3944–3948 Nucleic Acids Research, 1995, Vol. 23, No. 19

Mutation detection using immobilized mismatch binding protein (MutS)

Robert Wagner*, Paul Debbie and Miroslav Radman[1]

Genecheck Inc., 400 East Horsetooth Road, Fort Collins, CO 80525, USA and [1]Institute Jacques Monod, 2 Place Jussieu, Tour 43, 75251 Paris Cedex 05, France

Received June 15, 1995; Revised and Accepted August 20, 1995

ABSTRACT

An accurate and highly sensitive mutation detection assay has been developed. The assay is based on the detection of mispaired and unpaired bases by immobilized mismatch binding protein (*Escherichia coli* MutS). The assay can detect most mismatches and all single base substitution mutations, as well as small addition or deletion mutations. The assay is simple to use and does not require the use of either radioactivity or gel electrophoresis.

INTRODUCTION

A large number of human genetic diseases are caused by small genetic alterations, including single base substitutions and small additions or deletions. Such mutations may be inherited (inherited syndromes), arise *de novo* in the germline (sporadic diseases) or be acquired somatically (e.g. cancers). The development of diagnostic tests for such small DNA alterations will facilitate both the prevention and treatment of a wide variety of diseases. In addition, the ability to scan a large number of DNA samples for small differences will be useful for large scale studies of polymorphism in human and other species and in identifying unknown genes.

The methods that, to date, have been most successful in detecting small genetic alterations fall into two broad categories: (i) those based on sequence- or mismatch-dependent variability in electrophoretic mobilities; (ii) those based on proteins capable of detecting mispaired bases in heteroduplex DNA (1–7). The former class, while reasonably accurate, is technically demanding and requires the use of polyacrylamide gel electrophoresis, thus making the assays labor intensive, somewhat difficult to automate and difficult to apply to the rapid screening of a large number of samples. Mismatch detection assays also fall into two broad classes: (i) those which involve chemical or enzymatic cleavage of mismatch-containing heteroduplexes at the site of a mismatch (1,2,5,6); (ii) those which involve binding of mismatch-containing heteroduplexes (7). All mismatch cleaving assays and most mismatch binding assays require gel electrophoresis. Mismatch detection assays involving gel mobility shifts require the identification of protein–DNA complexes in polyacrylamide gels (3). Cleavage of mismatch-containing heteroduplexes requires subsequent identification of specific fragments via gel electrophoresis, as do mismatch binding assays involving nuclease protection (4).

Enzymatic mismatch cleavage recognizes distortions produced by disruptions in base pairing, such that those mismatches which produce maximal helical distortion and occur in A:T-rich regions are best recognized. However, the most frequently occurring replication errors arise from mismatches (or unpaired bases) which cause minimal helical distortion and occur most frequently in G:C-rich regions (8), which may make it difficult for enzymatic cleavage methods to detect some of the most commonly occurring mutations.

The specificity of mismatch binding proteins involved in mismatch, i.e. replication error, repair *in vivo* should make them ideally suited to mutation detection. Mismatch binding proteins recognize best those mismatches and unpaired bases which most resemble base pairs and which are, therefore, most likely to occur as replication errors. In addition, mismatch binding proteins recognize mismatches best in regions of high G:C content. However, the use of mismatch binding proteins in mutation detection has, heretofore, met with limited success. The results reported here indicate that immobilized mismatch binding protein exhibits enhanced ability to discriminate between DNA with and without mismatches relative to mismatch binding protein in solution. Thus immobilization facilitates mutation detection by mismatch binding protein and represents a novel approach to mutation/polymorphism detection. The assay is simple to use, accurate and readily amenable to automation.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides of the sequence biotin-GCACCTGACTC-CTGXGGAGAAGTCTGCCGT were annealed to unlabeled complementary oligonucleotides to form all possible mismatches (heteroduplexes) and a G:C base pair (homoduplex). Heteroduplexes were also prepared with unpaired bases by inserting the following bases between positions 15 and 16 of the complementary (non-biotinylated) strand of homoduplex molecules: (i) C, (ii) CA, (iii) CAG, (iv) CAGG.

Immobilized mismatch binding protein assay

MutS (Genecheck Inc.; 500 ng/well) in reaction buffer (20 mM Tris–HCl, pH 7.6, 5 mM $MgCl_2$, 0.1 mM dithiothreitol, 0.01 mM

* To whom correspondence should be addressed

EDTA) was bound to nitrocellulose pre-wet with reaction buffer in a 48-well slot blotting apparatus (Hoeffer). Reaction buffer without MutS was added to control wells. After 20 min at room temperature nitrocellulose was blocked with 200 µl/well 3% horseradish peroxidase (HRP)-free bovine serum albumin (BSA). After 1 h excess blocking solution was removed under vacuum and DNA (1 and 10 ng) was added in 20 µl reaction buffer plus 3% BSA. After 30 min at room temperature wells were washed five times with 100 µl reaction buffer. All washes were poured out rather than removed under vacuum. Streptavidin-conjugated HRP (100 µl, 0.05 µg/ml) in reaction buffer plus BSA was added to each well. After 20 min at room temperature any remaining solution was poured out and the wells washed five times with 100 µl reaction buffer as described above. The nitrocellulose sheet was removed from the apparatus, washed four times with 50 ml reaction buffer, blotted dry, immersed in 10 ml ECL development solution (Amersham) for 1 min, blotted dry and exposed to X-ray film.

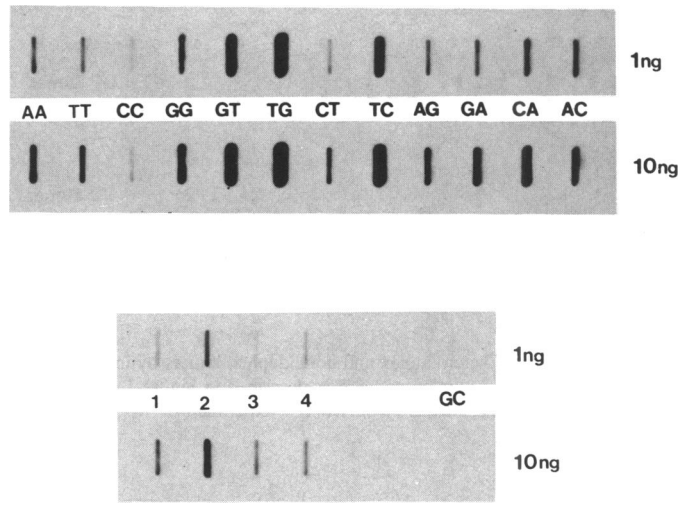

Figure 1. Binding of synthetic oligonucleotides by immobilized MutS. DNA (see Materials and Methods) with mismatches or a G:C base pair at position 15 or with one to four unpaired bases between positions 15 and 16 bound by MutS immobilized on nitrocellulose and revealed by chemiluminescence. Data are from a single experiment. Exposure time was 1 min.

Human genomic DNA

Human genomic DNA was PCR amplified to obtain the following specific fragments of the human glucokinase gene.

Exon 3.
Het-3a. The template was human genomic DNA known to be heterozygous for a transition mutation (G:C→A:T) in exon 3 of the glucokinase gene. The DNA was obtained from CEPH (Paris, France).
Het-3b. The template was human genomic DNA known to be heterozygous for a transversion mutation (G:C→C:G) in exon 3 of the glucokinase gene. The DNA was obtained from CEPH (Paris, France).
Hom-3. The template was human genomic DNA presumed to be homozygous in exon 3 of the glucokinase gene. The DNA was obtained from Sigma.
The primers were 5′-biotin–GGCTGACACACTTCTCTCT and 5′-GATGGAGTACATCTGGTGTT. The amplified fragment was 150 bp long.

Exon 6.
Het-6. The template was human genomic DNA known to be heterozygous for a transition mutation (G:C→A:T) in exon 6 of the glucokinase gene. The DNA was obtained from CEPH (Paris, France).
Hom-6a. The template was human genomic DNA presumed to be homozygous in exon 6 of the glucokinase gene. The DNA was obtained from Sigma.
Hom-6b. The template was human genomic DNA known to be homozygous in exon 6 of the glucokinase gene. The DNA was obtained from CEPH (Paris, France).
The primers were 5′-biotin–CAGCTTCTGTGCTTCTTG and 5′-TGAAGCCGTTTGTACACAG. The amplified fragment was 187 bp long.

Exon 2.
Het-2. The template was human genomic DNA known to be heterozygous for a transversion mutation (G:C→T:A) in exon 2 of the glucokinase gene. The DNA was obtained from CEPH (Paris, France).
Hom-2. The template was human genomic DNA presumed to be homozygous in exon 2 of the glucokinase gene. The DNA was obtained from Sigma.
The primers were 5′-biotin–GAAGGTGATGAGACGGAT and 5′-CCCAGGAGATTCTGTCTC. The amplified fragment was 230 bp long.

RESULTS AND DISCUSSION

Mismatch binding protein was immobilized by binding to nitrocellulose. The mismatch binding protein utilized in the experiments reported here is *Escherichia coli* MutS, which operates *in vivo* as the mismatch recognizing component of the *E.coli* mismatch repair system (9). Although all mismatches are not repaired with equal efficiency, either *in vivo* (10) or *in vitro* (11), MutS has been shown to bind *in vitro* to all mismatches and to heteroduplexes with one to four unpaired bases (11,12).

The results reported here are from experiments in which MutS was immobilized by binding to nitrocellulose. Other solid supports, including nylon and PVDF membranes, have been successfully employed as well (results not shown). The results of experiments utilizing synthetic 5′-biotinylated 30mers with and without mismatches or unpaired bases are shown in Figure 1. The sequence of the 30mers was taken from the β-globin gene at the region surrounding the sickle cell anemia mutation. The mismatches are at position 15 and the unpaired bases are between positions 15 and 16. Signals are generated by means of chemiluminescence. Immobilized mismatch binding protein readily detects all mismatches except C:C, which is the one mismatch which has been found to be generally refractory to repair by the *E.coli* mismatch repair system, both *in vivo* and *in vitro* (10). Heteroduplexes with one or two unpaired bases are readily detected. Heteroduplexes with three or four unpaired bases are somewhat less well detected. With immobilized MutS there is excellent discrimination between mismatched and non-mismatched oligonucleotides, the ratio of binding of G:T-containing to perfectly matched oligonucleotides

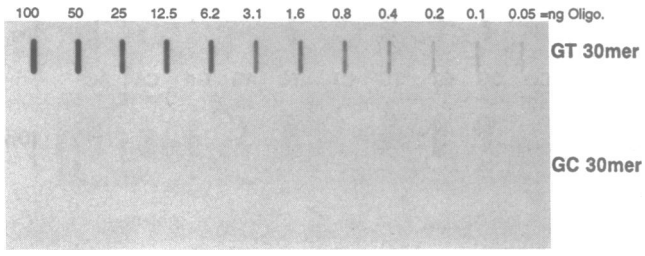

Figure 2. Binding of heteroduplex and homoduplex 30mers by immobilized MutS. DNA and assay conditions were as described in Figure 1 except that exposure time was 20 s. Data are from a single experiment.

(i.e. the ratio of the lowest concentrations at which a signal is detected) is of the order of 1000:1, whereas the ratio of binding with MutS in solution is only ~5:1 (Fig. 2 and Table 1).

Table 1. Binding of mismatch-containing DNA by MutS in solution

DNA (ng)	G:C base pair (c.p.m.)	G:T mismatch (c.p.m.)	Ratio mismatch: base pair
0.1	80	142	1.8
1.0	242	1252	5.2
10.0	1403	7236	5.2

Biotinylated oligonucleotides (described in the legend to Fig. 1) were labeled with ^{32}P by T4 polynucleotide kinase. Labeled oligonucleotides were annealed with unlabeled oligonucleotides as described in the legend to Figure 2 to produce 30mers without mismatches (homoduplexes, G:C) or heteroduplexes with G:T mismatches at position 15.
MutS (500 ng) was incubated in 20 µl reaction buffer (20 mM Tris–HCl, pH 7.6, 5 mM $MgCl_2$, 0.1 mM dithiothreitol, 0.01 mM EDTA) with DNA at room temperature for 30 min. The mixtures were spotted onto 25 mm nitrocellulose filters pre-wet with reaction buffer. The filters were washed five times with 2 ml reaction buffer by vacuum filtration and dried at 80°C for 15 min. Each filter was placed in 3 ml scintillation fluid and the radioactivity determined by scintillation counting. Background counts (no MutS) were subtracted. The results presented are the means of duplicate or triplicate experiments.

The finding that T:C mismatches are better detected than C:T mismatches suggests that mismatch recognition may depend on the sequence of the individual strands, i.e. the sequence in the vicinity of the mismatch and the orientation of the mismatch with respect to the strand, at least in relatively small oligonucleotides such as those used in these experiments. However, the detectable mismatches (all mismatches except C:C) are detectable independent of orientation. In addition, G:T and T:G mismatches are detected equally well, suggesting that well-detected mismatches are well detected independent of strand orientation. It cannot be excluded that some of the variation observed in the extent of binding is due to errors in the oligonucleotide synthesizing process. Some 30mers have been observed to give signals when in a homoduplex configuration (data not shown). However, these signals are generally weaker than the weakest signals considered to be indicative of mismatch-specific binding, i.e. weaker than or equal to the C:C signal in Figure 1.

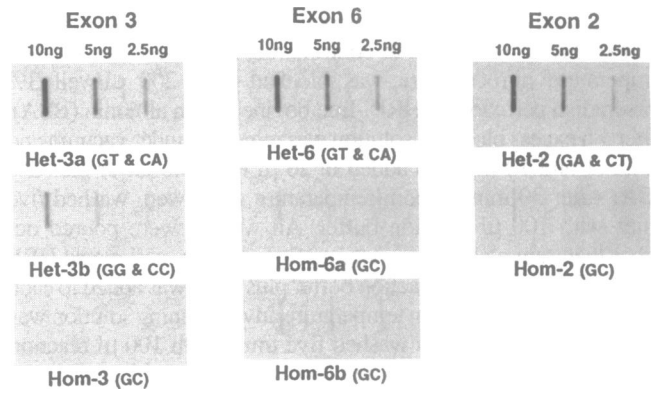

Figure 3. Detection of heterozygotes in the human glucokinase gene. Human genomic DNA from heterozygotes and homozygotes in the human glucokinase gene was amplified with primers specific to regions of exons 2, 3 and 6 (see Materials and Methods). Annealed PCR products were used in assays with immobilized MutS as described. Only DNA from heterozygotes should contain mismatches as indicated. Data from exons 3 and 6 are from a single experiment. Data from exon 2 are from a separate experiment. PCR mixture (100 µl): 10 mM Tris–HCl, pH 8.3, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 1.0 µM primers, 200 ng template DNA, 2.5 U Taq polymerase (Boehringer Mannheim). Thirty cycles: denaturation, 1 min at 94°C; annealing, 1 min, three cycles at 62°C, three cycles at 60°C, three cycles at 58°C, three cycles at 56°C, 18 cycles at 54°C; extension, 2 min at 72°C. Primers were removed with QIAquick spin columns (Qiagen). PCR products were eluted in 10 mM Tris–HCl, pH 8.0, adjusted to 0.1 M NaCl. Denaturation was at 100°C for 3 min. Annealing was for 1 h at 55°C, 4 min at 75°C and 30 min at 55°C. DNA was stored at 4°C until use. Exposure time 30 s.

The failure to bind C:C mismatches to a significant extent does not diminish the utility of this method for mutation detection, since every wild-type/mutant pairing gives rise to two different mismatches (e.g. G:G and C:C). G:G mismatches give strong signals.

Mutations in the human glucokinase gene are responsible for non-insulin-dependent diabetes (13). Regions of three glucokinase exons were PCR amplified from human genomic DNA known to be heterozygous for mutations in those regions and from human genomic DNA known or presumed to be homozygous for the wild-type sequence in those regions. In each case one of the primers contained a 5′-biotin, allowing detection by chemiluminescence. The same primers were used to amplify both heterozygous and homozygous genomic DNAs and the amplifications were performed simultaneously. Estimates of DNA quantities in the PCR products were obtained by polyacrylamide gel electrophoresis. (The DNA quantities shown in the figures are approximate. For the immobilized mismatch binding protein assay to produce accurate results it is sufficient to establish accurate relative quantities for homozygote and heterozygote comparisons. Any method capable of accurately obtaining such relative quantitation of the PCR products would be equally suitable.) The DNAs were denatured, by heating, allowed to re-anneal and tested for the presence of mismatches, i.e. heterozygotes, by testing their binding in an immobilized mismatch binding protein assay utilizing *E.coli* MutS.

The results are presented in Figure 3. In each case heterozygotes can be clearly distinguished from homozygotes. The actual ratios of mismatch-containing DNA binding to mismatch-free DNA binding (i.e. heteroduplex binding to homoduplex binding) are approximately twice the apparent ratios seen in Figure 3, since the

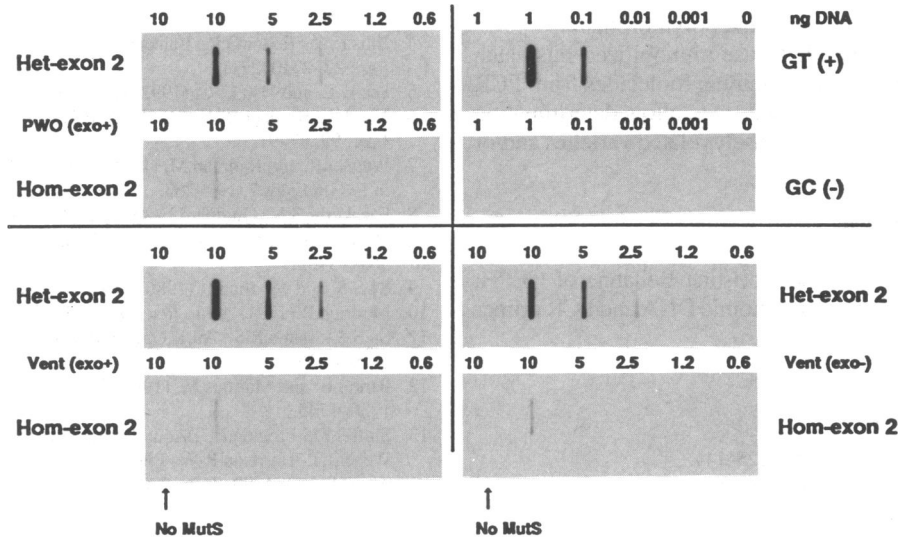

Figure 4. Comparison of products of different PCR polymerases using the immobilized MutS assay. DNA was amplified from human genomic DNAs heterozygous or homozygous in exon 2 of the glucokinase gene (see Materials and Methods). All data are from a single experiment. PCR mixtures (100µl): 0.25 mM dNTPs, 0.2 µM primer 1, 0.2 µM primer 2, 200 ng template DNA. PWO polymerase (Boehringer Mannheim): 10 mM Tris–HCl, pH 8.85, 25 mM KCl, 5 mM $(NH_2)_2SO_4$, 6 mM $MgSO_4$, 5 U DNA polymerase. Vent polymerase and Vent + exonuclease: 20 mM Tris–HCl, pH 8.8, 10 mM KCl, 2 mM $MgSO_4$, 10 mM $(NH_2)_2SO_4$, 0.1% Triton X-100, 2 U DNA polymerase. Thirty cycles: denaturation, 1 min at 94°C; annealing, 1 min, three cycles at 64°C, three cycles at 62°C, three cycles at 60°C, three cycles at 58°C, three cycles at 56°C, 15 cycles at 54°C; extension, 2 min at 72°C. Primers were removed with QIAquick spin columns (Qiagen). PCR products were eluted in 10 mM Tris–HCl, pH 8.0, adjusted to 0.1 M NaCl. Denaturation was at 100°C for 2 min. Annealing was for 1 h at 55°C, 4 min at 75°C and 30 min at 55°C. Exposure time 30 s.

heterozygote samples were randomly annealed. Therefore, half the molecules will be heteroduplexes and half will be homoduplexes. The strength of the heterozygote signal appears to be mismatch-dependent. In the case of exon 3, where two different mismatch pairs were studied, a strong signal is observed when the mismatches formed are G:T and C:A (Het-3a), whereas a somewhat weaker signal is observed with G:G and C:C mismatches (Het-3b), presumably due to the fact that only G:G mismatches are detected. The intermediate strength signal observed with the exon 2 fragment (Het-2) may reflect mismatch specificity, i.e. G:A and C:T mismatches appear to be somewhat less well recognized than G:T and A:C mismatches. However, the signal may also be somewhat lower because the molar concentration of mismatches is lower in the exon 2 fragment experiment than in the exon 3 fragment experiment, i.e. equal quantities of DNA were used and the fragments differ in length (230 versus 150 bp, respectively).

There is significantly increased binding of homoduplex DNA in these experiments relative to those with 30mer oligonucleotides (Fig. 2). It may be that the biotinylated primers occasionally initiate replication at sites other than the selected site. These fragments would be labeled and might be bound by immobilized MutS, either because they form mismatches when annealed with the genomic DNA from the homologous chromosome or because they form some secondary structure with mismatches. Alternatively, the homoduplex binding may be the result of polymerase errors or DNA damage occurring during amplification. Polymerase errors would be expected to occur relatively randomly throughout the amplified fragment, such that they would not be detectable by sequencing, but the cumulative effect of such errors could be to produce a sizable fraction of PCR products with some error. These would generally produce mismatches when denatured and annealed and thus contribute to positive signals in the immobilized mismatch binding protein assay. However, when the exon 2 fragment is amplified by four different polymerases, some of which have increased fidelity of replication and should, therefore, have a reduced rate of production of error-containing fragments, the ratio of heteroduplex to homoduplex binding does not change significantly (Figs 3 and 4).

The results presented here are concerned only with the detection of heterozygous mutations. The detection of homozygous mutations can easily be accomplished by adding known homozygous DNA to the test DNA before denaturation and annealing, either before or after amplification. Thus the use of immobilized mismatch binding protein assays for mismatch, mutation, heterozygosity or polymorphism detection involving single base substitutions and small additions or deletions seems to be limited only by the need to provide substrates free of labeled DNA with random mismatches, as discussed above. Immobilized mismatch binding protein provides a simple, accurate and easy to automate system for the following.

(i) Diagnostic screening for any disease causing mutation (or mutations), including single base substitutions and small additions or deletions, for which the sequence and location of the mutation(s) are known. It is possible to detect both carriers (heterozygotes) and affected patients (homozygotes) and to distinguish between them.

(ii) Rapid and large scale screening of human (or other) genomic DNA for single base change or small addition/deletion polymorphisms. The ease and speed of the system make it

possible to screen large numbers of individuals and to construct high resolution maps based on genomic polymorphism.

In addition, it may be possible to use immobilized mismatch binding protein to remove error-containing molecules from PCR samples, to bind heterozygous sequences to allow determination of identity by descent and to study closely related varieties and/or species to characterize biodiversity.

ACKNOWLEDGEMENTS

We thank Drs Nadine Cohen and Christine Bellanné of CEPH, Paris, France, for providing human genomic DNA and PCR primer sequences.

REFERENCES

1 Cotton,R.G.H. (1993) *Mutat. Res.*, **285**, 125–144.
2 Grompe,M. (1993) *Nature Genet.*, **5**, 111–116.
3 Lishanski,A., Ostrander,E.A. and Rine,J. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 2674–2678.
4 Ellis,L.A., Taylor,G.R., Banks,R. and Baumberg,S. (1994) *Nucleic Acids Res.*, **22**, 2710–2711.
5 Lu,A.-L. and Hsu,L.-C. (1992) *Genomics*, **14**, 249–255.
6 Youil,R., Kemper,B.W. and Cotten,R.G.H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 87–91.
7 Wagner,R. and Radman,M. (1995) In *Methods: A Companion to Methods in Enzymology*, **7**, 199–203.
8 Fazakerly,G.V., Quidnard,E., Woisard,A., Guschlbauer,W., van der Marel,G.A., van Boom,J.H., Jones,M. and Radman,M. (1986) *EMBO J.*, **5**, 3697–3703.
9 Su,S.S. and Modrich,P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5057–5061.
10 Modrich,P. (1991) *Annu. Rev. Genet.*, **25**, 229–253.
11 Su,S.S., Lahue,R.S., Au,K.G. and Modrich,P. (1988) *J. Biol. Chem.*, **263**, 6829–6835.
12 Parker,B. and Marinus,M. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1730–1735.
13 Stoffel,M., Frogule,P., Takeda,J., Zoulai,H., Vionnet,N., Nishi,S., Weber,I.T., Harrison,R.W., Pilkis,S.J., Lesage,S. *et al.* (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 7698–7702.