The mouse probes were donated by: R. Shemer (P1-clone covering 100 kb of the *Snrpn* region and the plasmid P3-*Bam*H1), S. Tilghman (plasmid 20267, 3′ to *Igf2*), W. Reik (λup12, a phage representing sequences upstream of *Igf2*), D. Barlow (pTCP and *Igf2r* gene clones LA2 and COS940), D. Ward (Chr6, a random cosmid clone from chromosome 6), R. Axel (olfactory receptor gene region phage clones 126 and 129 and YAC clone Y12), B. C. Holdener (85-M2, a BAC clone covering the deletion in $C^{112\kappa}$ mice), P. Fraser (mouse β-globin gene region plasmids pβ12g and βmajor), N. Benvenisti (15 kb TMP gene plasmid). P. Rotwein (cosIGF-5) and A. Chess (BAC clone 154f05 containing mouse IL-4 sequences, release I, Genome System Inc.). BAC clone 212a06 was randomly picked from the same library.

Replication timing by S-phase fractionation

EBV-transformed lymphoblasts, or an Abelson-transformed pre-B-cell subclone from Spretus/Musculus F_1 mice (donated by A. Chess), were labelled in 75 μM BrdU for 45 min before harvesting, and nuclei were sorted for cell-cycle fractions according to DNA content[9]. BrdU DNA was isolated from each fraction and assayed for specific sequence content by quantitative PCR[8] carried out for 29 (*Igf2*) or 25 (*Igf2r*) cycles (95 °C 40 s, 55 °C 40 s, 72 °C 40 s) using the primer pairs 5′-CTTGGACTTTGAGTCAAATTGG-3′ and 5′-GGTCGTGCCAATTACATTTCA-3′ (for human *Igf2*); and 5′-TGAGCAGTGGGGCACC TAGT-3′ and 5′-CACGCGTTAGAGGATCCGCA-3′ (for mouse *Igf2r*).

PCR products from the EBV DNA S-phase fractions were electrophoresed after cutting with the enzyme *Apa*I, which detects a polymorphism between the uncut (295-bp) and cut (231-bp) alleles in these human cells. Competitor DNA[8] which has an 85-bp deletion covering the *Apa*I site, is included in every reaction mix. Following SYBR green I staining and gel scanning, the relative amount of each allele in the different fractions was normalized to the level of competitor PCR product. A similar analysis was carried out on BrdU fractions from the Abelson pre-B cell line, using the enzyme *Hae*III to distinguish between the uncut paternal Spretus and cut maternal Musculus alleles which yield 90- and 100-bp bands after digestion. In this case, PCR was carried out without added competitor, but in the presence of [α-^{32}P]dCTP, and detection was by autoradiography. Total uncut PCR product was first quantitated on an initial gel. Each individual band was extracted, and the proportion of maternal or paternal alleles was determined from a second gel after cutting out the PCR product.

Received 26 May; accepted 9 September 1999.

1. Razin, A. & Cedar, H. DNA methylation and genomic imprinting. *Cell* **77,** 473–476 (1994).
2. Kitsberg, D. *et al.* Allele-specific replication timing of imprinted gene regions. *Nature* **364,** 459–463 (1993).
3. Knoll, J. H. M., Cheng, S.-D. & Lalande, M. Allele specificity of DNA replication timing in the Angelman/Prader-Willi syndrome imprinted chromosomal region. *Nature Genet.* **6,** 41–46 (1994).
4. Selig, S., Okumura, K., Ward, D. C. & Cedar, H. Delineation of DNA replication time zones by fluorescence *in situ* hybridization. *EMBO J.* **11,** 1217–1225 (1992).
5. Chess, A., Simon, I., Cedar, H. & Axel, R. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78,** 823–834 (1994).
6. LaSalle, J. M. & Lalande, M. Domain organization of allele-specific replication within the GABRB3 gene cluster requires a biparental 15q11-13 contribution. *Nature Genet.* **9,** 386–394 (1995).
7. Greally, J. M. *et al.* The mouse H19 locus mediates a transition between imprinted and non-imprinted DNA replication patterns. *Hum. Mol. Genet.* **7,** 91–95 (1998).
8. Tenzen, T. *et al.* Precise switching of DNA replication timing in the GC content transition area in the human major histocompatibility complex. *Mol. Cell. Biol.* **17,** 4043–4050 (1997).
9. Hansen, R. S., Canfield, T. KI., Lamb, M. M., Gartler, S. M. & Laird, C. D. Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* **73,** 1403–1409 (1993).
10. Kawame, H., Gartler, S. M. & Hansen, R. S. Allele-specific replication timing in imprinted domains: absence of asynchrony at several loci. *Hum. Mol. Genet.* **4,** 2287–2293 (1995).
11. Windham, L. Q. & Jones, P. A. Expression of H19 does not influence the timing of replication of the Igf2/H19 imprinted region. *Dev. Genet.* **20,** 29–35 (1997).
12. Wines, M. E., Tiffany, A. M. & Holdener, B. C. Physical localization of the mouse aryl hydrocarbon receptor nuclear translocator-2 (Arnt2) gene within the c112K deletion. *Genomics* **51,** 223–232 (1998).
13. McCarrey, J. R. in *Cellular and Molecular Biology of the Testis* 58–89 (Oxford Univ. Press, Oxford, 1993).
14. Bix, M. & Locksley, R. M. Independent and epigenetic regulation of the interleukin-4 alleles in $CD4^+$ T Cells. *Science* **281,** 1352–1354 (1998).
15. Riviere, I., Sunshine, M. J. & Littman, D. R. Regulation of IL-4 expression by activation of individual alleles. *Immunity* **9,** 217–228 (1998).
16. Gabriel, J. M. *et al.* A model system to study genomic imprinting of human genes. *Proc. Natl Acad. Sci. USA* **95,** 14857–14862 (1998).
17. Shemer, R. & Razin, A. in *Epigenetic Mechanisms of Gene Regulation* (eds Russo, V. E. A., Martienssen, R. A. & Riggs, A. D.) 215–229 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1996).
18. Shemer, R. *et al.* Dynamic methylation adjustment and counting as part of imprinting mechanisms. *Proc. Natl Acad. Sci. USA* **93,** 6371–6376 (1996).
19. Birger, Y., Shemer, R., Perk, J. & Razin, A. The imprinting box of the mouse Igf2r gene. *Nature* **397,** 84–88 (1999).
20. Szabo, P. & Mann, J. R. Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment and mechanisms of genomic imprinting. *Genes Dev.* **9,** 1857–1868 (1995).
21. Davis, T. L., Trasler, J. M., Moss, S. B., Yang, Bartolomei, M. S. Acquisition of the H19 methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* **58,** 18–28 (1999).
22. Caspary, T., Cleary, M. A., Baker, C. C., Guan, X.-J. & Tilghman, S. M. Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster. *Mol. Cell. Biol.* **18,** 3466–3474 (1998).
23. Sutcliffe, J. S. *et al.* Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nature Genet.* **8,** 52–58 (1994).
24. Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. & Tilghman, S. M. Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* **375,** 34–39 (1995).
25. Gunaratne, P. H., Nakao, M., Ledbetter, D. H., Sutcliffe, J. S. & Chinault, A. C. Tissue-specific and allele-specific replication timing control in the imprinted human Prader-Willi syndrome region. *Genes Dev.* **9,** 808–920 (1995).
26. Hogan, B., Beddington, R., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo. A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1994).
27. McCarrey, J. R., Hsu, K. C., Eddy, E. M., Klevecz, R. R. & Bolen, J. L. Isolation of viable mouse primordial germ cells by antibody-directed flow sorting. *J. Exp. Zool.* **242,** 107–111 (1987).
28. Bellve, A. *et al.* Spermatogenic cells of the prepuberal mouse, isolation and morphological characterization. *J. Cell Biol.* **74,** 68–85 (1977).
29. Harper, J. C. *et al.* Identification of the sex of human preimplantation embryos in two hours using an improved spreading method and fluorescent *in situ* hybridisation using directly labelled probes. *Hum. Reprod.* **9,** 721–724 (1994).
30. Dunnett, C. W. A multiple comparison procedure for comparing several treatments with a control. *J. Am. Stat. Assoc.* **50,** 1096–1121 (1955).

Acknowledgements

We thank T. Jakubowicz and E. Rand for their help in preparing the manuscript and figures, B. C. Holdener for providing the $C^{112\kappa}$ mice and an appropriate probe for detecting this deletion, R. Nicholls for providing hybrid cell lines and W. Reik for providing the parthenogenetic ES cell line. We also thank H. Friedlander-Klar for help with the statistical analysis. This work was supported by grants from the US–Israel Binational Science foundation (H.C. and J.M.) the Israel Academy of sciences (H.C.) and the Israel Cancer Research Fund (H.C.).

Correspondence and requests for material should be addressed to H.C. (e-mail: cedar@md2.huji.ac.il).

A kinetic proofreading mechanism for disentanglement of DNA by topoisomerases

Jie Yan*, Marcelo O. Magnasco† & John F. Marko*

** Department of Physics, The University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607, USA*
† Center for Studies in Physics and Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

Cells must remove all entanglements between their replicated chromosomal DNAs to segregate them during cell division. Entanglement removal is done by ATP-driven enzymes that pass DNA strands through one another, called type II topoisomerases. *In vitro*, some type II topoisomerases can reduce entanglements much more than expected, given the assumption that they pass DNA segments through one another in a random way[1]. These type II topoisomerases (of less than 10 nm in diameter) thus use ATP hydrolysis to sense and remove entanglements spread along flexible DNA strands of up to 3,000 nm long. Here we propose a mechanism for this, based on the higher rate of collisions along entangled DNA strands, relative to collision rates on disentangled DNA strands. We show theoretically that if a type II topoisomerase requires an initial 'activating' collision before a second strand-passing collision, the probability of entanglement may be reduced to experimentally observed levels. This proposed two-collision reaction is similar to 'kinetic proofreading' models of molecular recognition[2,3].

We consider the knotting state of a single circular DNA strand. Given a 'dumb' topoisomerase that merely passes DNA through itself with some fixed probability each time the molecule strikes itself (Fig. 1), the knot state of a circular DNA strand will come to thermodynamical equilibrium. It will sometimes be knotted (a fraction P^{eq}_{knot} of the time), and sometimes be unknotted (a fraction $P^{eq}_{unknot} = 1 - P^{eq}_{knot}$ of the time)[4,5]. The ratio $P^{eq}_{knot}/P^{eq}_{unknot}$ is equal to the ratio of the rate at which unknotted strands are knotted to the rate at which knotted strands are unknotted.

The equilibrium knot probability P^{eq}_{knot} expected theoretically for a 'dumb' topoisomerase[4,5] matches results from DNA knotting

experiments allowed to reach thermal equilibrium[6,7]. For DNA, P^{eq}_{knot} depends on molecule length and ionic conditions[4–7]. Under the conditions of ref. 1, $P^{eq}_{knot} = 0.031$ for 10-kb P4 DNA and $P^{eq}_{knot} = 0.017$ for 7-kb PAB4 DNA. The more complicated problem of interlinkage of two DNA strands has also been studied in this way; the linking probability P^{eq}_{link} depends on plasmid lengths and concentrations. For the conditions in ref. 1, $P^{eq}_{link} = 0.064$ for two 10-kb P4 DNA strands.

Experimental data[1] (Fig. 2) rule out the possibility that type II topoisomerases make 'random' strand passages. Certain type II topoisomerases (*Drosophila* topoisomerase II, *Escherichia coli* topoisomerase IV) suppressed the knotting of 10-kb P4 DNA to $P_{knot} = 0.00062$, and the knotting of 7-kb PAB4 DNA to $P_{knot} = 0.00019$. The mutual linking probability of P4 DNA strands was similarly suppressed to $P_{link} = 0.004$. Because these type II topoisomerases hydrolyse ATP, their strong suppression of entanglements does not violate the second law of thermodynamics. However, how these small topoisomerases ascertain the topology of large, flexible DNAs is uncertain.

Two mechanisms have been proposed: Rybenkov *et al.*[1] proposed a tracking scheme involving a topoisomerase-mediated synapse between three DNA segments, but without any quantitative analysis. Vologodskii[8] proposed that type II topoisomerases recognize knots by their tendency to have a DNA segment inside the bend of a second DNA segment; computer simulations showed that strand passages directed as in Fig. 3 suppressed P_{knot} to as low as $0.1P^{eq}_{knot}$. Thermal equilibrium is not reached because transitions that are the reverse of that of Fig. 3 are prohibited. This model is described by Fig. 1b, but with nonequilibrium steady-state rates, and a nonequilibrium knotting probability $P_{knot} < P^{eq}_{knot}$. Thus, a nonequilibrium $P_{knot} < P^{eq}_{knot}$ can be obtained by localized collision–knot-recognition events; however, the mechanism of Ref. 8 is unable to explain values of P_{knot} as small as those observed experimentally.

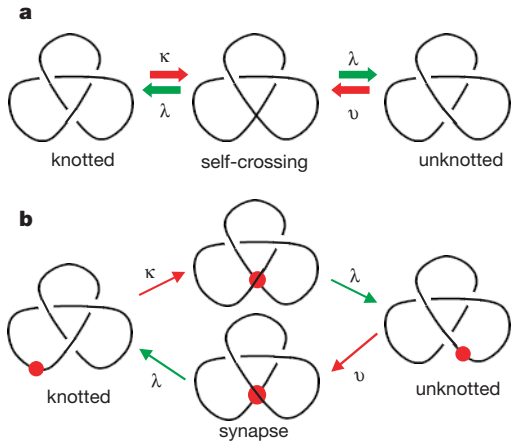

Figure 1 Simplest kinetic models of type II topoisomerases. **a**, Loop of DNA capable of freely crossing itself (a 'ghost' or 'phantom' polymer, such as a strand of DNA that can intermittently break and rejoin[6,7]) spends some fraction of time P^{eq}_{knot} as a knotted strand and some fraction P^{eq}_{unknot} as an unknotted strand. For the 3–10-kb DNA strands used here, knot–unknot free energy differences are entropic, and the strands at the self-crossing point (the X intersection) are not subject to forces that determine knotting or unknotting. Therefore, transitions away from the self-crossing point occur at equal rates λ for knotted and unknotted strands (green arrows); any knot–unknot discrimination must be based on the self-crossing rates κ and υ (red arrows). **b**, A more realistic 'one-way' or 'two-gate'[12] strand passage model of a type II topoisomerase distinguishes two DNA–topoisomerase-DNA 'synapse' states. On isolated circular DNA strands, transitions occur from knotted to unknotted states at a rate $\kappa\lambda$, and back again at a rate $\upsilon\lambda$. For a topoisomerase that does not consume stored energy, $(\upsilon\lambda)/(\kappa\lambda)$ must equal the thermal equilibrium ratio of knotted to unknotted strands for 'ghost' DNA as in **a**. If the topoisomerase uses stored energy, it is possible for the knot probability to be reduced below that expected at thermal equilibrium.

We propose that the type II topoisomerases described above suppress entanglements using a type of 'kinetic proofreading'. This was first discussed in general terms by Hopfield[2] and Ninio[3], who showed how the release of energy (for example, from ATP hydrolysis) could be used to make molecular recognition processes more specific than one would expect at thermal equilibrium. Figure 4 shows our proofreading scheme for type II topoisomerases interacting with DNA plasmids which may be knotted or unknotted. For simplicity, only one type of knot is considered; for the DNA strands of ref. 1, essentially all the knots are expected to be trefoils[9]. We limit our discussion to DNA-disentangling topoisomerases (such as eukaryote type II topoisomerases and *E. coli* type IV topoisomerases) that pass distant, disjointed DNA segments through one another.

We now describe our model in the context of the relatively simple knotting–unknotting single-DNA reaction. Linking–unlinking of two circular DNA strands can be discussed in a similar way for the dilute solution conditions of ref. 1. Starting from either knotted or unknotted strands with a topoisomerase attached (Fig. 4, state 1), a synapse can reversibly occur (Fig. 4, $1 \leftrightarrow 2$; numbers indicate how many DNA segments are bound by the topoisomerase). For ~10-kb plasmids we may assume that the rate of synapse formation on knotted strands exceeds the formation rate on unknotted strands by an amount of at least $P^{eq}_{unknot}/P^{eq}_{knot}$, given that it is possible for a single local recognition step to be used to obtain $P_{knot} < P^{eq}_{knot}$ (ref. 8). As the molecular motions leading to synapsis are the only dependents of these transitions on the knotting state (that is, the off-rate λ values are likely to be determined by the energetics of the topoisomerase-DNA binding rather than by global DNA conformation), we assume that the synapsis rates for knots (κ) and unknots (υ) satisfy $\upsilon/\kappa \leqslant P^{eq}_{knot}/P^{eq}_{unknot}$.

From state 2 (Fig. 4), an irreversible transition can occur to a transient state 1* with only one of the DNA segments bound to an 'activated' topoisomerase (the other DNA segment is released). Irreversibility could be enforced with ATP binding, possibly coupled to topoisomerase conformation change and cleavage of the bound DNA. As no overall DNA conformational change is involved in this step, the rate α should be insensitive to DNA topology.

The state 1* can 'decay' back to 1 at a topology-independent decay rate γ, giving spontaneous 'deactivation' of the topoisomerase with no change in DNA topology. Alternatively, a new synapse can reversibly form between a new DNA segment and the activated topoisomerase–DNA complex ($1^* \leftrightarrow 2^*$), which will lead to strand

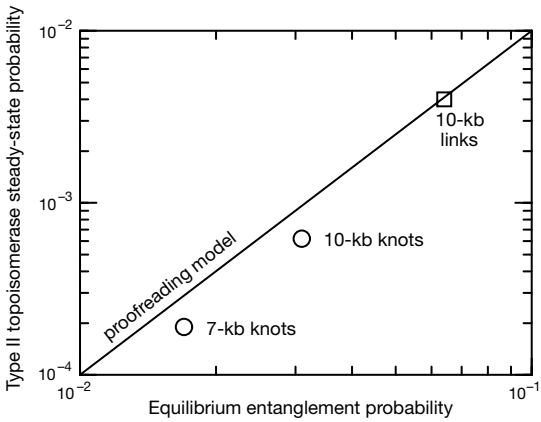

Figure 2 Experimental knotting (circles) and linking (square) probabilities from ref. 1, compared with our model. The horizontal axis shows the thermal equilibrium entanglement probability; the vertical axis shows the steady-state entanglement probability results in the presence of type II topoisomerases and ATP. The kinetic proofreading model (see text) is able to reduce the knotting probability to levels below the solid line, which is defined by (steady-state) = (equilibrium)^2.

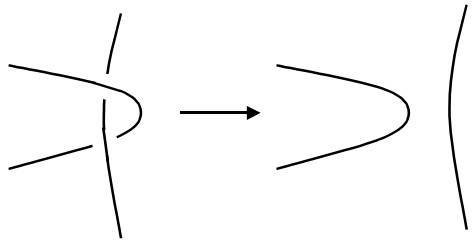

Figure 3 Nonequilibrium bend-recognizing topoisomerase mechanism of Vologodskii. Juxtapositions of sharp DNA bends with straight segments (left) occur more frequently on knotted DNA strands than on unknotted DNA strands[8]. By allowing only the forward strand passage (arrow) to occur, the knotting probability can be forced below that expected at thermal equilibrium. Therefore, knotted strands can be 'recognized' by topoisomerases on the basis of localized DNA geometry.

transfer. As before, the off-rate λ' values are assumed to be insensitive to DNA knotting. As we are only concerned with topology-changing events, we may assume that the ratio of the recollision rates for knots (κ') and unknots (υ') is $\upsilon'/\kappa' = P^{eq}_{knot}/P^{eq}_{unknot}$. We note that if the second synapse again requires specific types of collisions, such as that of Fig. 3, we could have $\upsilon'/\kappa' < P^{eq}_{knot}/P^{eq}_{unknot}$.

Finally, irreversible transitions from $2^* \rightarrow 1$ achieve strand passage, DNA religation, and release of the passed segment. At the end of the $2^* \rightarrow 1$ transitions, the topoisomerase is reset and ready for another cycle. ATP hydrolysis and product release could be involved in this step. These irreversible transitions should be controlled by local topoisomerase–DNA interactions, and should occur at a strand passage rate μ that is independent of DNA topology.

The steady-state knot–unknot ratio for this model is:

$$\frac{P_{knot}}{P_{unknot}} = \frac{(\gamma[\lambda' + \mu] + \kappa'\mu)\upsilon\upsilon'}{(\gamma[\lambda' + \mu] + \upsilon'\mu)\kappa\kappa'} \quad (1)$$

We assume that $\gamma[\lambda' + \mu] \gg \kappa'\mu$, as the recollision rate κ' requires a particular outcome of large-scale DNA conformational change, whereas the strand passage rate μ includes the process of ATP hydrolysis product release, both of which are expected to be slow relative to the rate of DNA release λ', and the decay rate γ. Similarly, $\gamma[\lambda' + \mu] \gg \upsilon'\mu$. These conditions simplify equation (1) to:

$$\frac{P_{knot}}{P_{unknot}} = \frac{\upsilon\upsilon'}{\kappa\kappa'} \leqslant \left(\frac{P^{eq}_{knot}}{P^{eq}_{unknot}}\right)^2 \quad (2)$$

This shows that the reaction of Fig. 4 can reduce the knotting probability below the square of $P^{eq}_{knot}/P^{eq}_{unknot}$, as is observed experimentally (Fig. 2). We note that all of the topology-independent rate constants have dropped out of the final knot probability, leaving a result with no adjustable parameters. The topology-dependent rates υ, υ', κ and κ' could be roughly estimated using molecular dynamics simulations.

The product of υ/κ and υ'/κ' appears in equation (1) because the inputs to the second synapsis are biased by the outcome of the first synapsis; the second stage 'proofreads' the first. For $P^{eq}_{knot} \ll P^{eq}_{unknot}$, as in experiments on 10-kb and 7-kb plasmids, our model explains how knots are so effectively removed by topoisomerases. Type II topoisomerases are also capable of suppressing the linkage probability to roughly the square of that expected for thermal equilibrium (Fig. 2). Further, the distribution of supercoiling in the presence of type II topoisomerases corresponds roughly to a squaring of the equilibrium distribution. This suggests that the discrimination step ($1 \leftrightarrow 2 \rightarrow 1^*$ in our model) recognizes knots, catenanes and supercoils; simulations[8] and structural studies of topos are needed to understand this in detail.

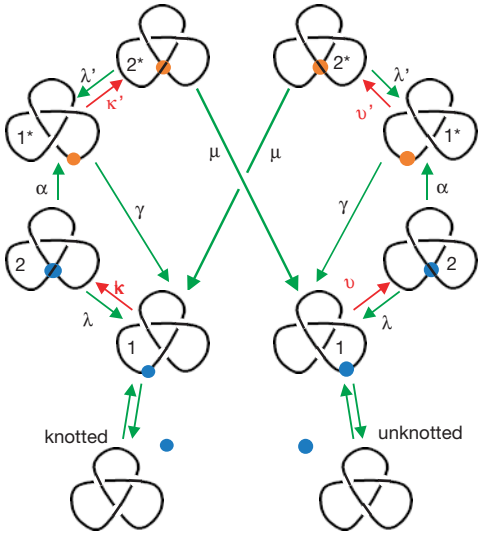

Figure 4 Proposed kinetic model for type II topoisomerase using kinetic proofreading of DNA topology. We note that reactions of the form of Fig. 1b occur twice along the knotting and unknotting pathways. The number of DNA segments bound to the topoisomerase is shown for each loop. 'Activated' topoisomerases (indicated with an asterisk) are able to pass DNA through DNA (see text). The topoisomerase itself is shown in blue when inactive and red when active. By cascading two synapsis events separated by irreversible transitions, the first synapsis ($1 \leftrightarrow 2$) delivers an excess of knotted strands over unknotted strands to the second synapsis ($1^* \leftrightarrow 2^*$); the second reaction 'proofreads' the first. Proofreading can reduce the knot–unknot ratio to below $(P^{eq}_{knot}/P^{eq}_{unknot})^2$. Most of the transitions do not depend on knottedness (green arrows); all discrimination of topology is based on synapse formation rates (red arrows), as in Fig. 1. The topology-independent rate constants do not contribute to the final knot–unknot steady-state ratio of this model.

The model of Fig. 4 is compatible with experimental constraints. The strand passage pathway is 'one-way', in accord with the experimentally supported[10] 'two-gate' model for type II topoisomerases in which the passed strand enters and exits the DNA–topoisomerase complex through different gateways. Other experiments show that a single round of strand passage can occur when non-hydrolysable ATP analogues are used[11,12], and support a ATP-binding-driven DNA-clamp model of topoisomerase function[10].

Our proofreading reaction is compatible with the DNA-clamping model. For example, the $2 \rightarrow 1^*$ transition could correspond to ATP binding, stimulated by the first synapsis. The decay of the 1^* state (γ) could then correspond to the closing of the DNA clamp. The second synapsis would then correspond to DNA recollisions occurring before this topoisomerase conformational change. Non-hydrolysable ATP would block completion of the $2^* \rightarrow 1$ transition, trapping the topoisomerase in an inactive state after one strand passage. Alternatively, given recent experiments indicating two sequential ATP hydrolysis events[13,14], it is tempting to imagine that the first ATP hydrolysis is somehow involved in the $2 \rightarrow 1^*$ transition; however, this is difficult to reconcile with clamp closure triggered by ATP binding.

Finally, proofreading may increase knotting under conditions where knotted strands are more likely than unknotted strands at thermal equilibrium, where topology-changing self-collisions on unknotted strands occur at higher frequency than on knotted strands. This is the case for large DNA strands: ~200-kb plasmids with equilibrated topology have more trefoils than any other type of knots, including unknotted strands[9], and on these molecules type II topoisomerases should generate more trefoils and fewer unknotted strands than expected in equilibrium. Avoidance of a situation in which topoisomerase becomes knot-generating rather than knot-removing suggests arranging large DNA strands into multiple 'loop' domains of less than 100 kb to ensure entanglement removal, as

'links' between the different loop domains. This provides a rationale for organization of, for example, the 4.5-Mb chromosome of *E. coli* into loop domains of approximately this size[15,16]. □

Received 22 December 1998; accepted 18 August 1999.

1. Rybenkov, V. V., Ullsperger, C., Vologodskii, A. V. & Cozzarelli, N. R. Simplification of DNA topology below equilibrium values by type II topoisomerases. *Science* **277,** 690–693 (1997).
2. Hopfield, J. J. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl Acad. Sci. USA* **71,** 4135–4139 (1974).
3. Ninio, J. Kinetic amplification of enzyme discrimination. *Biochimie* **57,** 587–595 (1975).
4. Frank-Kamenetskii, M. D., Lukashin, A. V. & Vologodskii, A. V. Statistical mechanics and topology of polymer chains. *Nature* **258,** 398–402 (1975).
5. Klenin, K. V., Vologodskii, A. V., Anshelevich, V. V., Dykhne, A. M. & Frank-Kamenetskii, M. D. Effect of excluded volume on topological properties of circular DNA. *J. Biomol. Struct. Dyn.* **5,** 1173–1185 (1988).
6. Shaw, S. Y. & Wang, J. C. Knotting of a DNA chain during ring closure. *Science* **260,** 533–536 (1993).
7. Rybenkov, A. V., Cozzarelli, N. R. & Vologodskii, A. V. Probability of DNA knotting and the effective diameter of the DNA double helix. *Proc. Natl Acad. Sci. USA* **90,** 5307–5311 (1993).
8. Vologodskii, A. V. in RECOMB 98: Proceedings of the second annual international conference on computational molecular biology 266–269 (Association for Computing Machinery, New York, 1998).
9. Deguchi, T. & Tsurusaki, K. A statistical study of random knotting using the Vassiliev invariants. *J. Knot Theory Ramific.* **3,** 321–353 (1994).
10. Roca, J. & Wang, J. C. DNA transport by a type II DNA topoisomerase: evidence in favor of a two-gate mechanism. *Cell* **77,** 609–616 (1994).
11. Roca, J. & Wang, J. C. The capture of a DNA double helix by an ATP-dependent protein clamp: a key step in DNA transport by type II DNA topoisomerases. *Cell* **71,** 833–840 (1992).
12. Roca, J., Berger, J. M. & Wang, J. C. On the simultaneous binding of eukaryotic DNA topoisomerase II to a pair of double-stranded DNA helices. *J. Biol. Chem.* **268,** 14250–14255 (1993).
13. Harkins, T. T. & Lindsley, J. E. Pre-steady-state analysis of ATP hydrolysis by *Saccaromyces cerevisiae* DNA topoisomerase II. 1. A DNA-dependent burst in ATP hydrolysis. *Biochemistry* **37,** 7292–7299 (1998).
14. Harkins, T. T. & Lindsley, J. E. Pre-steady-state analysis of ATP hydrolysis by *Saccaromyces cerevisiae* DNA topoisomerase II. 2. Kinetic mechanism for the sequential hydrolysis of two ATP. *Biochemistry* **37,** 7299–7312 (1998).
15. Worcel, A. & Burgi, E. On the structure of the folded chromosome of *Escherichia coli*. *J. Mol. Biol.* **71,** 127–147 (1972).
16. Sinden, R. R. & Pettijohn, D. E. Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling. *Proc. Natl Acad. Sci. USA* **78,** 224–228 (1981).

Acknowledgements

We thank D. Chatenay, N. R. Cozzarelli, G. B. Mindlin, V. Rybenkov, E. D. Siggia, A. V. Vologodskii and E. L. Zechiedrich for discussions. J.F.M. and J.Y. acknowledge the support of the NSF, the Research Corporation, the Petroleum Research Fund, and the Whitaker Foundation. M.O.M. acknowledges support of the Sloan Foundation and the Mathers Foundation.

Correspondence and requests for materials should be addressed to J.F.M. (e-mail: jmarko@uic.edu).

Table 1 Crystallographic data and refinement statistics

Data and refinement statistics (values in parentheses are for the resolution bin 2.53–2.40 Å)	
Space group	C2
Cell dimensions	$a = 165.51$ Å, $b = 95.87$ Å, $c = 211.77$ Å, $\beta = 106.83°$
Resolution range	25–2.4 Å
No. of reflections	103,327 (10,696)
Completeness	84.3% (59.9%)
Multiplicity	2.7 (1.7)
R_{merge}	0.132 (0.287)
R-factor	0.232 (0.312)
R_{free} value	0.265 (0.365)
No. of atoms	12,623
No. of reflections used in refinement	101,663
No. of reflections used for R_{free}	1,662
Solvent content	73.1%
No. of protein atoms (6 × 264 residues)	12,042
No. of water molecules	581
B-value from Wilson plot (3.5–2.4 Å)	35.3 Å^2
Mean B-value	35.2 Å^2
Average protein B-value	35.0 Å^2
Average water B-value	38.0 Å^2
Ramachandran plot of non-glycine and non-proline residues	
Most favourable regions	1,102 (78.5%)
Additional allowed regions	275 (19.6%)
Generously allowed regions	17 (1.2%)
Disallowed regions	10 (0.7%)
r.m.s. deviations from ideal values	
Bond distances	0.006 Å
Angles	1.4°

$R_{merge} = \Sigma_{hkl}\Sigma_i|I_{hkl} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_i\langle I_{hkl}\rangle$ where the sum i is over all separate measurements of the unique reflections hkl.
R-factor $= \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$
R_{free}, as R-factor but summed only over the test reflections
B − value $= 8\pi^2\langle u^2\rangle$ where $\langle u^2\rangle$ is the apparent mean square deviation from the atomic position.

A triple β-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein

Mark J. van Raaij*, Anna Mitraki†, Gilles Lavigne† & Stephen Cusack*

* *European Molecular Biology Laboratory, Grenoble Outstation, c/o Institut Laue Langevin, BP 156, 38042 Grenoble Cedex 9, France*
† *Institut de Biologie Structurale (CEA-CNRS), 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France*

Human adenoviruses[1] are responsible for respiratory, gastroenteric and ocular infections[2] and can serve as gene therapy vectors[3]. They form icosahedral particles with 240 copies of the trimeric hexon protein arranged on the planes and a penton complex at each of the twelve vertices. The penton consists of a pentameric base, implicated in virus internalization[4], and a protruding trimeric fibre, responsible for receptor attachment[5]. The fibres are homo-trimeric proteins containing an amino-terminal penton base attachment domain, a long, thin central shaft and a carboxy-terminal cell attachment or head domain. The shaft domain contains a repeating sequence motif with an invariant glycine or proline and a conserved pattern of hydrophobic residues[6]. Here we describe the crystal structure at 2.4 Å resolution of a recombinant protein containing the four distal repeats of the adenovirus type 2 fibre shaft plus the receptor-binding head domain. The structure reveals a novel triple β-spiral fibrous fold for the shaft. Implications for folding of fibrous proteins (misfolding of shaft peptides leads to amyloid-like fibrils) and for the design of a new class of artificial, silk-like fibrous materials are discussed.

The human adenovirus serotype 2 (Ad2) fibre is a trimer of 582 residues per monomer[7], of which the head domain is essential for trimerization and autonomously trimerizes when expressed[8]. The high-resolution structures of the heads of Ad2 and Ad5 are known[9,10] and are very similar: each head monomer forms an eight-stranded anti-parallel β-sandwich structure and the three monomers interact to form a three-bladed propeller. The head domain is responsible for binding to the cell receptor, which has been identified to be a human protein of unknown function: the coxsackievirus and adenovirus receptor[11]. This protein serves as the receptor for coxsackieviruses of subgroup B and adenoviruses of all subgroups except subgroup B (ref. 12). The primary sequence of the fibre shaft consists of 15-residue pseudo-repeats (22 of them for Ad2 and Ad5, ref. 7). Green *et al.*[6] predicted that these repeats contain two β-strands and two turns (the cross-β model). Stouten *et al.*[13] subsequently proposed a triple β-helical model, taking into account length measurements from electron microscopy and fibre diffraction patterns[14]. The full-length fibre is very stable, resistant to heat (its melting temperature is 85 °C, ref. 15) and detergents (at low temperatures[16]), and the shaft domain is highly resistant to proteases[17].

The Ad2 fibre unfolds through a stable intermediate in which the C-terminal head and distal part of the shaft remain folded and trimeric[16]. The stable domain has been identified to span residues 319–582 and has been cloned and expressed in *Escherichia coli* (M.J.v.R. *et al.*, unpublished results). The recombinant protein was