1609 (1985); R. S. Brown, C. Sander, P. Argos, *FEBS Lett.* **186**, 271 (1985).
4. A. M. Ginsberg, B. O. King, R. G. Roeder, *Cell* **39**, 479 (1984); J. Y. T'so, K. J. van den Berg, L. J. Korn, *Nucleic Acids Res.* **14**, 2187 (1986).
5. G. P. Diakun, L. Fairall, A. Klug, *Nature* **324**, 698 (1986).
6. D. R. Smith, I. J. Jackson, D. D. Brown, *Cell* **37**, 645 (1984).
7. M. S. Sands and D. F. Bogenhagen, *Nucleic Acids Res.* **19**, 1791 (1992).
8. J. Hayes, T. D. Tullius, A. P. Wolffe, *J. Biol. Chem.* **264**, 6009 (1989).
9. T. Pieler, J. Hamm, R. G. Roeder, *Cell* **48**, 91 (1987).
10. X. Liao, K. R. Clemens, L. Tennant, P. E. Wright, J. M. Gottesfeld, *J. Mol. Biol.* **223**, 857 (1992); K. R. Clemens, X. Liao, V. Wolf, P. E. Wright, J. M. Gottesfeld, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10822 (1992); J. Hayes and K. R. Clemens, *Biochemistry* **31**, 11600 (1992).
11. S. Sakonju and D. D. Brown, *Cell* **31**, 395 (1982).
12. J. J. Hayes and T. D. Tullius, *J. Mol. Biol.* **227**, 407 (1992).
13. J. Anderson and N. Delihas, *J. Biol. Chem.* **261**, 2912 (1986); F. Baudin, P. Romby, P. J. Romaniuk, B. Ehresmann, C. Ehresmann, *Nucleic Acids Res.* **17**, 10035 (1989); J. Christiansen, R. S. Brown, B. S. Sproat, R. A. Garrett, *EMBO J.* **6**, 453 (1987); T. Pieler, V. A. Erdmann, B. Appel, *Nucleic Acids Res.* **12**, 8393 (1984); P. J. Romaniuk, *Biochemistry* **28**, 1388 (1989).
14. M. S. Sands and D. F. Bogenhagen, *Mol. Cell. Biol.* **7**, 3985 (1987); Q. You, N. Veldhoen, F. Baudin, P. J. Romaniuk, *Biochemistry* **30**, 2495 (1991).
15. K. D. MacFerrin, M. P. Terranova, S. L. Schreiber, G. L. Verdine, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2304 (1990). The sequences of primers for the native $NH_2$-terminus and terminating primers for the COOH-terminal deletions are given in Liao *et al.* and in Clemens *et al.* [in (*10*)]. Initiating oligonucleotides for the $NH_2$-terminal deletions contained an Nde I restriction site of which the last three bases form the start codon (underlined) and were: (zf2-7) 5′-CGGCCAAGCTTAGGAGGAATATCATATGGAGAAACCATTTCCATGTAAGGAAG; (zf3-7) 5′-CGGCCAAGCTTAGGAGGAATATCATATGGAGAAACCATTTCCATGTAAGGAAG; (zf4-7) 5′-GTTGCTGAAGGACATATGTATGTGTGCCATTTTG; (zf5-7) 5′-GTT GCTGAAGGACATATGTACGAATGTCCTC. After amplification, polymerase chain reaction products were cleaved with Nde I and Eco RI and were ligated into similarly digested plasmid pRK172 [M. McLeod, M. Stein, D. Beach, *EMBO J.* **6**, 729 (1987)] and amplified in *Escherichia coli* strain DH5α and transferred to strain BL21(DE3) for protein expression [F. W. Studier and B. Moffat, *J. Mol. Biol.* **189**, 113 (1986)]. The sequence of each of the truncated polypeptides was determined from the expression clones. The zf1-9 had a single base change, G → T, which causes the mutation of $Ala^{257}$ to Ser. In the original cloning of zf1-7 [Clemens *et al.*, in (*10*)] single base mutations that affect the amino acid sequence of the protein were found in fingers 4 and 6. These mutations were converted to the wild-type sequence by oligonucleotide-directed mutagenesis [T. A. Kunkel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488 (1985)], and the sequence of zf1-7 was confirmed by DNA sequence analysis. Similar methods were used to introduce additional mutations into finger 6. Mutagenesis was performed in the vector pTZ19U (Bio-Rad) and the gene products were expressed from this vector.
16. A synthetic somatic-type 5*S* RNA gene cloned into pUC18 was the generous gift of P. J. Romaniuk [P. J. Romaniuk, I. Leal de Stevenson, H.-H. A. Wong, *Nucleic Acids. Res.* **15**, 2737 (1987)]. Plasmid DNA was cleaved with Dra I (Stratagene) and used as a template for transcription with T7 RNA polymerase (Stratagene) with [α-$^{32}$P]GTP (guanosine triphosphate) and unlabeled nucleoside triphosphates. The run-off transcript was 121 nucleotides long. RNA was purified by phenol extraction and concentrated by ethanol precipitation before use in binding reactions.
17. Note that any impurities in the recombinant polypeptide preparations would likely be similar for each protein because they were all isolated from the same bacterial strain by the same protocol.
18. Protein concentrations were determined by Coomassie blue staining of SDS-polyacrylamide gels and densitometry with known amounts of bovine serum albumin as a standard. After purification on a single heparin-Sepharose column (*10*), the purity of the recombinant polypeptides ranged from 25% (zfs 2-7 and 3-7) to near homogeneity. Concentrations of some of the polypeptides were determined by Western (immuno) blotting with a polyclonal serum raised against *Xenopus* TFIIIA with known amounts of the purest polypeptides (zfs 1-5 and 1-6) as standards. Protein concentrations used in $K_d$ determinations were adjusted appropriately.
19. M. S. Lee, G. P. Gippert, K. V. Soman, D. A. Case, P. E. Wright, *Science* **245**, 635 (1989); N. P. Pavletich and C. O. Pabo, *ibid.* **252**, 809 (1991).
20. K. E. Joho, M. K. Darby, E. T. Crawford, D. D. Brown, *Cell* **61**, 293 (1990).
21. O. Theunissen, F. Rudt, U. Guddat, H. Mentzel, T. Pieler, *ibid.* **71**, 679 (1992).
22. M. K. Darby and K. E. Joho, *Mol. Cell. Biol.* **12**, 3155 (1992).
23. T. J. Gibson, J. P. Postma, R. S. Brown, P. Argos, *Protein Eng.* **2**, 209 (1988).
24. K. R. Clemens, V. Wolf, P. Zhang, S. J. McBryant, J. M. Gottesfeld, P. E. Wright, in preparation.
25. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
26. E. Westhof *et al.*, *J. Mol. Biol.* **207**, 417 (1989).
27. Supported by grants from the National Institute of General Medical Sciences to J.M.G. (GM26453 and GM47530) and to P.E.W. (GM36643). X.L. was supported by the Cancer Research Institute/ Miriam and Benedict Wolf Fellowship and K.R.C. received support from a National Institutes of Health postdoctoral fellowship award (F32 CA09023). We thank B. Blair for her assistance in the preparation of this manuscript.

9 November 1992; accepted 8 February 1993

# Knotting of a DNA Chain During Ring Closure

Stanley Y. Shaw and James C. Wang

**The formation of knotted species on random ring closure of two DNAs that are 5.6 kilobase pairs (kbp) and 8.6 kbp in length was measured, and these data were used to calculate the effective DNA helix diameter as a function of sodium ion and magnesium ion concentration. In the presence of more than 50 mM magnesium ion, interactions between DNA segments appear to be attractive rather than repulsive. The free energy of formation of relaxed trefoil and figure-eight DNA knots and of supercoiled trefoil DNA knots was also evaluated.**

Descriptions of knots in 19th century atomic physics (*1*) presaged the formal development of knot theory, the study of the topology of deformable but nonbreakable closed curves in three-dimensional space. Concepts from knot theory have permeated such fields as algebraic topology, physics, and synthetic and polymer chemistry, sometimes revealing relations among seemingly disparate disciplines (*2*). In biology, the abundance of large ring-shaped molecules in living organisms has provided fertile ground for the study of knots, particularly since the discovery of knotted DNA rings (*3*). A number of DNA knots have been characterized in detail, and their knot-types have provided insights into the mechanisms of the enzymatic reactions that produced them (*4*).

When a particular DNA transaction can yield either a knotted or unknotted product, the energetics of DNA knotting might influence the course of the reaction (*5*). In contrast to the elegant and extensive work on knot typing and its mechanistic implications (*4*), however, no experimental datum exists for the probability (and hence the Gibbs free energy) of formation of even the simplest knotted DNA ring, the trefoil. It has been shown that for a polygon of $n$ sides, the probability that it is unknotted approaches zero as $n$ approaches infinity (*6*); thus, the probability of knotting in very large DNA rings is likely to be high. In solution, DNA behaves as a stiff molecule with a Kuhn statistical segment length (*7*) of about 1000 Å or 300 base pairs (bp). Theoretical models of knot formation in polymer chains suggest that for DNA molecules thousands of base pairs in length, the probability of knotting during ring closure is small but significant (*8–11*). Thus, in addition to yielding information on the energetics of knotted DNA rings, measurements of knot formation in randomly cyclized DNA should also provide experimental tests of the various theoretical treatments of knotting probabilities.

To determine experimentally the probability of knotting, we used two DNA molecules 5.6 kbp and 8.6 kbp in length. The DNAs were designed to have single-stranded ends identical to those of phage λ (*12*); because the chemical step of pairing such ends is the rate-limiting step in cyclization (*13*), the fraction of knotted rings in the product represents that in the equilibrium population.

Harvard University, Department of Biochemistry and Molecular Biology, 7 Divinity Avenue, Cambridge, MA 02138.

Agarose gel electrophoresis of cyclized DNA showed a band readily identifiable as a nicked unknotted circle, as well as a series of bands of decreasing intensity that migrated ahead of the unknotted nicked ring (Fig. 1, A and B). To confirm that these were DNA knots, we mixed the radiolabeled cyclization product with unlabeled nicked knots generated by eukaryotic DNA topoisomerase II (*14*). Superimposing the autoradiogram of a dehydrated agarose gel onto the ethidium-stained rehydrated gel showed that the bands observed on cyclization comigrate with the ladder of nicked knots generated by topoisomerase II (Fig. 1A).

With increasing NaCl concentration, the knotting probability $p$ gradually increased, reaching a maximum at 2 M NaCl (Fig. 2A). The knots observed were predominantly trefoils, knots with three irreducible nodes. The 8.6-kbp DNA displayed the same overall variation with NaCl concentration as did the 5.6-kbp substrate but produced a higher proportion of knots (Fig. 2A); trefoils, as well as figure-eight knots (with four irreducible nodes), were detected.

These results support the Metropolis–Monte Carlo calculations of Klenin *et al.* (*9*). They showed that the probability of knotting is sensitive to the effective DNA helix diameter $d_e$, which corresponds to the distance of closest approach between the helical axes of two nonadjacent DNA segments modeled as rigid cylinders to account for intersegmental repulsion. The theoretical results of Klenin *et al.* were combined with data shown in Fig. 2A to yield $d_e$ as a function of NaCl concentration (Fig. 3) (*15*); values of $d_e$ from two independent calculations of DNA as a polyelectrolyte (*16, 17*) are also included in the figure.

Comparable knotting probabilities were observed at $MgCl_2$ concentrations that were some 10- to 50-fold lower than the corresponding NaCl concentration, and the maximum value of $p$ attained was higher in the presence of 0.1 M $MgCl_2$ than in 2 M NaCl (Fig. 2B). The complexity of the knots observed at high $Mg^{2+}$ concentrations also increased; with the 8.6-kbp substrate, up to four knotted species were distinguished.

The theoretical work of Klenin *et al.* (*9*) predicted knotting probabilities for values of $d_e$ down to zero. In the presence of $Mg^{2+}$, our experimental knotting probabilities reached these "zero-diameter" limits (5.6-kbp substrate, 47 mM $MgCl_2$; 8.6-kbp substrate, 39 mM $MgCl_2$). At higher concentrations of $Mg^{2+}$, the observed knotting probabilities reached levels that would correspond to negative values of $d_e$.

A physical interpretation of these data is that $Mg^{2+}$ not only shields the negatively charged DNA backbone more effectively than do monovalent cations, but at concentrations higher than several hundredths molar may also induce an attractive potential between DNA segments. Lattice polygon calculations support the intuitive view that in poor solvents, which favor more collapsed polygon configurations and closer approaches of polymer segments, both the probability of knotting and the complexity of knots increase (*10*). Furthermore, studies using crystallography and electron microscopy indicate that $Mg^{2+}$ can facilitate the close approach of two DNA segments, sometimes at a DNA crossover (*18*); other work has confirmed that cross-linking two DNA segments affects knotting probability and complexity as $Mg^{2+}$ does (*14, 19*). An alternative interpretation, that $Mg^{2+}$ reduces the Kuhn statistical segment or persistence length of DNA (*20*) and thus favors knot formation, is less attractive in view of the insensitivity of the parameter to the ionic environment (*7*).

The free energy of converting a nicked unknotted DNA ring to a nicked knot can be obtained directly from the data on knot formation during ring closure; values calculated for the nicked trefoil and figure-eight knot are tabulated in Table 1. The free energy of forming a supercoiled knot from a

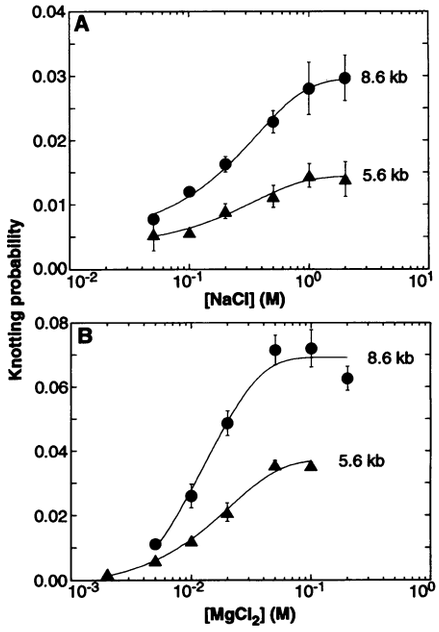


**Fig. 2.** (**A**) Probability of knot formation, $p$, in the presence of NaCl for 5.6-kbp and 8.6-kbp DNAs. The probability $p$ is defined as the fraction of cyclized molecules that are knotted. (**B**) Probability of knot formation in the presence of $MgCl_2$ for the same DNAs shown in (A). Error bars indicate the standard deviation calculated from multiple measurements at each point.

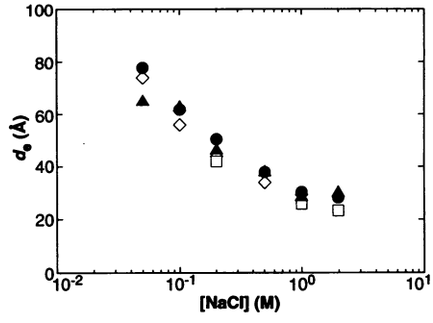


**Fig. 3.** Effective DNA helix diameter, $d_e$, as a function of NaCl concentration. For each concentration tested, we used the calculated curves of knotting probability versus helix diameter of Klenin *et al.* (*9*) to determine the corresponding value of $d_e$. The curve of Klenin *et al.* corresponding to a DNA chain of 20 Kuhn statistical segments, or ~5.9 kbp, was used to determine values of $d_e$ for the 5.6-kbp substrate (▲); similarly, the curve for a DNA chain of 30 statistical segments (or approximately 8.8 kbp) was used for the 8.6-kbp substrate data (●). Also plotted are values of $d_e$ obtained from Table 1 of (*16*) (◇) and Table 1 of (*17*) (□).

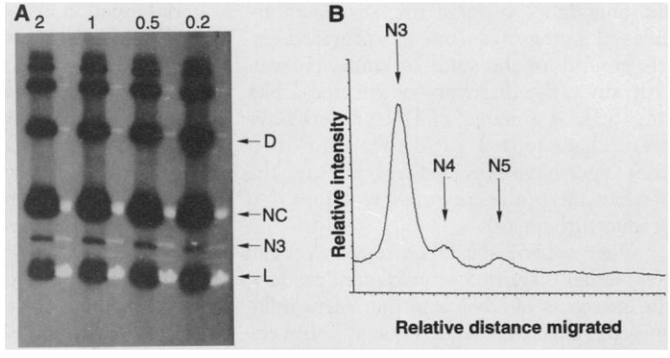


**Fig. 1.** Formation of DNA knots during random cyclization. (**A**) Simultaneous viewing of products of random cyclization (autoradiogram) and a ladder of nicked knots generated by *S. cerevisiae* DNA topoisomerase II (ethidium fluorescence). A radiolabeled 5.6-kbp DNA (*12*) was cyclized in the presence of 0.2, 0.5, 1, and 2 M NaCl, and unlabeled linear DNA and *S. cerevisiae* topoisomerase II–knotted DNA were added to each sample before electrophoresis (*28*). The cyclization conditions favored forms with cohered ends, and these forms were stable during electrophoresis (*13*); the radiolabeled linear molecules in the cyclized samples were due to the presence of linear DNA without the proper 12-nucleotide overhangs. L, linear monomer; N3, nicked trefoil; NC, unknotted nicked circle; and D, linear dimer. (**B**) Densitometric tracing of knots resulting from random cyclization of an 8.6-kbp substrate (*12*) in the presence of 0.1 M $MgCl_2$, showing the formation of nicked DNA knots with three, four, and five nodes.

nicked unknotted DNA ring is then the sum of the free energy of knotting a nicked ring plus the free energy of converting a nicked knot to a supercoiled knot. We have determined the latter quantity for a trefoil.

Covalent closure of a nicked DNA ring by DNA ligase yields topological isomers (topoisomers) of the DNA ring that differ only in their linking numbers (*21*). For unknotted DNA, the distribution of topoisomers is Gaussian, corresponding to a quadratic dependence of the free energy of supercoiling on the linking difference: $\Delta G = K_\tau (Lk - Lk°)^2$, where $\Delta G$ is the Gibbs free energy, $K_\tau$ is a length-dependent constant, and $Lk$ and $Lk°$ are the linking numbers of a given topoisomer and of the thermodynamically most stable topoisomer, respectively (*21*). For DNA rings larger than 2000 bp, the quantity $NK_\tau$, where $N$ is the length of the DNA in base pairs, has been shown to assume a constant value of 1100 $RT$ (*21*, *22*), where $R$ is the gas constant and $T$ is the temperature. The thermal distribution of the linking number topoisomers of a 5.6-kbp DNA trefoil, obtained by nicking and religating the knot, is shown in Fig. 4; the fit to a Gaussian function with $NK_\tau$ = 1760 $RT$ is excellent (*23*). For the unknotted circle, $NK_\tau$ was 1310 $RT$, in good agreement with the average value of 1100 $RT$ for large DNAs (*21*, *22*).

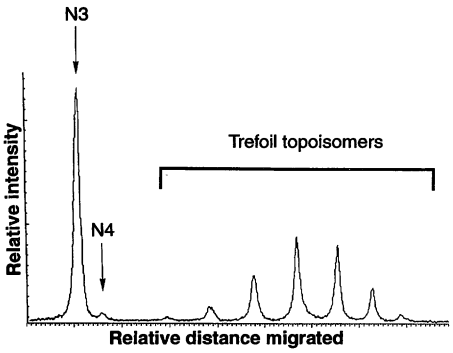


**Fig. 4.** A thermal population of trefoils of different linking numbers, produced by ligation of a nicked trefoil. N3, nicked pCC56cos (*12*) trefoil; N4, nicked figure-eight knot of pCC56cos, which contaminated the nicked trefoil during its isolation by gel electrophoresis. N3 was obtained by treatment of negatively supercoiled pCC56cos with phage T4 DNA topoisomerase, followed by nicking of the product (*29*); the nicked trefoil so prepared contained predominantly molecules with three negative nodes (*25*).

The free energy cost of introducing supercoils into a trefoil is significantly greater than that for unknotted DNA; constraints due to the presence of the three irreducible trefoil nodes presumably make it more difficult to change the linking number from its relaxed value. For a relaxed DNA ring 5 to 10 kbp in size, our data show that the trefoil is disfavored by 2 to 4 kcal mol$^{-1}$ in free energy depending on the reaction conditions. For a supercoiled DNA, the trefoil may be even more disfavored, due to its higher value of $NK_\tau$. The values of $NK_\tau$ for the unknotted and trefoil 5.6-kbp ring would predict that $\Delta G°$ is higher by 5, 20, or 40 kcal mol$^{-1}$, respectively, for the trefoil with $\Delta Lk$ = −10, −20, or −30 (corresponding to a specific linking difference, or superhelical density, of −0.019, −0.038, or −0.056). It is likely, however, that the value of $NK_\tau$ obtained from the trefoil topoisomer distribution shown in Fig. 4 is more representative of the free energy cost of accommodating a trefoil knot when $Lk$ begins to deviate from $Lk°$; further changes in $Lk$ may be accompanied by lower free energy increments, and predictions based on a constant $NK_\tau$ value may overestimate the free energy required to introduce higher degrees of supercoiling into a trefoil. Direct measurements at larger $\Delta Lk$ values are needed to ascertain the free energy differences between unknotted and knotted DNAs that are more supercoiled than the topoisomers in a thermal population. Nevertheless, it is clear from the present results that the presence of a knot in a relaxed or supercoiled DNA ring is associated with a substantial free energy cost; in enzyme-catalyzed reactions that yield knotted products, this free energy cost must be compensated by a favorable free energy term, such as that derived from protein-DNA interactions (*5*, *24*).

**Table 1.** Free energies of formation of a nicked knot from an unknotted, nicked circle. All measurements were performed at 34°C. ND, not detectable.

| DNA length (kbp) | NaCl | | | $MgCl_2$ | | |
|---|---|---|---|---|---|---|
| | Concentration (M) | ΔG°, trefoil (kcal mol$^{-1}$) | ΔG°, figure-8 (kcal mol$^{-1}$) | Concentration (M) | ΔG°, trefoil (kcal mol$^{-1}$) | ΔG°, figure-8 (kcal mol$^{-1}$) |
| 5.6 | 0.05 | 3.17 | ND | 0.002 | 3.80 | ND |
| | 0.10 | 3.14 | ND | 0.005 | 3.10 | ND |
| | 0.20 | 2.87 | ND | 0.010 | 2.68 | ND |
| | 0.50 | 2.73 | ND | 0.020 | 2.37 | 4.22 |
| | 1.00 | 2.57 | ND | 0.050 | 2.10 | 3.19 |
| | 2.00 | 2.60 | ND | 0.100 | 2.07 | 3.51 |
| 8.6 | 0.05 | 2.95 | ND | 0.005 | 2.73 | ND |
| | 0.10 | 2.69 | ND | 0.010 | 2.31 | 3.50 |
| | 0.20 | 2.54 | 4.11 | 0.020 | 1.95 | 3.08 |
| | 0.50 | 2.33 | 4.03 | 0.050 | 1.73 | 2.80 |
| | 1.00 | 2.19 | 4.04 | 0.100 | 1.71 | 2.82 |
| | 2.00 | 2.15 | 4.14 | 0.200 | 1.78 | 2.92 |

## REFERENCES AND NOTES


1. W. Thompson, *Philos. Mag.* **34**, 15 (1867); P. G. Tait, in *Scientific Papers* (Cambridge Univ. Press, London, 1898), vol. 1, p. 273.
2. V. F. R. Jones, *Bull. Am. Math. Soc.* **12**, 103 (1985); L. H. Kauffman, *Knots and Physics* (World Scientific, Teaneck, NJ, 1991); F. Y. Wu, *Rev. Mod. Phys.* **64**, 1099 (1992); E. Wasserman, *J. Am. Chem. Soc.* **82**, 4433 (1960); H. L. Frisch and E. Wasserman, *ibid.* **83**, 3789 (1961); G. Schill, *Catenanes, Rotaxanes and Knots* (Academic Press, New York, 1971); D. M. Walba, *Tetrahedron* **41**, 3161 (1985); C. O. Dietrich-Buchecker and J.-P. Sauvage, *Angew. Chem. Int. Ed. Engl.* **28**, 189 (1989); C. O. Dietrich-Buchecker *et al.*, *ibid.* **29**, 556 (1990); J. E. Mueller *et al.*, *J. Am. Chem. Soc.* **113**, 6306 (1991); S. F. Edwards, *Proc. Phys. Soc.* **91**, 513 (1967); P.-G. de Gennes, *Macromolecules* **18**, 2445 (1984).
3. L. F. Liu *et al.*, *J. Mol. Biol.* **106**, 439 (1976); L. F. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5498 (1981).
4. S. A. Wasserman, J. M. Dungan, N. R. Cozzarelli, *Science* **229**, 171 (1985); S. A. Wasserman and N. R. Cozzarelli, *ibid.* **232**, 951 (1986).
5. M. Gellert and H. Nash, *Nature* **325**, 401 (1987); J. C. Wang and G. Giaever, *Science* **240**, 300 (1988).
6. D. W. Sumners and S. G. Whittington, *J. Phys. A. Math. Gen.* **21**, 1689 (1988); N. Pippenger, *Disc. Appl. Math.* **25**, 273 (1989).
7. P. J. Hagerman, *Biopolymers* **20**, 1503 (1981); Z. Kam, N. Borochov, H. Eisenberg, *ibid.*, p. 2671.
8. A. V. Vologodskii *et al.*, *Sov. Phys. JETP* **39**, 1059 (1974); M. D. Frank-Kamenetskii, A. V. Lukashin, A. V. Vologodskii, *Nature* **258**, 398 (1975); M. D. Frank-Kamenetskii and A. V. Vologodskii, *Usp. Fiz. Nauk* **134**, 641 (1981).
9. K. V. Klenin, A. V. Vologodskii, V. V. Anshelevich, A. M. Dykhne, M. D. Frank-Kamenetskii, *J. Biomol. Struct. Dyn.* **5**, 1173 (1988).
10. E. J. Janse van Rensburg and S. G. Whittington, *J. Phys. A Math. Gen.* **23**, 3573 (1990).
11. J. P. J. Michels and F. W. Wiegel, *Proc. R. Soc. London Ser. A* **403**, 269 (1986); T. Schlick and W. K. Olson, *J. Mol. Biol.* **223**, 1089 (1992); *Science* **257**, 1110 (1992).
12. DNA oligonucleotides were synthesized such that two DNA duplexes would form during annealing. The sequence of duplex L (5′-P-CGGGTCCTTTC-CGGTGATCCGACAGGTTACG-3′:5′-HO-AG-GTCGCCGCCCCGTAACCTGTCGGATCACCG-GAAAGGACCCGCATG-3′) has a 4-nucleotide 3′ overhang identical to that produced by the restriction endonuclease Sph I, and a 12-nucleotide 5′ overhang identical to the left end of bacteriophage λ; duplex R (5′-HO-GGGCGGC-GACCTCGCGGGTTTTCGCTATTTATG-AAAATTTTCG-3′:5′-P-TCGACGAAAATTTTCAT-AAATAGCGAAAACCCGCG-3′) has the right end of phage λ at one end and a 5′ overhang Sal I end at the other. These two duplexes were ligated onto the ends of linear DNA fragments with Sph I and Sal I or Xho I ends to produce substrates for random cyclization via the phage λ cohesive ends. The plasmid pCC56 (5.6 kbp) was constructed by insertion of a 2.7-kbp Kpn I–Eco RI fragment from the λ genome into the corresponding sites in pGEM3 (Promega). To prepare the cyclization substrate, we cut pCC56 at the Sal

I and Sph I sites in the pGEM3 polylinker and ligated it to a 5- to 7.5-fold molar excess of each of duplexes L and R. Excess oligomers were removed from the 5.6-kbp DNA by precipitation in the presence of polyethylene glycol. The plasmid form of the 5.6-kbp cyclization substrate, incorporating duplexes L and R, was designated pCC56cos. The plasmid pCC86 (8.8 kbp) consists of a 5.9-kbp Sal I–Eco RI fragment from the λ genome ligated to a 2.8-kbp Sal I–Eco RI fragment from pGEM3. The 8.6-kbp cyclization substrate was constructed by ligation of duplexes L and R onto an 8.5-kbp Sph I–Xho I fragment of pCC86; subsequent purification of the 8.6-kbp substrate was identical to that of its 5.6-kbp counterpart described above.


13. J. C. Wang and N. Davidson, *J. Mol. Biol.* **15**, 111 (1966).
14. T. Hsieh, *J. Biol. Chem.* **258**, 8413 (1983).
15. A similar dependence of $d_e$ on NaCl concentration obtained from knotting probabilities is described in V. V. Rybenkov, N. R. Cozzarelli, and A. V. Vologodskii (*Proc. Natl. Acad. Sci. U.S.A.*, in press).
16. D. Stigter, *Biopolymers* **16**, 1435 (1977).
17. A. A. Brian, H. L. Frisch, L. S. Lerman, *ibid.* **20**, 1305 (1981).
18. Y. Timsit, E. Westhof, R. P. P. Fuchs, D. Moras, *Nature* **341**, 459 (1989); M. Adrian *et al.*, *EMBO J.* **9**, 4551 (1990).
19. N. K. Annan, P. R. Cook, S. T. Mullins, G. Lowe, *Nucleic Acids Res.* **20**, 983 (1992).
20. D. Porschke, *J. Biomol. Struct. Dyn.* **4**, 373 (1986).
21. R. E. Depew and J. C. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4275 (1975); D. E. Pulleyblank, M. Shure, D. Tang, J. Vinograd, H. Vosberg, *ibid.*, p. 4280.
22. D. S. Horowitz and J. C. Wang, *J. Mol. Biol.* **173**, 75 (1984).
23. The data were analyzed as in (*21*); regression analyses were performed with Kaleidagraph version 2.1.
24. T. J. Pollock and H. A. Nash, *J. Mol. Biol.* **170**, 1 (1983).
25. S. A. Wasserman and N. R. Cozzarelli, *J. Biol. Chem.* **266**, 20567 (1991).
26. L. F. Liu *et al.*, *Cell* **19**, 697 (1980).
27. R. C. Parker, R. M. Watson, J. Vinograd, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 851 (1977); D. Shortle, P. Grisafi, S. J. Benkovic, D. Botstein, *ibid.* **79**, 1588 (1982).
28. Cyclization of radiolabeled DNA was carried out by heating at 70°C for 5 to 7 min to unpair any preexisting cohered single-stranded ends and then annealing for 6 to 8 hours at 34°C. Unlabeled linear and *Saccharomyces cerevisiae* DNA topoisomerase II–knotted (*14*) pCC56cos DNA were added to the labeled samples as carriers before electrophoretic analysis in agarose gel at 4°C, as described in (*14*). Gel slabs were dried and analyzed in a Phosphor-Imager (Molecular Dynamics); the radioactivity of the various bands was quantitated with MD ImageQuant Software version 3.15.
29. T4 DNA topoisomerase was used to generate a series of supercoiled knots from negatively supercoiled pCC56cos as described (*25, 26*); these were then nicked by Hind III in the presence of ethidium bromide (*27*), and the nicked trefoil was gel-purified. The trefoil was radiolabeled by nick translation and ligated by *Escherichia coli* DNA ligase. Electrophoresis was done at room temperature in 0.9% agarose gels in the presence of 90 mM tris base, 90 mM boric acid, and 1.8 mM $Na_2EDTA$.


30. We thank N. R. Cozzarelli, M. D. Frank-Kamenetskii, P. Hagerman, H. Nash, D. Sumners, and B. Zimm for helpful discussions and K. Kreuzer for phage T4 topoisomerase. We are grateful to many of our colleagues, particularly R. Hanai, N. Lue, J. Lindsley, and J. Berger for materials and discussions. We thank also S. Field and M. Greenberg at the Harvard Medical School for use of the Phosphor-Imager. Supported by grants from the National Science Foundation (DMB-88-07067 and 21283F). S.Y.S. is a Howard Hughes Medical Institute predoctoral fellow.

9 July 1992; accepted 26 February 1993


# Negative Regulation of G1 in Mammalian Cells: Inhibition of Cyclin E–Dependent Kinase by TGF-β


Andrew Koff, Masahiko Ohtsuki, Kornelia Polyak, James M. Roberts, Joan Massagué*



**Transforming growth factor–β (TGF-β) is a naturally occurring growth inhibitory polypeptide that arrests the cell cycle in middle to late G1 phase. Cells treated with TGF-β contained normal amounts of cyclin E and cyclin-dependent protein kinase 2 (Cdk2) but failed to stably assemble cyclin E–Cdk2 complexes or accumulate cyclin E–associated kinase activity. Moreover, G1 phase extracts from TGF-β–treated cells did not support activation of endogenous cyclin-dependent protein kinases by exogenous cyclins. These effects of TGF-β, which correlated with the inhibition of retinoblastoma protein phosphorylation, suggest that mammalian G1 cyclin-dependent kinases, like their counterparts in yeast, are targets for negative regulators of the cell cycle.**


Progression through the cell cycle is regulated by cyclin-dependent protein kinases (CDKs) (*1*). The CDKs are necessary for the start of the S phase (*2*) and mitosis (*3*). The mitotic role of the CDKs is thought to be executed by the prototypic CDK, Cdc2 (*3, 4*). Cdc2 is positively regulated by cyclin B (*5*), which is synthesized during the S and G2 phases. The binding of cyclin B to Cdc2 induces phosphorylation of the complex, which leads to its activation at the G2 to M transition (*6, 7*). In the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, this kinase is also necessary for starting the S phase (*8*). In higher eukaryotes, however, the CDKs have evolved into a small gene family (*9*), and at least one additional family member, Cdk2, also functions during the S phase (*10, 11*). Cdk2 kinase activity is first evident during the middle of G1 (*12, 13*), in which it may phosphorylate the retinoblastoma gene product (Rb) (*14–16*). Cdk2, like Cdc2, is regulated by its association with cyclins. It assembles in a complex with cyclin E in the middle of G1 (*13*); this is followed at the start of the S phase by formation of a complex with cyclin A (*11, 12, 17*). Cyclin E–Cdk2 may function in controlling progression through G1 (*18*), and cyclin A–Cdk2 may function in controlling the start of DNA synthesis (*19*).

Transforming growth factor–β inhibits cell proliferation by delaying or arresting progression through the late portion of G1 (*20, 21*). In Mv1Lu lung epithelial cells, TGF-β prevents Rb phosphorylation during G1 (*21*), retaining Rb in a hypophosphorylated state that may suppress progression into the S phase (*15, 16, 22*). Because the activity of cyclin-dependent kinases seems critical for the G1 to S transition (*13, 18, 19*), negative regulators of this transition might operate by affecting the activity of these kinases. We investigated whether a known G1 kinase, the cyclin E–Cdk2 complex, might be inhibited by TGF-β.

Proliferating Mv1Lu cells were arrested by growth to high density in serum-containing medium. We then released these contact-inhibited cells from quiescence by sparsely seeding them in this medium. As reported (*21*), the transit through the middle of G1, 6 hours after the release from contact inhibition, involved Rb phosphorylation and was blocked by TGF-β added at 6 hours (Fig. 1A). Protein immunoblots with cyclin E antibodies and Cdk2 antibodies showed that these proteins were present in contact-inhibited cells (Fig. 1B). As cells progressed through G1, a faster migrating form of Cdk2 appeared (Fig. 1B) that corresponds to the phosphorylated, catalytically active state of this protein (*23*). To follow the kinetics of Cdk2 association with cyclin E, we immunoprecipitated cell extracts with cyclin E antibodies and analyzed the immune complexes by immunoblot using Cdk2 antibodies (*13*) (Fig. 1C). A small amount of slow-migrating Cdk2 form was present in cyclin E immune complexes 3 hours after the start of G1 and thereafter (Fig. 1C, inset). The amount of fast Cdk2 form associated with cyclin E increased 6 hours after the start of G1 and became the predominant cyclin E–associated Cdk2 form 9 hours after the start of G1 and later (Fig. 1C). The increase in cyclin E–associated fast Cdk2 form was correlated closely with an increase in cyclin E–associated histone H1 kinase activity (Fig. 1D) over a basal background level and preceded by several hours the appearance of kinase activity precipitable with Cdc2 antibodies. Thus, the natural progression of Mv1Lu cells from quiescence into the S phase


A. Koff and J. M. Roberts, Department of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98104.
M. Ohtsuki, K. Polyak, J. Massagué, Cell Biology and Genetics Program and Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, NY 10021.

*To whom correspondence should be addressed.