve University Medical School, Cleveland, OH 44106.
B. D. Weintraub, Molecular and Cellular Endocrinology Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892.

*Present address: Division of Endocrinology, Department of Medicine, The Children's Hospital, 300 Longwood Avenue, Boston, MA 02115.
†Present address: Division of Endocrinology, Department of Medicine, Beth Israel Hospital, 330 Brookline Avenue, Boston, MA 02215.
‡To whom correspondence should be addressed.