Volume 15 Number 23 1987 Nucleic Acids Research

A novel missense mutation in the factor VIII gene identified by analysis of amplified hemophilia DNA sequences

Barbara Levinson, Robert Janco[1], John Phillips III[1] and Jane Gitschier

The Howard Hughes Medical Institute and Department of Medicine, University of California, San Francisco, CA 94143 and [1]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

Received August 25, 1987; Revised and Accepted October 29, 1987

ABSTRACT

To date the only point mutations demonstrated to cause hemophilia are C to T transitions in *Taq*I sites. These were detected by screening Southern blots with cloned factor VIII probes. During the development of improved methods for detecting and analyzing mutations in genomic DNA, a novel G to C transversion mutation has been identified. This rare transversion results in a missense mutation, with proline being substituted for arginine in one of the active domains of the factor VIII molecule. The results suggest that the improved methods will be useful for detecting mutations in hemophilia as well as in other genetic disorders. In this method, specific DNA sequences in genomic DNA are amplified using oligonucleotide primers and a heat-resistant DNA polymerase. Mutations are detected and localized in the amplified samples by RNase A cleavage, and the altered region is then sequenced.

INTRODUCTION

Hemophila A, an X-linked bleeding disorder that affects one male in 5000 worldwide, is caused by a defect in blood coagulation factor VIII. Because the disease shows a wide range of clinical severity, it is thought to result from a heterogeneous collection of mutations in the factor VIII gene. In addition, the high incidence of non-familial cases reflects the introduction of spontaneous mutations into the population and suggests that hemophilia might provide an excellent model for the study of mutagenesis in man (1).

With the cloning of the factor VIII gene, it is now possible to determine the molecular basis of hemophilia (2). To date, deletions and *Taq*I restriction site mutations have been found by Southern blotting in 10% of hemophilia patients studied (3-7). In every TaqI site mutation examined, a C to T transition has occurred at the CpG dinucleotide in the recognition sequence (TCGA). These sites appear to be hot spots for mutation since multiple C to T transitions have occurred independently at each position resulting in either a stop codon (TGA) or more rarely a glutamine missense mutation (CAA). The cytosine at CpG is known to be methylated and spontaneous deamination to thymine results in a C to T transition (8-10).

Other mutations in the factor VIII gene must result in hemophilia, but the large size of the gene (186 kb) (2) and inaccessible mRNA (11) have prevented their identification. We report here the application and improvement of two recently described techniques, DNA amplification (12) and RNase cleavage (13) to detect and localize new mutations in the factor VIII coding region.

In the RNase cleavage procedure developed by Myers et al. (13), a radiolabeled RNA copy of a wild-type cDNA template is hybridized to denatured DNA or RNA. The resulting hybrid is digested with RNase A, and the protected, labeled RNA fragments are analyzed on a denaturing gel. Under optimal conditions, using both sense and antisense probes, approximately 50-70% of single-base mismatches can be identified with RNase A (13). Although RNase cleavage has been used to detect mutations in mRNA for other proteins (14,15), the low abundance of factor VIII mRNA precludes the use of this approach (11). Myers et al. (13) also used RNase cleavage to detect a previously characterized β-globin gene mutation in genomic DNA directly. We initially attempted a similar approach for hemophilia A using an RNA probe containing 12 of the 26 exons, however a variable and weak signal discouraged us from using this method as a routine screen. To augment the RNase cleavage assay, we have chosen to amplify coding sequences in the factor VIII gene by oligonucleotide primer-directed DNA synthesis (12) using the heat-stable DNA polymerase from Thermus aquaticus (16,17). Amplification has the added benefit that mutations located by RNase cleavage can be easily analyzed by cloning and sequencing the amplified DNA (18), or as has been recently described, by direct sequencing (19). Using these techniques, a novel G to C transversion has been identified resulting in an arginine to proline substitution in one of the active domains of the factor VIII molecule.

MATERIALS AND METHODS

Oligonucleotide Synthesis

Oligonucleotide primers and probes were synthesized on an Applied Biosystems 380A DNA Synthesizer by the methoxy phosphoramidite method (20).

Genomic DNA Isolation

DNA was prepared from a cell line with 4X chromosomes and from peripheral blood leukocytes as previously described (21).

RNase Cleavage

Sense and antisense RNA probes were transcribed from linearized plasmid pF8.11 using a Promega Biotec transcription kit. Plasmid pF8.11 is a pGEM-3 (Promega Biotec) derived vector containing factor VIII cDNA sequences for exons 20-25 (1050 bp, *Hpa*II/*Pvu*II fragment). Probes of 1 kb were uniformly labeled with ^{32}P using 50 μCi [^{32}P]UTP (100 Ci/mmole, 50 μM, Amersham), 1 μg template DNA, and either bacteriophage SP6 or T7 RNA polymerase. For the RNase A cleavage, approximately 3-5 ng of amplified DNA was denatured and hybridized to 1.5 x 10^5 cpm probe as described by Myers et al. (13) for cloned DNA. After hybridization, 350 μl containing 35 μg/ml RNase A (Sigma) and 1 μg/ml RNase T1 (Sigma, included to degrade small labeled oligonucleotides) in 10 mM Tris-HCl, pH 7.6, 1 mM EDTA, 300 mM NaCl, 100 mM LiCl was added and incubated for 30 min at 20°C. The samples were run on 6% polyacrylamide-urea gels. The autoradiogram was exposed for 3 h. at -70°C.

TABLE I. FACTOR VIII OLIGONUCLEOTIDE PRIMERS AND PROBES

	exon 22	exon 23
PRIMERS	14.1 5'GTGGATCTGTTGGCACCAATG	13.1 5'GTCTTCTTTGGCAATGTGGATTC
	14.2 5'CATTAAGCTTCCAGTGGAATTTCC	13.5 5'CTACTTAAGTCACAGCCCATCAACTC
PROBES		
wild-type	14.3 5'GACTTATCGAGGAAA	13.6 5'TATTGCTCGATACAT
nonsense	14.4 5'GACTTATTGAGGAAA	13.7 5'TATTGCTTGATACAT
missense	14.5 5'GACTTATCAAGGAAA	13.8 5'TATTGCTCAATACAT
	14.6 5'GACTTATCCAGGAAA	

Primer sequences were chosen to match the sequences at the ends of exons 22 and 23 (2, 21) but were modified slightly at the 5' ends to include restriction enzyme sites for cloning. Probes are 15 nt and sequences corresponding to TaqI sites are underlined. Probe 14.6 was used to confirm the V99 sequence.

DNA Amplification

DNA sequences were amplified as described by Kogan et al. (17) in a 100 μl-reaction volume containing 1 μg genomic DNA, 1.5 mM each dATP, dGTP, dCTP, and dTTP, 50 pmoles each primer (Table I), and 10% dimethylsulfoxide, in 1X reaction buffer [16.6 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8 at 25°C), 6.7 mM $MgCl_2$, 10 mM B-mercaptoethanol, 6.7 μM EDTA, and 170 μg/ml bovine serum albumin]. Samples were heated 5 min. at 95°C to denature the DNA and briefly cooled to room temperature. One unit *Taq* polymerase (New England Biolabs) was added to each sample and transfered to 65°C for 1 min. 40 sec. for primer-directed DNA synthesis of a 200-bp region. Subsequent rounds of 30-sec. denaturation, 30 sec. cool-down at room temperature and 65°C synthesis continued for 30 cycles, with 1 unit *Taq* polymerase added at rounds 10 and 20. Approximately 100 ng target DNA was synthesized, resulting in a 2 million-fold amplification.

Slot Blots and Oligonucleotide Hybridization

Slot blots of amplified DNA were prepared by denaturing 10% of the amplified DNA sample (10 μl) in 200 μl containing 0.4 M NaOH, 25 mM EDTA. The samples were heated for 2 min. at 95°C, chilled on ice, then loaded on a Zetaprobe nylon filter (Schleicher and Schuell) that had been pre-wetted with 20X SSPE (200 mM NaH_2PO_4, 3 M NaCl, 20 mM EDTA) and placed in a slot-blot manifold. The filters were washed with 20X SSPE, baked at 80°C for 1 h., pre-hybridized in 5X SSPE, 5X Denhardt's, 0.5% SDS at 37°C for 1 h., and hybridized overnight at 37°C with the appropriate oligonucleotide probe end-labeled with [γ^{32}P]ATP (Amersham) added to the hybridization solution. Blots were washed for 30 min. at room temperature in 6X SSC and twice for 20 min at 49°C in 3 M tetramethylammonium chloride solution as described by Wood et al. (22).

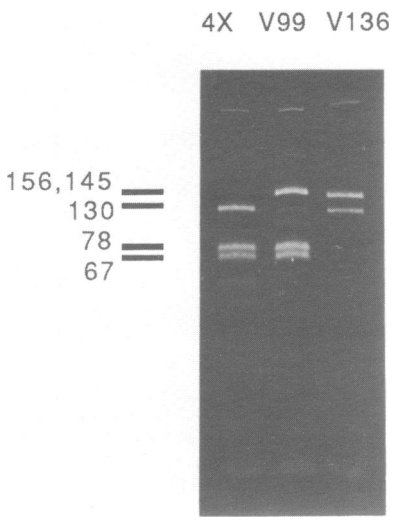

Figure 1 Amplification of exon 22 and 23 sequences with *Taq* polymerase and digestion with *Taq*I. Genomic DNA samples isolated from a cell line with four X chromosomes (4X, ref. 3) and from peripheral blood leukocytes of patients V99 and V136 were amplified using *Taq* polymerase (16,17). The oligonucleotide primers 14.1 and 14.2 direct synthesis at exon 22 and 13.1 and 13.5 at exon 23, resulting in unique amplified fragments of sizes 156 and 145 bp, respectively. 20 μl of the amplification reaction was digested with *Taq*I, and the digestion products were run on a 10% polyacrylamide mini-gel. *Taq*I digestion of exon 22 yields two fragments of 130 and 26 bp and that of exon 23 two fragments of 78 and 67 bp.

DNA Sequencing

The exon 22 primers contained *Sau*3A and *Hin*dIII restriction sites, enabling the exon 22 amplified sequences to be readily cloned into the appropriately digested M13mp19 vector and sequenced by the dideoxy chain termination method (23).

RESULTS

Two hemophilia mutations of known location but unknown sequence were chosen to test our ability to identify mutations by amplification and RNase A cleavage. These mutations were discovered as aberrant *Taq*I restriction sites in exons 22 and 23 of hemophilia DNA samples V99 and V136, respectively (data not shown). Exon 22 and 23 sequences from the two hemophilia DNAs and 4X chromosome control DNA were amplified simultaneously using the two sets of oligonucleotide primers listed in Table I. The amplification was conducted at 65°C, using the heat-stable DNA polymerase isolated from Thermus aquaticus (*Taq* polymerase; 16,17). At this elevated temperature, the hybridization of a pair of primers to a genomic template is sufficiently stringent so that only the appropriate target sequence is amplified resulting in a discrete band. The fidelity of the amplification was confirmed by analysis of the amplified

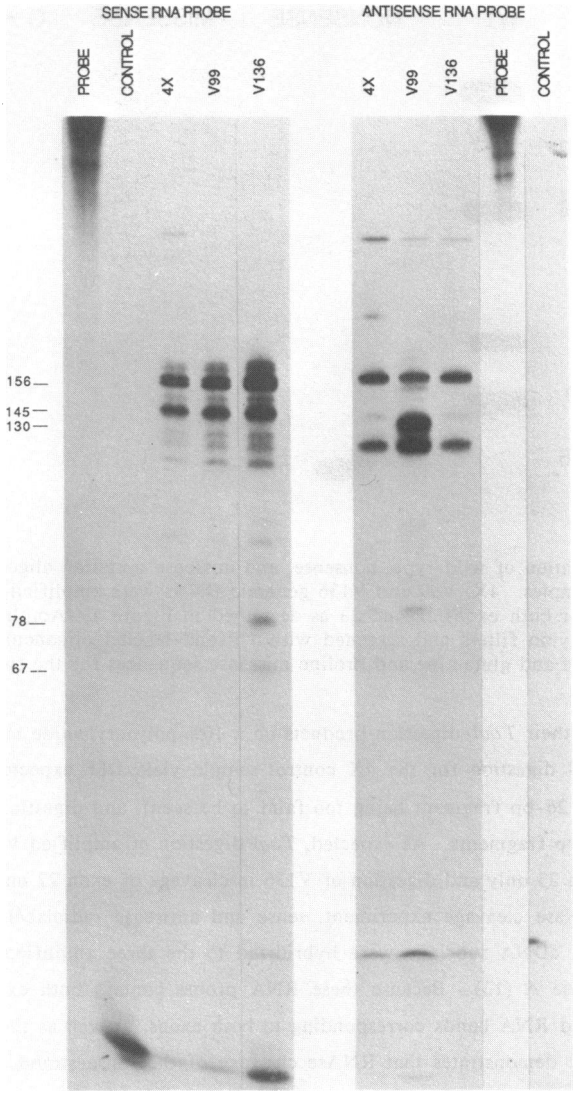

Figure 2 RNase cleavage of sense and antisense RNA probes hybridized to amplified genomic sequences. The three amplified DNA samples, 4X, V99, and V136, were analyzed by the RNase cleavage reaction with ^{32}P-labeled sense and antisense RNA probes that contain factor VIII exons 22 and 23. Exon 22 and 23 amplified DNA protect 156-nt and 145-nt RNA molecules, respectively, if perfect homology exists. An autoradiogram of the RNase digestion is shown and the sizes of the resulting products are indicated. The sizes of the protected bands are approximate due to mismatches between the probe and DNA at the primer sequence and possible "breathing" between the RNA:DNA hybrid (13). Cleavage at these mismatched sites is dependent on whether a sense or antisense probe is used for the hybridization due to the specificity of RNase A for certain mismatches. Also shown are the probes and the control lanes with RNase-digested probes.

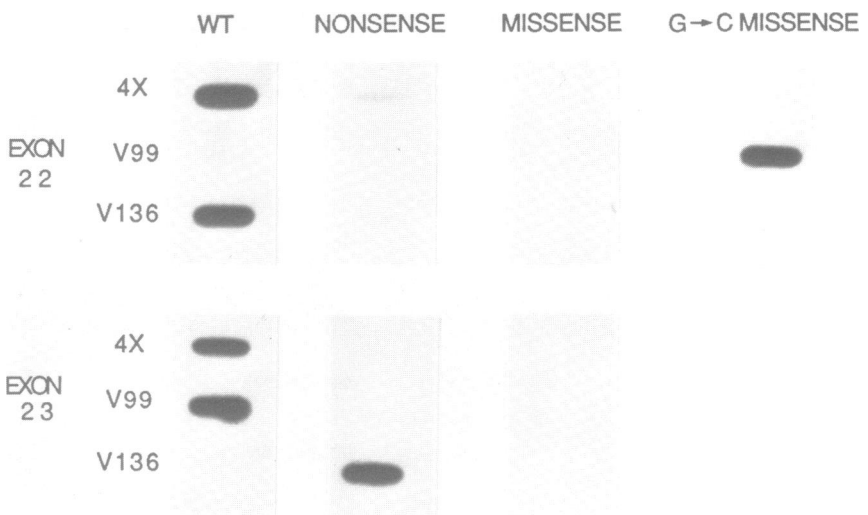

Figure 3 Hybridization of wild-type, nonsense, and missense mutation oligonucleotide probes to amplified DNA samples. 4X, V99 and V136 genomic DNAs were amplified simultaneously with primers specific for both exons 22 and 23 as described in Figure 1. Amplified sequences were slot-blotted onto nylon filters and screened with ^{32}P-end-labeled oligonculeotides representing wild-type, nonsense and glutamine and proline missense sequences for the two exons (Table 1).

DNA samples and their *Taq*I-digestion products on a 10% polyacrylamide mini-gel, as shown in Figure 1. A *Taq*I digestion for the 4X control sample yields the expected 130-bp fragment from exon 22 (the 26-bp fragment being too faint to be seen), and digestion of exon 23 results in 67-bp and 78-bp fragments. As expected, *Taq*I digestion of amplified V99 sequences results in cleavage of exon 23 only and digestion of V136 in cleavage of exon 22 only.

For the RNase cleavage experiment, sense and antisense radiolabeled RNA transcripts from a factor VIII cDNA subclone were hybridized to the three amplified DNA samples and digested with RNase A (13). Because these RNA probes contain both exon 22 and exon 23 sequences, protected RNA bands corresponding to both exons, as well as the cleavage products, are seen. Figure 2 demonstrates that RNAse cleavage of the sense-strand probe hybridized to V136 results in the appearance of specific 78- and 67-nt bands. In this case, partial cleavage of the exon 23 band to the predicted sizes has occurred. The 67- and 78-nt bands do not appear in the 4X or V99 lanes which contain normal exon 23 sequences, even with a twenty-fold longer exposure. Using the antisense RNA probe, exon 23 is not cleaved in any of the three DNA samples. In contrast, the mutated exon 22 in V99 appears normal when hybridized to the sense RNA probe, but the anti-sense probe is cleaved to a band of predicted size (130 nt).

Because *Taq*I-site mutations in the factor VIII gene routinely occur by a C to T transition, a series of oligonucleotide probes corresponding to the predicted nonsense and

missense mutations were synthesized to rapidly analyze *Taq*I mutations. To determine the sequence in V99 and V136, the amplified DNA samples were hybridized to labeled oligonucleotides that match the wild-type and the predicted nonsense and missense sequences in exons 22 and 23 (total of six probes, Table I). The V136 sample hybridized to the exon 23 nonsense probe proving that the mutation causes a TGA stop codon by a C to T transition. However, the V99 sample did not hybridize to any of the three exon 22 probes and therefore appears to be different from any previously described mutation.

To determine the sequence of the exon 22 *Taq*I site mutation in V99, the amplified DNAs from 4X and V99 were cloned into an M13 vector using the artificial restriction enzyme sites incorporated into the 5' ends of the oligonucleotide primers. The amplified 4X DNA was found to have the normal exon 22 sequence. Sequencing the V99 DNA revealed an unexpected G to C transversion altering the *Taq*I site sequence to TCCA. The new sequence in V99 was confirmed by hybridization to two independently amplified samples with an oligonucleotide probe homologous to the TCCA missense mutation (probe 14.6, Table 1, Fig. 3). This missense mutation results in a change from arginine to proline at position 2116 (21) in the factor VIII protein. The V99 and V136 mutations result in RNA:DNA single-base mismatches (C:C and C:A respectively) which are known to be digested with RNase A (13) and are therefore consistent with the RNase cleavage results.

DISCUSSION

G to C transversions are rarely observed in either bacteria (24) or mammalian genes (25) and although the mechanism for this transversion is unknown, it may result from DNA replication error (26). In contrast, C to T transitions are common and the mechanism is known to be deamination of methyl cytosine, resulting in an unrepaired thymine (8-10). This transversion in V99 is the first to be described in the factor VIII gene. Although this missense mutation may be neutral and not cause hemophilia in patient V99, this substitution of proline for arginine is likely to result in a destabilization of the factor VIII molecule (27). This mutation occurs in the CI domain of the factor VIII protein, known to be a region important for biological activity (28). Both patients V99 and V136 suffer from severe hemophilia, having less than 1% normal factor VIII clotting activity; the severity is consistent with both the destabilizing effects of a proline missense mutation and with a truncated protein molecule.

The combined use of RNase cleavage and amplification of genomic DNA sequences provides a means to localize mutations which cannot be detected by Southern blot analysis. An additional advantage of specific oligonucleotide primer-directed amplification is the ability to then rapidly sequence the regions of interest. Futhermore, *Taq* polymerase allows simultaneous and specific priming at multiple sites and generation of amplified fragments as long as 2000 bp (data not shown). This method is applicable to hemophilia as well as to other genetic disorders. Currently we are screening hemophilia DNA samples by these methods for mutations in regions

of the factor VIII gene that are required for biological activity, including the promoter region and obligate proteolytic cleavage sites.

ACKNOWLEDGEMENTS

We are indebted to Richard Myers for many helpful discussions and encouragement regarding the RNase A cleavage technique. We thank Scott Kogan who developed the amplification protocol with *Taq* polymerase, Marie Doherty for the preparation of oligonucleotides, and Pamela Orlando for DNA preparation. Drs. Philip Chance and Gordon Bray kindly provided patient samples after obtaining informed consent, and we thank the patients for their cooperation. We wish to thank Y.W. Kan and members of his laboratory for their interest and support of this project. Special thanks to Bernard Morley, Kevin Shannon, and Jennifer Gampell for many helpful suggestions in the preparation of this manuscript. Part of this work was funded by NIH grant #HL35097 to R. Janco. J. Gitschier is an Assistant Investigator of the Howard Hughes Medical Institute.

REFERENCES

1. Haldane,J.B.S. (1935) J. Genet. **31**, 317-326.
2. Gitschier,J., Wood,W.I., Goralka,T.M., Wion,K.L., Chen,E.Y., Eaton,D.H., Vehar,G.H., Capon,D.J.and Lawn,R.M. (1984) Nature **312**, 326-330.
3. Gitschier,J., Wood,W.I., Tuddenham,E.G.D., Shuman,M.A., Goralka,T.M., Chen,E.Y. and Lawn,R.M. (1985) Nature **315**, 427-430.
4. Antonarakis,S.E., Waber,P.G., Kittur,S.D., Patel,A.S., Kazazian,H.H.,Jr., Mellis,M.A., Counts,R.B., Stamatoyannopoulos,G., Bowie,E.J.W., Fass,D.N., Pittman,D.D, Wozney,J.M. and Toole,J.J. (1985) N. Engl. J. Med. **313**, 842-848.
5. Youssoufian,H., Antonarakis,S.E., Aronis,S., Tsiftis,G., Phillips,D.G. and Kazazian,H.H., Jr. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 3772-3776.
6. Youssoufian,H., Kazazian,H.H., Phillips,D.G., Aronis,S., Tsiftis,G., Brown,V.A. and Antonarakis,S.E. (1986) Nature **324**, 380-382.
7. Gitschier,J., Wood,W.I., Shuman,M.A. and Lawn,R.M. (1986) Science **232**, 1415-1416.
8. Coulondre,C., Miller,J.H., Farabaugh,P.J. and Gilbert,W. (1978) Nature **274**, 775-780.
9. Duncan,B.K. and Miller,J.H. (1980) Nature **287**, 560-561.
10. Bird,A.P. (1980) Nucl. Acids. Res. **8**, 1499-1504.
11. Wion,K.L., Kelly,D., Summerfield,J.A., Tuddenham,E.G.D. and Lawn,R.M. (1985) Nature **317**, 726-729.
12. Saiki,R.A., Scharf,S, Faloona,F., Mullis,K.B., Horn,G.T., Erlich,H.A. and Arnheim,N. (1985) Science **230**, 1350-1354.
13. Myers,R.M., Larin,Z. and Maniatis,T. (1985) Science **230**, 1242-1246.
14. Winter,E., Yamamoto,F., Almoguera,C. and Perucho,M. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 7575-7579.
15. Gibbs,R.A. and Caskey,C.T. (1987) Science **236**, 303-305.
16. Chien,A., Edgar,D.B. and Trela,J.M. (1976) J. Bact. **127**, 1550-1557.
17. Kogan,S., Doherty,M. and Gitschier,J. (1987) N. Engl. J. Med. In press.
18. Scharf,S.J., Horn,G.T. and Erlich,H.A. (1986) Science **233**, 1076-1078.
19. Wrischnik,L., Iiguchi,R.G., Stoneking,M., Erlich,H.A., Arnheim,N. and Wilson,A.C. (1987) Nucl. Acids. Res. **15**, 529-543.
20. Beaucage,S. L. and Caruthers,M.H. (1981) Tetrahedron Lett. **22**, 1859-1862.
21. Wood,W.I., Capon,D.J., Simonsen,C.C., Eaton,D.L., Gitschier,J., Keyt,B., Seeburg,P.H., Smith,D.H., Hollingshead,P., Wion,K.L., Delward,E., Tuddenham,E.G.D., Vehar,G.A. and Lawn,R.M. (1984) Nature **312**, 330-337.

22. Wood,W.I., Gitschier,J., Lasky,L.A. and Lawn,R.M. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1585-1588.
23. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5467.
24. Miller,J.C. (1983) Ann. Rev. Genet. **17**, 215-238.
25. Orkin,S.H., Kazazian,H.H.,Jr., Antonarakis,S.E., Goff,S.C., Boehm,C.D., Sexton,J.P., Waber,P.G. and Giardina,P.J.V. (1982) Nature **296**, 627-631.
26. Glickman,B.W. (1985) In Scarfe,J., Leach,D., Galizzi,A. (eds) Genetics of Bacteria, pp. 41-47. Academic Press, London.
27. Chou,P.Y. and Fasman,G.D. (1974) Biochemistry **13**, 222-245.
28. Eaton, D., Rodriguez, H., Vehar, G.A. (1986) Biochemistry **25**, 505-512.