# MBP Fusion Protein with a Viral Protease Cleavage Site: One-Step Cleavage/ Purification of Insoluble Proteins

*BioTechniques 30:1194-1198 (June 2001)*

Production of recombinant proteins as affinity-tagged fusions in *E. coli* has become standard laboratory procedure. However, the choice of expression vector and, hence, of the purification scheme and protease used for fusion-cleavage, is case-dependent. Among the factors affecting this choice are expression level and final yield, solubility of the fusion, specificity of the protease, specificity and affinity of the purification resin, and the associated time and material costs.

For example, we are interested in the expression of milligram quantities of the C-terminal region of Par-4, an apoptosis-associated protein (14,15). Attempts to express the fragment as a GST- (16) or His-tagged (2,6,13) fusion produce largely insoluble protein, which is difficult to refold after solubilization in denaturants. In addition, thrombin digestion to separate the fusion tag results in proteolytic degradation of the Par-4 fragment. The use of a commercially available maltose-binding protein (MBP) fusion system (pMAL-c2X; New England Biolabs, Beverly, MA, USA) (3,9) increases solubility. However, Factor Xa digestion (10,11) to remove the MBP fusion is costly and results in unacceptable levels of secondary cleavage within the Par-4 fragment. The secondary cleavages are due, at least in part, to the propensity of Factor Xa to nonspecifically cleave unfolded or partially folded polypeptides (4,12). A method that has been developed to suppress secondary cleavages by Factor Xa through the reversible acylation of lysine side chains (18) may not be generally applicable to other proteins, and it may be preferable to avoid covalent modification of the substrate protein altogether. In addition, the subsequent removal of Factor Xa requires the design and use of a second purification step. Finally, the propensity of MBP or MBP fusions to prematurely elute from amylose resin can decrease the final yield and hinder the complete separation of the protein product from MBP after fusion cleavage. Time-dependent leakage also renders attempts at fusion cleavage while bound to amylose resin inefficient.

To address these issues, we have modified an existing plasmid to create a novel expression vector that we call H-MBP-3C. This vector combines the favorable solubilization features of an MBP fusion with the use of a His tag for affinity purification (Figure 1). A human rhinovirus protease site [e.g., 3C protease from HRV-14 (1,17)] placed between the fusion partners provides highly specific fusion cleavage, even in the case of unfolded or partially folded proteins. A second construct was engineered to express HRV-14 3C protease as an MBP fusion with an N-terminal His tag (i.e., H-3Cpro). This construct can be used to produce 3C protease at low cost. After immobilizing the fusion protein of interest on metal affinity resin, the addition of H-3Cpro causes the protein to elute, and MBP and the protease remain bound to the resin. Any leakage of residual MBP or H-3Cpro can be easily removed by a quick additional step, resulting in a high yield of greater than 98% pure protein. Alternatively, the fusion cleavage can be performed with other sources of 3C protease such as GST-tagged 3C protease (PreScission protease; Amersham Pharmacia Biotech, Piscataway, NJ, USA), which subsequently can be removed with GSH resin.

The H-MBP-3C vector was constructed via a circuitous route for historical reasons. (*i*) A DNA duplex

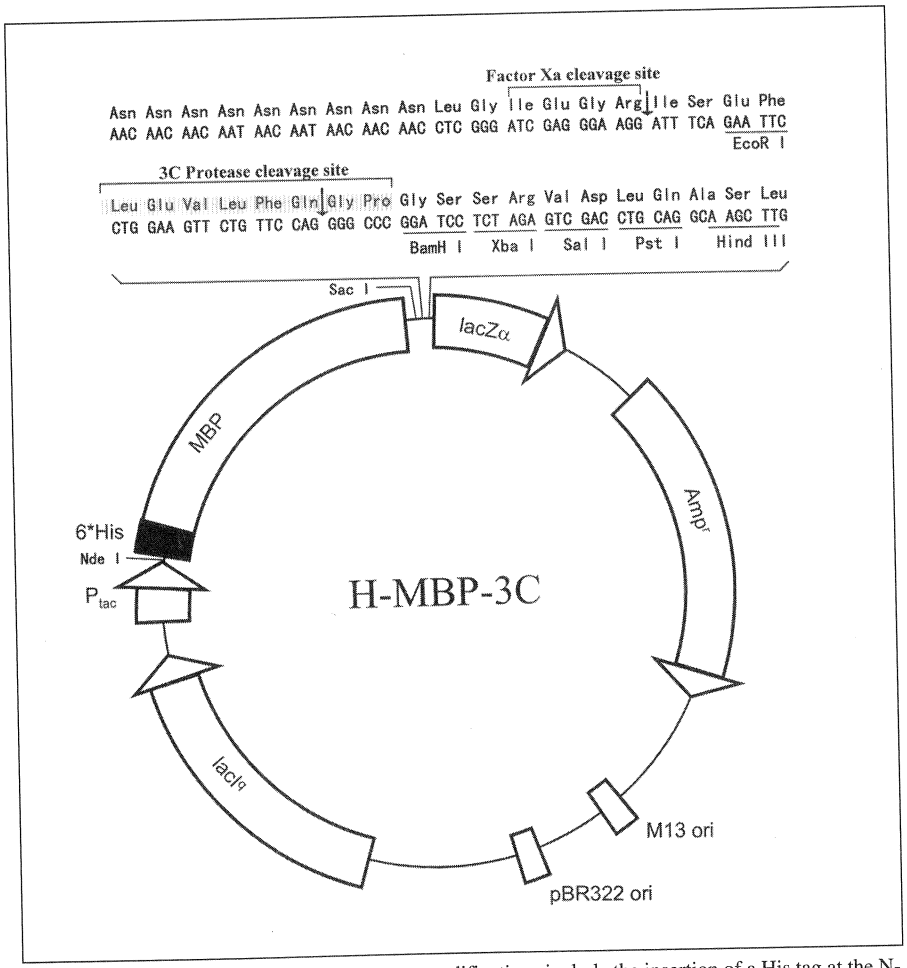


**Figure 1. H-MBP-3C vector map.** Important modifications include the insertion of a His tag at the N-terminus of MBP and a 3C protease site between the *Eco*RI and *Bam*HI sites of the MCS. The MBP fusion increases solubility, and the His tag and 3C site optimize purification and fusion cleavage.

(5′-AATTCCTGGAAGTTCTGTTCC-AGGGGCCCGp-3′, 5′-GATCCGGG-CCCCTGGAACAGAACTTCCAGG-p-3′) coding for an HRV-14 3C protease site (i.e., LEVLFQGP) with flanking *Eco*RI and *Bam*HI sites was inserted between the *Eco*RI and *Bam*HI sites in the MCS of the pMal-c2 vector. (*ii*) As part of another project, an unrelated DNA sequence was inserted between the MCS *Bam*HI and *Sal*I sites. This sequence was later removed, and the full MCS was restored through the use of PCR. (*iii*) PCR (primers: 5′-GGGAATTCCATATGCATCACC-ATCACCATCACATGAAAACTGAA-GAAGGTAAACTGG-3′, 5′-CGCCA-GGGTTTTCCCAGTCACGAC-3′) was used to amplify the above construct between the MBP N-terminus and a site 3′ to the MCS, with modification of the 5′ end to add an *Nde*I site and a six-His tag. This PCR product was then inserted between the *Nde*I and *Sal*I sites of the pMal-c2X vector. To ensure that only one copy of the 3C protease site had been inserted in the first step, an *Eco*RI digest of the final product was followed by gel purification, religation, retransformation, and sequencing. Figure 1 shows a map of the resulting construct.

To create an expression construct for the Par-4 C-terminal region, a PCR product coding for amino acids 286–332 of human Par-4 with flanking *Bam*HI and *Sal*I sites was inserted between the H-MBP-3C *Bam*HI and *Sal*I sites using standard procedures.

To create a construct expressing His-tagged HRV-14 3C protease, cDNA corresponding to human rhinovirus isotype 14 was used as a template for PCR. Primers were designed to extract DNA coding for amino acids 1–182 of 3C protease with the addition of a 5′ *Bam*HI site and a 3′ stop codon, followed by a *Sal*I site. After restriction cleavage, this DNA was inserted into the pMal-c2T vector (New England Biolabs), which is similar to pMAL-c2X, except the Factor Xa site is replaced by a thrombin site. The thrombin site and 3Cpro-containing region of this plasmid was then excised using *Sac*I and *Sal*I restriction enzymes and inserted between the *Sac*I and *Sal*I sites into the H-MBP-3C vector. Thus, the final construct produces 3C protease with an N-terminal six-His and MBP fusion, which may be removed by thrombin cleavage.

The H-MBP-3C Par-4 construct above was transformed into BL21 *E. coli* cells (Novagen, Madison, WI, USA). After approximately 24 h of growth at 37°C, a single colony was chosen to inoculate 50 mL ampicillin-containing LB media, which was agitated overnight at 37°C. Ten milliliters of this culture were taken to inoculate 1 L LB in a 2-L baffle flask, which was agitated at 37°C until $A_{600}$ reached approximately 0.6. At this point, isopropyl-β-D-thiogalactoside (IPTG) was added to a concentration of 0.4 mM to induce protein expression, and 4 h later, cells were harvested by centrifugation at 4200× *g* for 20 min and stored overnight at -20°C. The expression of H-3Cpro proceeded along similar lines in BL21 (DE3) codon-plus cells, except the growth temperature was reduced to 28°C after the addition of IPTG, and cells were harvested after 6–12 h. The DE3 insertion is not required here, but its presence did not appear to affect expression. However, codon-plus cells were required because of the presence of multiple, rare arginine codons (5,7,8) in the 3Cpro sequence.

The protease itself was purified using the following procedure (all of the steps were carried out at 4°C or on ice unless otherwise noted). After thawing, cells that had been pelleted from 500 mL culture were resuspended in 25 mL buffer A (40 mM sodium phosphate, pH 7.8, 200 mM NaCl, 5 mM BME, and 5% glycerol). Cells were lysed by three passages through a French press, and the lysate was centrifuged at 17 000× *g* for 30 min, passed through a 0.22-µm filter, and rotated for 2 h with 3 mL Talon cobalt-affinity resin, which had been pre-equilibrated with buffer A. Note that while nickel-affinity resin can be substituted, in general, we find more specific binding with cobalt resin. In addition, cobalt resin does not require a wash with a low imidazole concentration buffer before elution, and elution can take place with a 3× lower concentration of imidazole. The resin was then pelleted by centrifugation at 1500× *g* for 2 min, the supernatant was decanted, and the beads were washed six times with buffer A. To elute, the resulting 5 mL slurry was transferred to a 7-mL disposable column with end cap in place, and 0.75 mL buffer AI (i.e., buffer A with 1 M imidazole) was added. The column was covered and rotated for 20 min, after which we collected approximately 3 mL eluent containing 98% or purer H-3Cpro. To remove imidazole, this solution was twice dialyzed in 1 L buffer A or run through a size exclusion column (SE 100/17; Bio-Rad Laboratories, Hercules, CA, USA) with buffer A as the

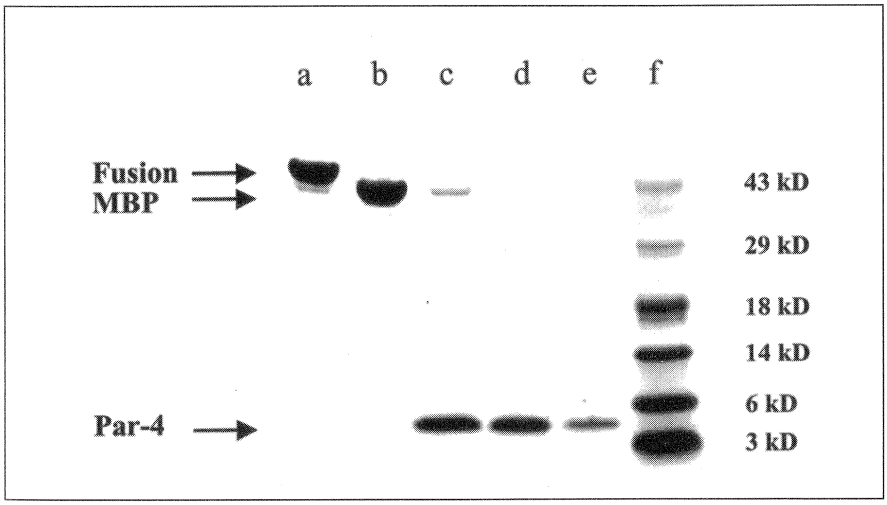


**Figure 2. Purification and cleavage.** Coomassie® Blue-stained SDS-PAGE. (a) Talon resin after incubation with soluble extract containing H-MBP-3C Par-4 and (b) Talon resin from (a) after incubation with a 1:100 ratio of H-3Cpro, transfer to a disposable column, and separation of eluent. Cleavage is greater than 99% complete. (c) Eluent. Elution of Par-4 is more than 99%. Less than 5% of the MBP present after cleavage elutes with the Par-4, and it can be further removed by (d) incubation with amylose resin or (e) passage through a 30-kDa cut-off filter.

running buffer. After adding ice-cold glycerol to 20%, the protease was divided into aliquots containing 0.2 mg protease and was stored at -20°C. Total fusion protein yield is typically 30 mg from a 500-mL culture.

The H-MBP-3C Par-4 fusion protein was purified using the following procedure. Cells were resuspended in 15 mL PBS buffer (140 mM NaCl and 12 mM NaP, pH 7.5), to which we added one EDTA-free, protease-inhibitor cocktail tablet (Roche Molecular Biochemicals, Indianapolis, IN, USA) and 1 mg/mL lysozyme. This mixture was incubated at 37°C for 15 min, sonicated, and centrifuged at 20 200× *g* for 40 min to remove debris. The supernatant was passed through a 0.45-µm filter and rotated for 2 h with 4 mL Talon cobalt-affinity resin, which had been pre-equilibrated with PBS. The resin was then washed twice with PBS, three times with PBS plus 1 M NaCl, and twice again with the original PBS. The yield is typically 30–50 mg purified H-MBP-3C Par-4 fusion protein/L culture.

Separation of the H-MBP tag and the Par-4 fragment (5 kDa) was achieved through the use of H-3Cpro. A total of 50 µg H-3Cpro (1/4 of an aliquot as prepared above) was added to 5 mg H-MBP-3C Par-4 bound to 0.5 mL Talon resin in PBS. This mixture was incubated and gently rotated for 12 h at 4°C. The resin and supernatant were then transferred to a 5-mL disposable column, and the Par-4 was eluted with PBS and then passed through a 0.22-µm filter. Residual MBP eluting with the Par-4 was removed by nutation with 0.8 mL amylose resin for 1 h at 4°C or by passing through a 3-kDa cutoff spin concentrator.

The problem of insolubility of the Par-4 fragment when expressed as a GST-fusion or with a six-His tag was largely overcome through the use of the pMAL-c2X vector to create an MBP fusion. The expression of the Par-4 fragment from the H-MBP-3C vector (Figure 1) offers the same solubility advantages, with the added benefit of replacing the Xa digest with a 3C digest. As shown in Figure 2, the fusion protein is more than 98% pure after a single binding step with cobalt resin (lane a). Incubation for 12 h with a 1:100 weight ratio of H-3Cpro to Par-4 fusion results in cleavage of more than 99% of the fusion protein. Note that a 1:400 weight ratio of 3C is sufficient to cleave more than 95% of the fusion protein in 12 h (data not shown). Approximately 95% of the MBP remains bound to the beads (lane b), while the Par-4 fragment elutes with the additional 5% of the MBP (lane c). Due to the presence of both a six-His tag and an MBP tag on the fusion protein, 1 h incubation with amylose resin (lane d) can be used to reduce residual MBP to approximately 1%. Alternative approaches, such as passage through a 30-kDa cutoff spin concentrator (lane e) can remove MBP to within detectable limits.

The procedures described here have been applied to a small, relatively insoluble, and partially folded polypeptide, which presented difficulties when expressed from several commercially available plasmids. The same procedures should also prove efficient and applicable to proteins with varying or unknown properties and sizes, such as is becoming common in proteomics applications. A 500-mL culture H-3Cpro produces an amount of protease sufficient to cleave approximately 3 g fusion protein to 99% completion or 12 g to 95%. If desired, higher concentrations of H-3Cpro can be used to perform the cleavage step in a shorter time since no secondary cleavages were observed with H-3Cpro/fusion protein ratios up to 1:1 (data not shown). After storing the protease as glycerol stock at -20°C for several months, we were unable to detect appreciable loss in proteolytic activity as determined by the ability to cleave fusion protein (data not shown).

## REFERENCES


1. **Cordingley, M.G., P.L. Callahan, V.V. Sardana, V.M. Garsky, and R.J. Colonno.** 1990. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. J. Biol. Chem. *265*:9062-9065.
2. **Crowe, J., H. Dobeli, R. Gentz, E. Hochuli, D. Stuber, and K. Henco.** 1994. 6xHis-Ni-NTA chromatography as a superior technique in recombinant protein expression/purification. Methods Mol. Biol. *31*:371-387.
3. **di Guan, C., P. Li, P.D. Riggs, and H. Inouye.** 1988. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. Gene *67*:21-30.
4. **Doskeland, A.P., A. Martinez, P.M. Knappskog, and T. Flatmark.** 1996. Phosphorylation of recombinant human phenylalanine hydroxylase: effect on catalytic activity, substrate activation and protection against non-specific cleavage of the fusion protein by restriction protease. Biochem. J. *313*:409-414.
5. **Forman, M.D., R.F. Stack, P.S. Masters, C.R. Hauer, and S.M. Baxter**. 1998. High level, context dependent misincorporation of lysine for arginine in *Saccharomyces cerevisiae* a1 homeodomain expressed in *Escherichia coli*. Protein Sci. *7*:500-503.
6. **Hochuli, E., H. Dobeli, and A. Schacher.** 1987. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J. Chromatogr. *411*:177-184.
7. **Hu, X., Q. Shi, T. Yang, and G. Jackowski.** 1996. Specific replacement of consecutive AGG codons results in high-level expression of human cardiac troponin T in *Escherichia coli*. Protein Expr. Purif. *7*:289-293.
8. **Kane, J.F.** 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr. Opin. Biotechnol. *6*:494-500.
9. **Maina, C.V., P.D. Riggs, A.G.D. Grandea, B.E. Slatko, L.S. Moran, J.A. Tagliamonte, L.A. McReynolds, and C.D. Guan.** 1988. An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. Gene *74*:365-373.
10. **Nagai, K. and H.C. Thogersen.** 1984. Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. Nature *309*:810-812.
11. **Nagai, K. and H.C. Thogersen.** 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. Methods Enzymol. *153*:461-481.
12. **Nakashima, A., K. Mori, T. Nagatsu, and A. Ota.** 1999. Expression of human tyrosine hydroxylase type I in *Escherichia coli* as a protease-cleavable fusion protein. J. Neural Transm. Gen. Sect. *106*:819-824.
13. **Porath, J., J. Carlsson, I. Olsson, and G. Belfrage.** 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature *258*:598-599.
14. **Sells, S.F., D.P. Wood, Jr., S.S. Joshi-Barve, S. Muthukumar, R.J. Jacob, S.A. Crist, S. Humphreys, and V.M. Rangnekar.** 1994. Commonality of the gene programs induced by effectors of apoptosis in androgen-dependent and -independent prostate cells. Cell Growth Differ. *5*:457-466.
15. **Sells, S.F., S.S. Han, S. Muthukkumar, N. Maddiwar, R. Johnstone, E. Boghaert, D. Gillis, G. Liu et al.** 1997. Expression and function of the leucine zipper protein Par-4 in apoptosis. Mol. Cell Biol. *17*:3823-3832.
16. **Smith, D.B. and K.S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene *67*:31-40.
17. **Walker, P.A., L.E. Leong, P.W. Ng, S.H. Tan, S. Waller, D. Murphy, and A.G. Porter.** 1994. Efficient and rapid affinity purification of proteins using recombinant fusion proteases. Biotechnology *12*:601-605.
18. **Wearne, S.J.** 1990. Factor Xa cleavage of fu-

sion proteins. Elimination of non-specific cleavage by reversible acylation. FEBS Lett. *263*:23-26.

*The authors wish to thank Dr. Wai-ming Lee and Dr. Roland R. Rueckert for the kind gift of DNA corresponding to HRV-14, Akiko Koide and Keith Nehrke for discussions regarding molecular cloning, and New England Biolabs for the pMal-c2T vector. This project was funded in part by the Department of Defense Prostate Cancer Research Program grant no. DAMD17-97-1-7295. Address correspondence to Dr. Steven M. Pascal or Andrei Alexandrov, Department of Biochemistry and Biophysics, Box 712, University of Rochester Medical Center, Rochester, NY 14642, USA. e-mail: pascal@oxbow.biophysics.rochester.edu, aalexand@mc.rochester.edu*

Received 17 November 2000; accepted 19 February 2001.

**Andrei Alexandrov, Kaushik Dutta, and Steven M. Pascal**
*University of Rochester Medical Center, Rochester, NY, USA*

## Sensitive Ribonuclease Protection Assay Employing Glycogen as a Carrier and a Single Inactivation/ Precipitation Step

*BioTechniques 30:1198-1204 (June 2001)*

The ribonuclease protection assay (RPA) is a sensitive method for the detection and quantitation of specific messenger RNAs in total RNA or mRNA samples (7). Both of these parameters are enhanced using RPAs compared to other RNA detection procedures that use solid support, such as northern blot hybridization. However, RPAs are time consuming, and the protocol is difficult to perform consistently, in part because of the number of manipulations required to process the samples. We describe a modification of the traditional RPA that simplifies the assay by combining the ribonuclease inactivation and RNA precipitation steps into a single step using one solution. In addition, the sensitivity of the method was enhanced by including glycogen in either the RNase digestion buffer or the hybridization buffer, which resulted in increased quantitative recovery of dsRNAs. Commercial RPA kits are available that utilize these modifications, but this protocol allows reagent solutions to be assembled at a fraction of the cost, using common, less-expensive chemicals generally available in most molecular biology laboratories.

Essentially, the RPA is a solution hybridization protocol in which an excess of labeled antisense RNA probe is incubated with an RNA sample. Complementary probe and target mRNA sequences anneal, forming dsRNA hybrids. Ribonucleases, enzymes that digest only unpaired ssRNAs, are added to the mixture. dsRNA molecules are resistant to ribonuclease digestion; thus, the target sequence is "protected." Ribonucleases are inactivated, RNA fragments are precipitated, and the protected fragments are separated by denaturing polyacrylamide gel electrophoresis. Detection is either by autoradiography or nonisotopic procedures, depending on the type of labeled probe utilized (3,6,7).

Typically, the RNA regions protected are designed to be small, usually in the range of 100–300 bp, so that they may be readily resolved during electrophoresis. These fragments can be difficult to recover quantitatively, so carriers such as tRNA are often added to aid the in their precipitation. Normally, tRNA is included in the ribonuclease inactivation solution, but because tRNA precipitates in the presence of isopropanol or ethanol, it was necessary to perform separate ribonuclease inactivation and nucleic acid precipitation steps. We rationalized that if the carrier could be included in the hybridization buffer or ribonuclease digestion buffer, the ribonuclease inactivation and RNA precipitation steps could be integrated into a single step, further streamlining this multistep protocol. Purified glycogen was chosen as the co-precipitate to test. Glycogen is a relatively inert molecule that can be used for precipitation of DNA or RNA, has demonstrated efficiency in the quantitative recovery small nucleic acid fragments, and is less expensive than tRNA (2,5).

The protocol summarized in Table 1 is a modification of one previously published (4). Basically, the method we began with is as follows. Total RNA was prepared from human β-globin locus yeast artificial chromosome (β-YAC) transgenic mouse 12-day fetal blood samples by the method of Chomczynski and Sacchi (1). Three radiolabeled, antisense RNA probes were synthesized using T7 RNA polymerase (Ambion, Austin, TX, USA) in the presence of $^{32}$P-UTP (Amersham Pharmacia Biotech, Pisscataway, NJ, USA), mouse α-globin (128-bp protected fragment), human $^{A}$γ-globin exon II (170-bp protected fragment), and human β-globin exon II (205-bp protected fragment). Unincorporated nucleotides were removed by centrifugal gel filtration through mini Quick Spin™ RNA columns (Roche Molecular Biochemicals, Indianapolis, IN, USA). Fifty nanograms of RNA were mixed with $10^6$ cpm of each probe (specific activity was approximately $4 \times 10^8$ cpm/μg); the mixture was vacuum-dried and resuspended in 20 μL RNA hybridization buffer (80% formamide, 400 mM NaCl, 40 mM PIPES, pH 6.4, 1 mM EDTA). The RNA was denatured by heating the