Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 2141–2144, March 1995
Genetics

Antiparallel, intramolecular triplex DNA stimulates homologous recombination in human cells

SEAN M. ROONEY† AND PETER D. MOORE

Department of Genetics, University of Illinois, Chicago, IL 60612

Communicated by Frederick W. Alt, Howard Hughes Medical Institute, Boston, MA, December 5, 1994 (received for review July 5, 1994)

ABSTRACT **The DNA motif 5′-AAGGGAGAAXGGGTATAGGGYAAGAGGGAA-3′ (named XY32) is an H-palindrome and has been shown to undergo a superhelix-induced, pH-dependent structural transition to H-form (pyrimidine•purine○pyrimidine triplex) DNA when X=Y=A (AA32) or X=Y=G (GG32), but when X=A and Y=G (AG32) or X=G and Y=A (GA32), the transition is much more difficult [Mirkin, S. (1987) *Nature (London)* 330, 495–497]. Furthermore, AA32, GG32, and GA32 triplexes have the proper sequence structure to potentially form pyrimidine•purine○purine (*H-form) triplexes, but AG32 does not [Beal, P. A. & Dervan, P. B. (1992) *Nucleic Acids Res.* 20, 2773–2776]. Using an *in vivo* plasmid–plasmid recombination assay system in cultured human cells, we have found that AA32, GA32, and GG32 stimulate homologous recombination between plasmids 3- to 5-fold when both recombination substrates contain these triplex-forming sequences, whereas AG32, which differs from the others by only 1 or 2 bp, does not significantly affect the frequency of recombination. Double-strand breaks, which destroy supercoiling, nullify the stimulation. Therefore, stimulation of homologous recombination between plasmids containing these sequences correlates with their triplex-forming potential. Crosses in which the triplex-forming sequence is inserted into only one substrate exhibit an intermediate stimulation, suggesting that the inserts are acting alone as intramolecular triplexes.**

The idea that triplex DNA may be an intermediate in homologous recombination was first advanced in 1984 (1). More recently, results have been published that suggest that RecA protein acts via long three-stranded DNA intermediates. Specifically, it has been reported that three-stranded complexes exist as intermediates for RecA-mediated strand-transfer reactions *in vitro* and that a three-stranded DNA structure can be stable even after RecA protein is removed (2–5). However, whether a triple DNA helix is actually an intermediate is controversial. Some investigators (6) have reported that most RecA-mediated three-stranded DNA structures are unstable without RecA protein present. Others (7) assert that any such stable structure is an artifact attributable to impurities in RecA protein preparation. It has been shown also that these structures do not include interactions with the guanine N-7 atom, which all previously reported triplex DNA structures do utilize (8). A structure called R-DNA has been proposed to explain this problem (5). R-DNA is a triple helical structure in which homologous (as opposed to complementary) DNA strands are parallel to each other with respect to their 5′-3′ orientations and which does not require interactions involving guanine N-7. Although R-DNA is still very hypothetical, triplexes in which homologous strands are antiparallel have been extensively studied in naked DNA. These structures are formed by perfect or near-perfect homopurine-homopyrimidine repeats. There are two forms of these triplexes, H-DNA, or Y•R○Y (pyrimidine•purine○pyrimidine, where "•" represents a Watson–Crick base-pairing bond, and "○" represents a Hoogsteen bond), and *H-DNA, or Y•R○R. In cloned genomic DNA there exist homopurine-homopyrimidine repeats that can form intramolecular triplexes of either H or *H configuration (9, 10). These are potential hotspots of recombination, but there is to date no data available as to their direct role. We have attempted to address the question of the role of homopurine-homopyrimidine repeats in homologous recombination. These sequences can form both inter- and intramolecular triplexes. We have used a mutational approach to look for the link between these structures and recombination in human cells.

The sequence studied, the homopurine-homopyrimidine H-palindrome motif 5′-AAGGGAGAAXGGGTATAGGGYAAGAGGGAA-3′ (named XY32) has been shown to undergo a superhelix-induced, pH-dependent structural transition to intramolecular H-DNA when X=Y=A or X=Y=G, named AA32 and GG32, respectively, but when X=A and Y=G (AG32) or X=G and Y=A (GA32), the transition is much more difficult (11). Furthermore, AA32, GG32, and GA32 have the proper sequence structure to potentially form intramolecular *H triplexes at neutral pH, but AG32 does not (12).

Mammalian cells in culture can mediate homologous recombination between transfected plasmids. This study takes advantage of this ability and uses a plasmid–plasmid *in vivo* recombination assay with the pSV2neo shuttle vector in the human bladder carcinoma cell line EJ. We compared the frequency of recombination between two pSV2neo-derived plasmids, each containing a specific deletion, with the frequency of recombination between the same plasmids containing XY32 inserts. Our results show that the triplex-forming XY32 sequences stimulate homologous recombination and that they are acting alone as intramolecular DNA triplexes.

MATERIALS AND METHODS

Constructs. The constructs are derived from the shuttle vector pSV2neo (Fig. 1), which contains the neomycin phosphotransferase gene under control of the simian virus 40 promoter (13). Two recombination substrates, named pDL (deletion left) and pDR (deletion right) were previously constructed (14). The versions of pDL and pDR used in this study were slightly modified at a few restriction sites (15). Putative recombination-stimulating sequences were excised from the four plasmids pXY32, obtained from Sergei Mirkin, University of Illinois, Chicago and described elsewhere (11), with *Eco*RI and *Hin*dIII; the *Eco*RI site was filled in and the resulting fragment was inserted between the *Stu* I and *Hin*dIII sites of pDL and pDR, outside but near the region within which the recombination event must occur to produce a functional *neo* gene, with the 3′ end of the above-listed sequence oriented toward the start of the *neo* gene. Twelve recombination substrates were constructed and named pDLXY or pDRXY. The pSV2neoXY control plasmids were constructed similarly using pSV2neo as the target vector.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†To whom reprint requests should be addressed.

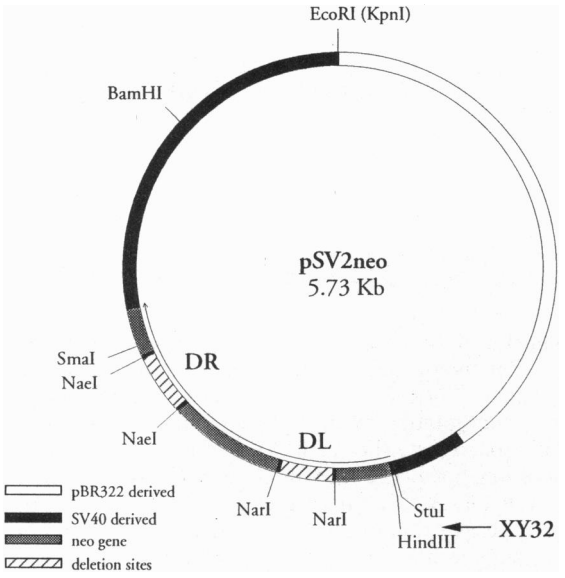

FIG. 1. Construction of recombination substrates. pSV2neo confers resistance to G418 in transfected mammalian cells. Deletion left (pDL) was made by removing the 248-bp *Nar* I–*Nar* I fragment, and deletion right (pDR) was made by removing the 283-bp *Nae* I–*Nae* I fragment. Neither deletion plasmid alone will confer G418 resistance. Restriction-site polymorphisms exist at several positions in the deletion plasmids, which were created for an earlier study. In the plasmids used here, the *Eco*RI site of pSV2neo is replaced with a *Kpn* I site. XY32 sequences were inserted as shown (XY32 is 5′-AAGGGAGAAXGGGGTAT-AGGGYAAGAGGGAA-3′). SV40, simian virus 40.

Tissue Culture and Transfections. The EJ cell line, from a human bladder carcinoma, was provided by Raju Kucherlapati (Albert Einstein College of Medicine). Cells were maintained in DMEM/10% fetal calf serum. Transfections were by the calcium phosphate precipitation method (16). Five micrograms of pSV2neo or 50 μg of each recombination substrate plasmid DNA was introduced into ≈5 × 10^6 cells on five 60-mm plates. Cells were incubated in normal medium for 24 hr after transfection and then transferred to 100-mm plates in medium containing the aminoglycoside G418 at 400 μg/ml. The cells were fed every 3–4 days for 16 days, and the surviving colonies were stained with Giemsa stain and counted.

RESULTS

Triplex DNA Promotes Homologous Recombination. Our assay for homologous recombination is well characterized and has been used to identify other sequences that stimulate recombination (14, 15, 17). The recombination substrates were mixed together and introduced into EJ cells by calcium phosphate-mediated transfection. If a homologous recombination event occurs between the plasmids in the region between the deletions, a wild-type *neo* gene is reconstituted, and the cell becomes stably resistant to G418 when the recombination product integrates into the genome. Transfected cells were maintained in selective medium for 16 days, and the surviving colonies were counted. In most experiments representative colonies were isolated and expanded for analysis. The apparent recombination frequency is calculated as the number of colonies per microgram of input DNA that arise from cells cotransfected with two recombination substrate plasmids divided by the colonies per microgram from cells transfected with pSV2neo. None of the recombination substrates can confer G418 resistance alone, and none of the three inserts tested has an effect on *neo* gene expression when inserted into the parental pSV2neo plasmid, as measured by colony counts (Table 1). Insertion of AA32, GA32, or GG32 sequence in both substrates resulted in stimulation of recombination by 3- to 5-fold over the unsubstituted cross DL × DR, whereas enhancement by AG32 sequence, which forms neither H- nor *H-DNA triplex, is insignificant (Fig. 2). Though the ability of these sequences to form *H-DNA triplex has not been shown directly, as it has for H-DNA triplex, the fact that GA32 sequence stimulates equally as well as AA32 and GG32 sequences suggests that *H-DNA triplex is the form involved in these reactions.

Table 1. Controls

DNA	Exp., no.	Total colonies	Total DNA, μg	Average colonies per μg
Recombination substrates transfected alone				
DLAA	4	0	200	0.0
DRAA	4	0	200	0.0
DLAG	4	0	200	0.0
DRAG	4	0	200	0.0
DLGA	4	0	200	0.0
DRGA	4	0	200	0.0
DLGG	3	0	150	0.0
DRGG	4	0	200	0.0
XY32 inserts in parental pSV2neo plasmid				
pSV2neo	27	5309	151	35.23
pSV2neoAA	3	520	15	34.7
pSV2neoAG	3	470	15	31.3
pSV2neoGG	3	467	15	31.1

Double-Strand Breaks Negate the Triplex-Induced Stimulation. Formation of triplex DNA, whether H or *H form, depends on negative superhelical stress. Throughout this study, cesium chloride-purified, and thus closed-circular supercoiled, plasmid DNA was used. When double-strand breaks were introduced with the restriction enzyme *Kpn* I, which has a specific site outside and far from the region of interest (see Fig. 1 legend), there was a minor increase in recombination frequency between unsubstituted plasmids, as has been reported (18). The presence of the triplex-forming sequence GG32 resulted in no significant additional increase (Fig. 2); this result further supports the hypothesis that formation of triplex DNA, rather than the XY32 sequence *per se*, is relevant in these experiments.

The Triplex Structures Involved Are Intramolecular. When GG32 sequence is in only one substrate, there is a slight stimulation, of marginal statistical significance, but less than when both substrates contain GG32 sequence (Fig. 2). In reactions where the triplex-forming insert is in both substrates, the stimulation for GG32 sequence is ≈2-fold that seen in these "heterozygous" crosses, an additive effect. If the two inserts interacted in some way to stimulate recombination, such as forming an intermolecular triplex, a synergistic effect would be expected. Furthermore, when AA32 sequence is in one substrate and GG32 sequence is in the other, there is an ≈5-fold enhancement of recombination (Fig. 2). Both AA32 and GG32 sequences can form intramolecular triplexes, but it should be quite difficult for them to form an intermolecular triplex between them because they are mismatched at 2 bases. Therefore, it is unlikely that an intermolecular triplex structure is involved in the stimulation. Taken together, the above data suggest that intramolecular, rather than intermolecular, triplexes are responsible for the observed stimulation of recombination.

DISCUSSION

We have shown that sequences that can form an intramolecular DNA triplex can stimulate homologous recombination between plasmids introduced into human cells. The sequences named AA32 and GG32 can form intra- and intermolecular antiparallel triplexes where the base triplets are either pyrimidine•

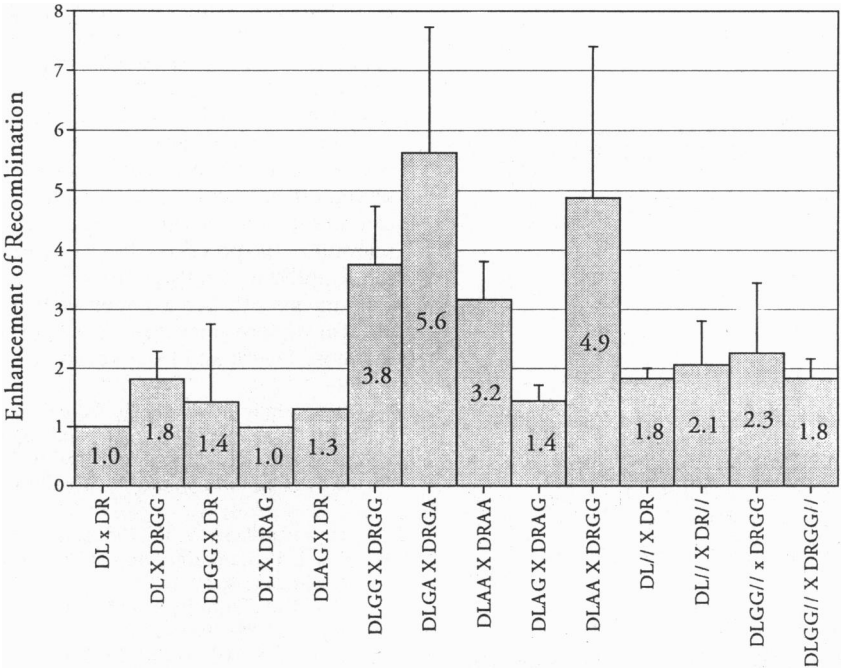

FIG. 2. Results of recombination assays. In each experiment 100 μg of DNA was transfected into 5×10^6 cells by calcium phosphate transfection, the cells were grown in selective medium for 16 days, and surviving colonies were counted. Enhancement ratio was calculated as the number of colonies per microgram from a given set of substrates divided by the number of colonies per microgram from the unsubstituted cross DL × DR. Error bars represent 1 SD. // represents a double-strand break.

purine∘pyrimidine (H-DNA) or pyrimidine•purine∘purine (*H-DNA), and GA32 sequence can form only *H-DNA. [Formation of *H-DNA with these inserts has not been shown directly but was inferred from basic principles (12)]. All three sequences stimulate recombination 3- to 5-fold. Double-strand breaks, which destroy the triplex-forming potential of the inserts by relaxing the necessary supercoils, abrogate the stimulation. The sequence named AG32, which differs from either AA32 or GG32 by only 1 bp each and at different positions, can only with great difficulty form either type of triplex (again, formation or lack of formation of *H form has not yet been shown directly), and its stimulation of homologous recombination is insignificant. Therefore, the ability to form triplex DNA correlates with stimulation of homologous recombination between plasmids; this result reinforces the hypothesis that triplex DNA is involved in homologous recombination. The observations that there is an intermediate stimulation when the triplex-forming insert is in only one substrate, that either substrate contributes an equal amount of stimulation in these experiments, and that crosses involving mismatched triplex-forming inserts stimulate equally as well as when both substrates match suggest that an intramolecular, rather than an intermolecular, triplex is involved and further strengthens the case for triplex DNA being necessary to the observed stimulation. The fact that GA32 sequence stimulates at least as well as AA32 and GG32 sequences implies that *H-, vs. H-DNA is the form of triplex that is relevant for this effect.

Homologous recombination in human cells is almost surely an enzyme-mediated process, involving mechanisms similar to more well-characterized recombination reactions. If the human recombinase involved is similar to RecA protein in bacteria, then it is likely that its mechanism includes a triplex DNA intermediate. In support of this idea, our results show that sequences that more easily form triplex DNA increase the frequency of recombination when they are included in the recombination substrates. However, it must be emphasized that the intramolecular, antiparallel triplexes discussed here are fundamentally different from (intermolecular, parallel) R-DNA, the proposed intermediate for RecA-mediated recombination. In addition to the structural differences, these triplexes have strict sequence requirements, whereas R-DNA can theoretically form using any DNA sequence, as long as the substrates are homologous. It is possible that our inserts are involved in recombination indirectly. For example, the intramolecular triplex structure necessitates the formation of a stretch of single-stranded DNA equal in length to half the palindrome. The presentation of this single-stranded stretch could itself be recombinogenic. Alternatively, either this single-stranded stretch or some other aspect of the triplex structure could be a target for nucleases, resulting in single- or double-stranded cuts in the DNA, which could themselves be recombinogenic. Whatever the mechanism, it remains that sequences that more easily form this structure are recombinogenic and are therefore likely to represent recombination hotspots. These sequences could be binding sites for a recombinase or some other part of a recombination complex, perhaps an enzyme that forms an intramolecular triplex as an intermediate to present single strands. This enzymatic help would relax the strict sequence requirements for triplex formation. The XY32 sequences have some similarity to the hypervariable minisatellite sequences [the sequence used in the previous work (15) was a synthetic consensus hypervariable minisatellite of the form $(AGAGGTGGGCAGGTGG)_{6.5}$] that have been previously shown to be both recombinogenic (15) and specific binding sites for a nuclear protein in HeLa cells (19). In fact, a human antiparallel triplex DNA-binding protein has been reported as well (20), which somewhat reinforces this possibility. We have shown that sequences differing by only 1 bp have very different abilities to stimulate homologous recombination. It is unlikely that this small a difference in sequence would have so great an effect on protein binding as is demonstrated here for the complete reaction, unless such binding is enhanced by triplex formation or triplex-forming potential. In addition, the similarity with hypervariable minisatellite sequences works both ways. These minisatellite sequences may also form triplexes more easily than other DNA, and this may explain or at least contribute to their recombinogenic activity. The synthetic minisatellite sequences tested stimulate recombination in this system signif-

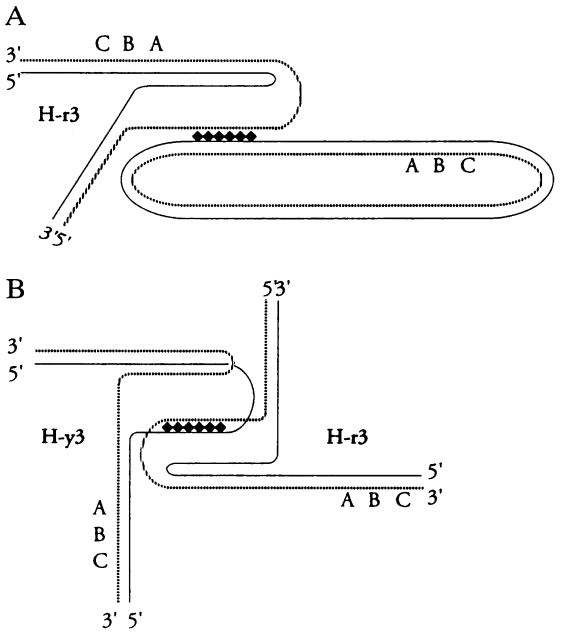

FIG. 3. Mechanistic models. (*A*) The intramolecular triplex structure provides a region of single-stranded DNA, which then interacts with its homologous duplex counterpart by strand invasion, etc. (*B*) One plasmid forms an intramolecular H-DNA (H-y3) triplex, and the other forms an intramolecular *H–DNA (H-r3) triplex. Their respective single-stranded regions can base-pair, forming a heteroduplex. Solid lines, purine strand (with respect to triplex-forming inserts); dashed lines: pyrimidine strand; ABC denotes orientation of *neo* gene. H-r3, *H-DNA (the 3′ half of the purine strand bends back to form the triplex); H-y3; H-DNA (the 3′ half of the pyrimidine strand bends back).

icantly more than the XY32 sequences, ≈13- to 15-fold (15). One reason for this difference may be that the minisatellite sequences consisted of multiple direct repeats and were ≈100 bp in total length, vs. one repeat of 32 bp for XY32. Aside from enhancing mechanisms such as specific protein binding, this would also allow them to form intramolecular triplexes more easily than a single repeat.

Several mechanistic models to explain our observations are possible. In comparing these reactions with those mediated by RecA protein, it is important to remember that R-DNA, the putative RecA intermediate, is an intermolecular, parallel triplex structure (3, 4). Two models that use intramolecular antiparallel triplexes are illustrated. The intramolecular triplex structures formed by XY32 sequence include a single-stranded region of DNA, which would be expected to be recombinogenic (Fig. 3A). Recombination hotspots, perhaps in the form of minisatellite DNA, may work in this way. Another model is possible if one plasmid forms H-DNA and its partner forms *H-DNA, so that their respective single-stranded regions can pair, bringing the molecules into close contact and making a recombination event more likely (Fig. 3B). In a previous study of the effect on recombination of Z-DNA-forming sequences, Wahls *et al.* (17) proposed a similar model, in which adjacent B- and Z-DNA regions could form homologous (complementary) single-stranded stretches that could interact.

Even without regard to mechanisms of recombination, the correlation between *in vitro* triplex-forming ability and an observed effect *in vivo* indicates that intramolecular triplex DNA can, in fact, form inside human cells and that it is functionally relevant to the process of homologous recombination. Although the hypothesis is difficult to support from these data, further experiments are being undertaken to test the possibility that H-form triplexes may assist in the formation of R-form triplex, the more logical and likely recombination intermediate.

We gratefully acknowledge Sergei Mirkin for the gift of XY32 plasmids, the drawings for Fig. 3, and countless helpful discussions, as well as Roya Tavasolli for technical assistance. This work was supported by National Institutes of Health Grant GM45141.

1. Howard-Flanders, P. (1984) *Nature (London)* **309,** 215–220.
2. Hsieh, P., Camerini-Otero, C. S. & Camerini-Otero, R. D. (1990) *Genes Dev.* **4,** 1951–1963.
3. Rao, B. J., Dutreix, M. & Radding, C. M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2984–2988.
4. Umlauf, S. W., Cox, M. M. & Inman, R. B. (1990) *J. Biol. Chem.* **265,** 16898–16912.
5. Zhurkin, V. B., Raghunathan, G., Ulyanov, N. B., Camerini-Otero, R. D. & Jernigan, R. L. (1994) *J. Mol. Biol.* **239,** 181–200.
6. Rao, B. J., Chiu, S. K. & Radding, C. M. (1993) *J. Mol. Biol.* **229,** 328–343.
7. Bedale, W. A., Inman, R. B. & Cox, M. M. (1993) *J. Biol. Chem.* **268,** 15004–15016.
8. Jain, S. K., Inman, R. B. & Cox, M. M. (1992) *J. Biol. Chem.* **267,** 4215–4222.
9. Kohwi, Y. & Panchenko, Y. (1993) *Genes Dev.* **7,** 1766–1778.
10. Weinreb, A. D., Collier, B. K., Birshtein, R. D. & Wells, R. D. (1990) *J. Biol. Chem.* **265,** 1352–1359.
11. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. & Frank-Kamenetskii, M. D. (1987) *Nature (London)* **330,** 495–497.
12. Beal, P. A. & Dervan, P. B. (1992) *Nucleic Acids Res.* **20,** 2773–2776.
13. Southern, P. J. & Berg, P. (1982) *J. Mol. Appl. Genet.* **1,** 327–341.
14. Kucherlapati, R. S., Eves, E. M., Song, K. Y., Morse, B. S. & Smithies, O. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3153–3157.
15. Wahls, W. P., Wallace, L. J. & Moore, P. D. (1990) *Cell* **60,** 95–103.
16. Lowy, D. R., Rands, E. & Scolnick, E. M. (1978) *J. Virol.* **26,** 291–298.
17. Wahls, W. P., Wallace, L. J. & Moore, P. D. (1990) *Mol. Cell. Biol.* **10,** 785–793.
18. Song, K. Y., Schwartz, F., Maeda, N., Smithies, O. & Kucherlapati, R. S. (1985) *Mol. Cell. Biol.* **5,** 3331–3336.
19. Wahls, W. P., Swenson, G. & Moore, P. D. (1991) *Nucleic Acids Res.* **19,** 3269–3274.
20. Kiyama, R. & Camerini-Otero, R. D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10450–10454.