Reactive site polymorphism in the murine protease inhibitor gene family is delineated using a modification of the PCR reaction (PCR+1)

Frank Borriello and Kenneth S.Krauter*
Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York, NY 10461, USA

Received June 5, 1990; Revised and Accepted August 8, 1990

ABSTRACT

Murine protease inhibitor (α_1-PI) proteins are encoded by a multigene family which has undergone recent duplication. It has been suggested that the evolution of diversity within this gene family may be driven by unusual selection for novel function at the reactive site of the duplicated members (1,2,3). In an attempt to use polymerase chain reaction (PCR) to generate and sequence clones spanning the polymorphic reactive site region, a PCR artifact was identified and determined to result from heteroduplex formation during the co-amplification of the related sequences in this multigene system. This artifact results in sequences which are combinatorial mosaics of the template sequences. We present a simple and general method (PCR+1) for overcoming this artifact and demonstrate its application in delineating five distinct α_1-PI reactive site sequences in C57BL/6 mice, thus providing sequence information to generate gene-specific probes. The significance of the reactive site diversity in this protease inhibitor gene family is discussed as well as the general applications and limitations of the PCR+1 technique.

INTRODUCTION

The protease inhibitors of man and mouse control proteolysis involved in many physiological processes including blood clotting, complement activation, development and metastasis (4). It has been demonstrated by genetic analysis, that α_1-AT deficiency in man is associated with the development of emphysema (7) resulting from unopposed neutrophil elastase activity on the lung tissue. The multiple serine protease inhibitors (α_1-PI) in mice and their single human homologue α_1-antitrypsin (α_1-AT), function as suicide inhibitors by forming a stable 1:1 covalent adduct with a target protease active site (5) which dissociates slowly to yield active protease and inactive cleaved inhibitor (1). A short peptide loop on the surface of the inhibitor protein is recognized by the cognate protease and is cleaved (6). However, unlike the proteolysis of other substrates, the protease fails to dissociate due to conformational changes in the inhibitor which traps the protease in a complex and thus is inactivated in a stoichiometric fashion (6). The specific amino acid sequence of the inhibitor peptide loop is critical to the specificity of this reaction and is called the reactive site.

The human α_1-AT and murine α_1-PI genes are 12 kb and 10 kb long, respectively (8,9) and contain 5 exons. The inhibitors are expressed at a high level in a liver abundant manner (10). Preliminary analysis in mice has suggested that there are at least four α_1-PI genes clustered on murine chromosome 12, 8−12 centimorgans centromeric to the immunoglobulin heavy chain locus (11,12). Although it has been shown that purified murine α_1-PI can effectively inhibit neutrophil elastase *in vitro* it is not clear which members of the family are responsible for this activity or what activities the other members may possess (13).

Partial DNA sequence analysis of two independently isolated cDNA clones and three genomic DNA subfragments from BALB/c mice has shown there to be less than 4% overall sequence divergence between members of the gene family in protein coding regions as well as introns, 3′ untranslated regions, and 5′ flanking regions (9,15, and Smith and Krauter unpublished observation). One of the cDNA clones analyzed encodes a tyrosine while the other encodes a methionine in the critical P1 position of the reactive site like the human α_1-AT gene (14). Based on limited DNA sequence information and by comparing one of the two available mouse protein sequences with those of related serine protease inhibitors, it has been proposed that the α_1-PI and α_1-AT gene(s) of mice and man have undergone an unusual evolutionary selection process which has created diversity at the reactive site (15). The evolution of several antiproteases has been postulated to be dependent and responsive to evolutionary changes in their cognate substrates (2). In order to better understand the extent and significance of this diversity we have develped a new PCR strategy to determine the complexity of the multigene family with respect to the number of α_1-PI genes and the DNA sequences present at each of the reactive sites. We determined that there are at least five α_1-PI reactive sites in C57BL/6 mice which are expressed and are likely to encode different α_1-PI activities.

* To whom correspondence should be addressed

MATERIALS AND METHODS

Preparation of Oligonucleotides

Oligonucleotides shown in figure 3 were synthesized on an Applied Biosystems 380B DNA Synthesizer and resuspended at a concentration of 100 μg/ml in water using the arbitarary extinction coefficient of 25 μg/OD260, regardless of the specific oligonucleotide sequence. Oligonucleotide probes were labeled with T4 polynucleotide kinase (PNK) (16). Probes were added directly to the hybridization mixture without further processing at a final concentration of 10 ng/ml.

Oligonucleotide hybridization conditions

Hybridization of oligonucleotides was carried out in 5× SSPE, 5% SDS, 0.2% dry milk and 100μg/ml heat denatured salmon sperm DNA. Hybridization was allowed to proceed for 8 hours at an empirically determined optimal temperature. Washing was carried out at the hybridization temperature in a solution of 2× SSPE, 0.2% SDS. Blots were exposed at −70°C with intensifying screens.

PCR Amplificiation

PCR was carried out according to instructions supplied by Perkin-Elmer-Cetus, using 5μg C57BL/6 mouse genomic DNA and 200 ng of each amplimer in a total volume of 100μl. The reaction mixture contained: 200μM dNTP (each), 50mM KCl, 100mM Tris-HCl, pH 8.3, 15mM $MgCl_2$, 0.1% (w/v) gelatin and 2.5U Taq polymerase. The PCR cycle parameters were 94°C denaturation for 2 min., 55°C annealing for 2 min., and 72°C extension for 3′ for a total of 30 cycles. All tubes and tips were silanized and sterilized. The reaction products were extracted twice with phenol:chloroform (50:49 v/v) followed by chloroform extraction. The aqueous layer was transferred to a new tube and DNA precipitated by the addition of 100ul of 4M ammonium acetate and 2 volumes 95% ethanol at −70°C. DNA was recovered by centrifugation, the pellet washed with 70% ethanol, and then dried under vacuum.

Colony Lift and Screening Protocol

PCR products were cloned as described in the results section. Clones were screened by standard colony filter hybridization techniques (16). The number of insert containing PCR clones varied from 5−20% of the total number of colonies present. Plasmid DNA was isolated from positive colonies and subjected to DNA sequencing using dideoxy sequencing of double stranded DNA as described previously (17).

RESULTS

PCR Amplification of α_1-PI Reactive Site Sequences

To determine the number of different α_1-PI genes in C57BL/6 animals, we used the PCR technique to amplify the most highly polymorphic segment of the genes and then subjected the PCR products to DNA sequence analysis after cloning. Based on limited DNA sequence information from BALB/c mice, it appeared that the reactive site of the α_1-PI genes would be the best candidate segment to amplify (9,15). Oligonucleotides were chosen for the PCR amplification of a 140 bp segment of genomic DNA which encodes the reactive site peptide loop. The chosen sequences, amplimers 1 and 2, flanked the reactive site and corresponded to DNA sequences which were completely conserved within exon 5 among all of the α_1-PI sequences available (figure 3 and ,9,15). Extension of amplimer 1 produced the sense strand while amplimer 2 produced the antisense strand. Amplimer 2 carried a HindIII site within 8 nucleotides of its 5′ end which was not complementary to α_1-PI sequences but facilitated subsequent cloning of the resulting PCR products. After amplification of 5μg of C57BL/6 genomic DNA, a small aliquot was analyzed by electrophoresis on a 1.5% agarose gel to verify amplification. Under these conditions, one major band was observed, in agreement with the expected size of 193 bps which corresponds to the amplified region plus the length of the two amplimers. Direct DNA sequence determination of PCR products driven by these amplimers would have been impossible to interpret because of the known level of polymorphism within this region. Therefore the PCR products were cloned into a plasmid vector (pSP72; Promega) and subjected to DNA sequence analysis.

DNA Sequences of PCR Clones Suggest an Artifact

DNA sequence analysis of 20 independently isolated clones produced fifteen different but closely related sequences. Since previous estimates of the number of different α_1-PI genes were far less than 15 (11), we suspected that our methods had introduced an artifact which resulted in sequence mutation. A comparison of each clone to the others showed that the cloned segments contained a limited number of short DNA sequence motifs which appeared to be recombined in each sequence (data not shown).

One previously described mechanism to account for this result, termed 'PHLOP' amplification, involves the incomplete extension of an amplimer in one PCR cycle followed by its further extension on a different gene template to form a complete strand in subsequent cycles (18) . This mechanism results in the joining of sequences from multiple genes to create mosaic genes. An alternative explanation of our inital results was that heteroduplex dsDNA was formed between two complete strands derived from different genes during the denaturation-annealing steps of the PCR (figure 1, step 1a and 2a). The cloning of these heteroduplexes into bacteria capable of DNA mismatch repair lead to independent repair to one strand or the other, at each point of mismatch. The PHLOP mechanism would generate homoduplexes with mosaic single strands and result in sequences which were stably scrambled during the amplification procedure. These mosaic sequences could contain sequences from two or more genes. The second mechanism however would produce sequences that were not scrambled until the heteroduplexes were repaired in bacteria. Furthermore, these sequences would be expected to be mosaics of only two different genes.

No method exists to circumvent the PHLOP artifact, although changing the reaction conditions could decrease its incidence (18). If however, the latter mechanism of bacterial mismatch repair was responsible for the majority of the observed artifacts, then it should have been possible to obtain unscrambled sequences if heteroduplexes could be eliminated prior to cloning.

The PCR+1 Reaction

To test our hypothesis that the observed artifacts were due to the mechanism of bacterial mismatch repair, we designed a method which permitted the selective cloning of homoduplex molecules (figure 1, steps 1b, 2b and 3). This method used one additional amplimer (amplimer 3) which was similar to amplimer 1 but encoded a unique restriction site (BamHI) at its 5′ end. One cycle of denaturation and synthesis using amplimer 3 as a primer should produce a new population of PCR products which

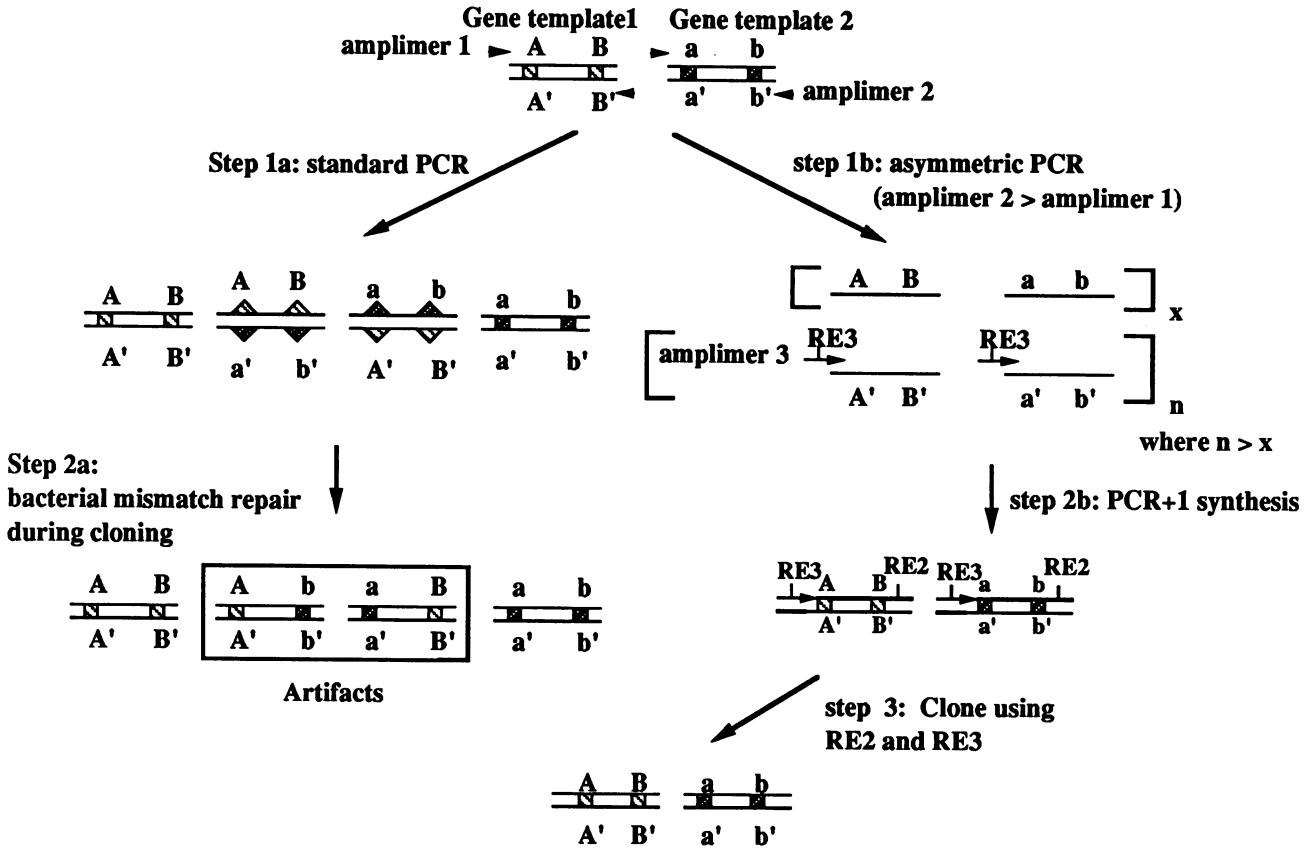


Figure 1. *The heteroduplex mediated mismatch repair artifact detected during the PCR amplification of a multigene family and the PCR+1 solution.* The simplest case of two related genes which differ at two loci denoted as A and B are amplified. When opposite sense strands from the two genes anneal, heteroduplexes are formed (step 1a). Upon cloning into bacteria capable of mismatch repair, these heteroduplexes are repaired at all points of mismatch which can lead to formation of mosaics of the two original template sequences (step 2a). The PCR+1 method employs one additional amplimer in one more PCR cycle. Genomic DNA is amplified using asymmetric PCR where amplimer 2 is in excess of amplimer 1 (step 1b). The resulting PCR products contain an excess of the strand primed by amplimer 2 (denoted in brackets where n > x). Amplimer 3 is then employed in one addition cycle of melting, annealing and synthesis followed by no further melting (step 2b). Since amplimer 3 encodes a novel restriction site, the resulting homoduplexes can be specifically cloned (step 3b). RE2 and RE3 denote restriction enzyme sites encoded on amplimers 2 and 3, respectively.

have a unique restriction site at one end. The new strand, primed by amplimer 3 and fully extended, must be homoduplex and can be specifically cloned using the unique amplimer 3 encoded restriction site. Heteroduplexes can only arise with subsequent cycles of denaturation and reannealing which are omitted here.

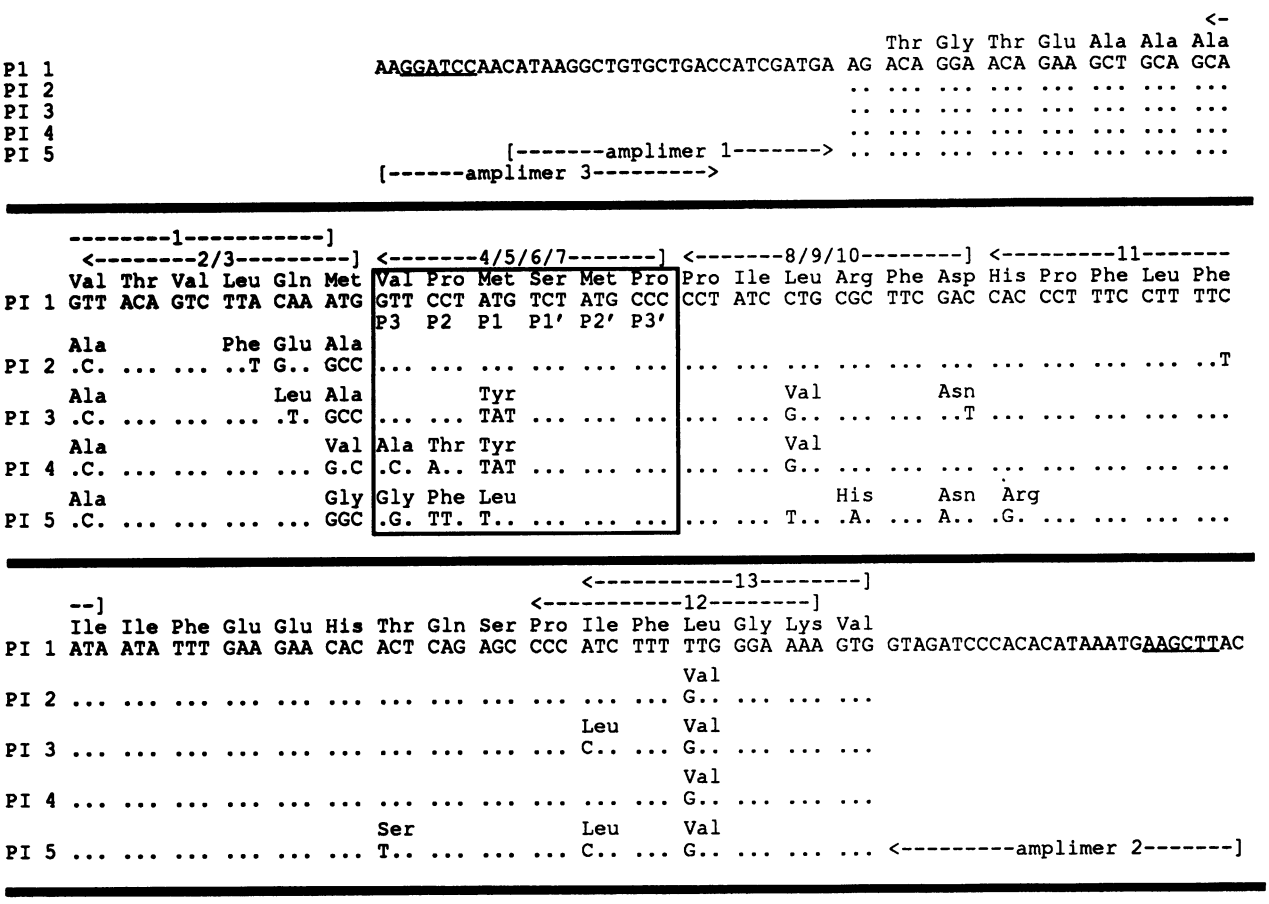
In order to increase the efficiency of the PCR+1 reaction, the genomic DNA was first amplified by an asymmetric PCR reaction which used the same conditions as the standard PCR reaction with the exception that only one tenth the amount of amplimer 1 (20 ng) was added (figure 1, step 1b)(16). Under conditions of limited amplimer 1, thirty cycles of PCR leads to the exhaustion of amplimer 1 in solution and creates an excess of the antisense strand. The resulting asymmetry in PCR strands increases the efficiency of the PCR+1 product because amplimer 3 will have diminished competition for the antisense strand by either amplimer 1 or the sense strand.

One quarter of the products of the asymmetric PCR reaction were mixed with 200ng of amplimer 3 and subjected to a single denaturation-extension cycle (figure 1, step 2b). Although heteroduplexes which did not undergo the additional cycle surely contaminated the reaction products, they did not contain the additional amplimer 3 encoded restriction site required for the cloning strategy and should therefore be eliminated. The PCR+1 products were digested with BamHI and HindIII, ligated to identically cleaved pSP72 vector DNA and transformed into bacteria. Clones containing PCR product inserts were identified by colony hybridizations with PCR amplimer 3 under stringent conditions. Plasmid DNA from the positive colonies was characterized by DNA sequence determination of the entire 193bp insert using the appropriate oligonucleotide primers flanking the insertion site.

Reconstruction and Correction of the Artifact

In a reconstruction experiment two clones of PCR products, obtained from the earlier amplification of the genomic DNA, which differed at only two nucleotide positions were used as templates and were co-amplified. These two nucleotide differences provided template specific markers which could be followed through the course of the experiment. The templates were mixed, co-amplified using the standard PCR conditions, and the products cloned. Fifteen clones were isolated and the DNA sequence of the inserts were determined. Nine clones were identical to one of the two template plasmids while six contained a rearrangment of template markers which indicated mosaicism. Thus the artifact was reproduced in a defined system. When the same two test templates were co-amplified using the PCR+1

A. C57BL/6 α_1-Protease Inhibitor Reactive Sites



B. BALB/c α_1-Protease Inhibitor Reactive Sites

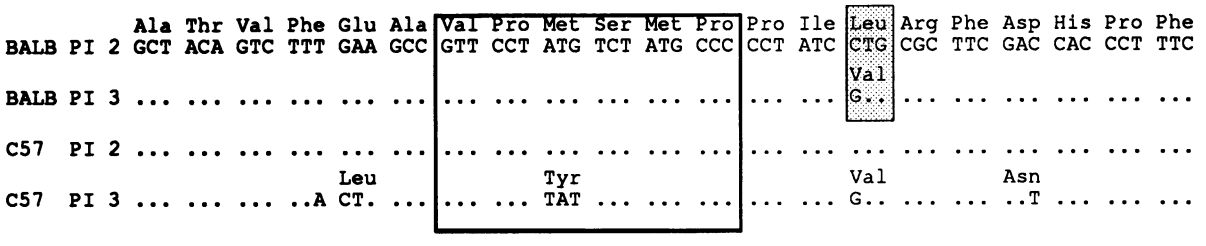


Figure 2. *Panel A*: The nucleic acid and amino acid sequences of the five amplified C57BL/6 α_1-PI reactive sites are shown. Oligonucleotides used for screening and PCR are shown with arrows indicating their orientation. Nucleic acid and amino acid changes relative to α_1-PI-1 are indicated. Identity with the C57BL/6 α_1-PI-I sequence is indicated by a dot(.). The box shows the reactive site peptide loop where P1 is the reactive site methionine. *Panel B*: The nucleic acid and amino acid sequences of two of the BALB/c α_1-PI gene reactive sites which differ by a single nucleotide over the 140 base fragment amplified. The box shows the reactive site as seen in *Panel A* above while the shaded box shows the altered codon. Dots indicate identity with the BALB/c PI-2 gene on the top line of *Panel B*.

protocol shown in figure 1, fifteen out of fifteen clones characterized produced template sequences. No mosaic sequences were observed. Since incomplete primer extension should not be affected by the PCR+1 methodology, this reconstruction experiment indirectly showed that heteroduplex formation is the likely source of the artifact, under the conditions used, and not incomplete primer extension.

If indeed heteroduplex molecules were the source of the artifact, and repair in *E. coli* was causing the scrambling of the sequences, then transformation of the standard PCR products into repair-deficient, recombination-deficient bacteria should reduce the frequency of sequence scrambling. We therefore transformed the PCR products of the test templates described above, into a mutS:tn10 repair deficient strain rather than into LM1039, our standard transformation host strain. We observed no scrambling of the sequences in the 10 clones sequenced (data not shown)

whereas nearly half of all clones sequenced from wild type host cells showed scrambling. We did not systematically pursue this method of eliminating the artifact further because these strains are limited in their ability to suppress repair of heteroduplex regions of more than a few base pairs (20) and is thus limited in its general applicability. However, this result provides support for our hypothesis that it is indeed repair of heteroduplexes which causes the sequence scrambling.

These results implied that the PCR reaction produces a fraction of molecules which denature but are not subsequently primed by amplimers. Instead, they reanneal with other single strands to form heteroduplexes. It is possible that in the terminal stages of the PCR reaction, the concentration of product becomes so high that the rate of reannealing of product to product is more rapid than the rate of annealing of oligomer to template thus resulting in an increase in the number of heteroduplexes. On the other hand, it may be that at each cycle, a fraction of molecules does not prime and can reanneal to form heteroduplexes. We have only used the conditions recommended as optimal for genomic DNA amplification (18) and have not investigated whether changing the number of cycles or reaction conditions can reduce the formation of heteroduplexes. Since we used relatively popular conditions, it is fair to say that this artifact may occur more frequently than has been previously appreciated. The artifact is such that for the simplest case of a two template system (figure 1) with two polymorphic sites, the complexity of the products doubles. In the case of five or more genes with multiple polymorphisms, the complexity of products obtained can be enormous.

Amplification and DNA Sequence Determination of the α_1-PI reactive site

We applied the PCR+1 method to the characterization of C57BL/6 α_1-PI reactive sites using genomic DNA as a template. One hundred independent clones which contained the unique amplimer 3 sequence were subjected to detailed analysis either by DNA sequencing or hybridization to a panel of oligonucleotide probes. DNA sequence analysis of 15 randomly chosen clones revealed five different sequences (figure 2). Based on these five sequences , five oligonucleotide probes (2,4,5,6, and 7), which could discriminate among the genes were prepared, along with eight additional probes to permit unambiguous assignment of each clone to one of the five groups depicted in figure 2. The combination of 13 probes allowed us to screen for clones which represented scrambling of two or more sequences during PCR cloning.

Hybridization of the set of 13 different oligonucleotide probes to a panel of 100 PCR+1 α_1-PI clones indicated that 84 of 100 screened clones contained inserts with precisely one of the five α_1-PI reactive site sequences. Eight clones had lost their inserts or contained amplimer sequence but no intervening α_1-PI sequence and an additional eight clones had a single nucleotide difference with respect to one of the five putative reactive site sequences. These latter eight clones most likely represent Taq DNA polymerase induced misincorporation which, under the conditions used in these studies, is approximately 1 in 400 nucleotides in the final product (19).

As further evidence that there are at least five α_1-PI gene family members in C57BL/6 mice, we have isolated 97 α_1-PI cDNA clones from a C57BL/6 mouse liver library using an α_1-PI probe containing exons 3, 4, and 5. Of these cDNA clones, 76 contained the fifth exon reactive site. Screening with the panel of 13 oligonucleotide probes showed that all of the 76 cDNA clones hybridized to one of the five gene sequences determined by PCR+1 with no exceptions. The number of clones of each type ranged from 7 to 27. This suggests that the level of expression of each mRNA is approximately equal within statistical limits, assuming that the library is representative of the abundance of all members. This corroborates the PCR+1 results and shows that all five genes are expressed in liver. Furthermore, since only two genes have the cannonical Met-Ser in their reactive sites, it is possible that the three other genes have evolved new substrate specificity (Borriello and Krauter, In prepartion).

We have thus delineated the extent of reactive site diversity among the liver expressed α_1-PI genes. These data do not rule out the possibility that the amplimers chosen for PCR failed to hybridize to additional α_1-PI genes but it does suggest, based on the distribution of cDNA clones that those additional genes, if any, are not expressed to levels comparable to the other five. The possibility also exists that two cDNAs that we have grouped together may actually differ at a point(s) outside the reactive site.

#	Sequence 5'-> 3'	PI1	PI2	PI3	PI4	PI5
1	TTTGTAAGACTGTAGCTG	+c	-	-	-	-
2	GGCTTCAAAGACTGTAG	-	+	-	-	-
3	GGCTAGTAAGACTGTGA	-	-	+	-	-
4	GGGCATAGACATAGGAAC	+	+	-	-	-
5	GGGCATAGAATAAGGAAC	-	-	+	-	-
6	GGGCATAGAATAAGTAGC	-	-	-	+c	-
7	GGGCATAGACAAAAAACC	-	-	-	-	+c
8	GTCGAAGCGCACGATAGG	-	-	+	+	-
9	GTTGAAGTGCAAGATAGG	-	-	-	-	+c
10	GTCGAAGCGCAGGATAGG	+	+	-	-	-
11	TATAAAAAGGAAAGGGTG	-	+	-	-	-
12	TTTTCCCACAAAGAGGGG	-	-	+	-	+
13	CACTTTTCCCAAAAAGAT	+c	-	-	-	-

amplimer 1 CATAAGGCTGTGCTGACCATCGAT
amplimer 3 AAGGATCCAACATAAGGCTGTGCTGAC

amplimer 2 GTAAGCTTCATTTATGTGTGGGATCATC

Figure 3 Listed are the oligonucleotides used in this work. Their number, sequence and complementarity with the five distinct α_1-PI reactive sites of C57BL/6 are indicated by +/−. Lower case (c) indicates that these probes discriminate C57BL/6 from BALB/c PCR products generated under the same conditions. All probes are negative strand except for amplimer 1 and amplimer 3 which are positive strand. Restriction enzyme sites are underlined in amplimer 2 and 3 which contain a HindIII and BamHI site, respectively.

DISCUSSION

We have described a novel PCR artifact that interferes with attempts to clone PCR amplified products from a multi-template system. Such amplification results in the formation of heteroduplexes during the course of the PCR reaction. Attempts to clone such heteroduplexes in bacteria capable of mismatch repair can generate a large overestimate of the size of a multigene family by producing many sequences which are mosaics of two of the original template sequences. While in theory, the cloning of PCR products into mismatch repair deficient bacterial host strains can eliminate this artifact such bacterial mutants appear to be limited in their ability to suppress the repair of increasingly

long heteroduplex regions (20). Therefore, the more general PCR+1 method was developed.

The PCR+1 method permits the production of homoduplexes in an additional PCR cycle employing a third amplimer which encodes an additional restriction site. This allows selective cloning of the homoduplexes and rapid screening of large numbers of independently isolated clones. Further characterization of the members of a multigene family by DNA sequence analysis within the defined region of polymorphism is therefore easily achieved. Taq polymerase fidelity remains an inherent problem of PCR technology but can be somewhat compensated for by screening a large number of clones. Sequences obtained from this rapid method can be used to develop gene-specific probes for further analysis and extension of the PCR+1 results.

We have used the PCR+1 method to determine the reactive site sequences of five α_1-PI genes in C57BL/6 mice (figure 2). These sequences confirm and extend the previous observation that clustered mutations at the reactive sites of the α_1-PI proteins are characteristic of the evolution of this gene family. Despite a reported high level of DNA sequence conservation among these five genes overall, the inferred amino acid sequences of the five different reactive site regions showed a significant number of polymorphisms within a rather narrow region surrounding the reactive site methionine (figure 2, P1 position). Two genes, α_1-PI-1 and α_1-PI-2, are similar to the human α_1-AT gene in that they encode a critical methionine at the P1 position, while two, α_1-PI-3 and α_1-PI-4, encode a tyrosine and one, α_1-PI-5, encodes a novel leucine at this position. The clustering of amino acid substitutions within the protease recognition loop of the antiproteases suggests that each has a different spectrum of cognate protease substrate specificities and affinities, although this remains to be proven. A complete analysis of the full length cDNAs will more clearly determine the patterns of evolution of the sequences of this multigene family.

The α_1-PI sequence determination of fifteen randomly chosen clones produced five different sequences. The probability of there being a sixth sequence in the reactive site that was missed by purely statistical error is given by Bayes' Theorem and in this case is <0.025. This assumes that the amplimers bound equally well to all gene members and that each gene is present only once in the genome. The oligonucleotide hybridizations scanned an additional 85 informative PCR+1 clones and revealed no additional sequences. Applying a purely statistical argument to these data implies that there is essentially no chance that there is a sixth gene. However, this does not rule out base changes outside the region spanned by the oligonucleotides nor truly systematic reasons for a failure to detect a sixth gene. Nevertheless, it seems reasonable to conclude that there are at least five genes with only a small probability of there being additional members.

It is obvious that statistical arguments about the number of different genes are irrelevent if the methods used to isolate clones bias the definition of an independent gene. If a pair of genes differs outside of the region screened or is subtlely different, it may escape detection. For example, we have carried out a preliminary PCR+1 study of the α_1-PI genes in another inbred mouse strain, BALB/c. The data show there to be at least five genes which appear to be somewhat diverged from the C57BL/6 homologues (data not shown). The reactive sites of two of the five genes identified in BALB/c were found to differ from one another by a single nucleotide change within the 140 bp amplified region (see figure 2B). This was confirmed by isolation of representative clones of each type from a BALB/c genomic library. Therefore, it appears that two of the BALB/c genes have diverged less than 1% in a highly polymorphic region. This example serves to point out that for any gene, it is virtually impossible to rule out the existance of additional closely related members. The PCR+1 method provides an approach to this problem which permits the rapid DNA sequence analysis of independently isolated clones of genes of interest. Sufficiently large numbers of clones can be analyzed to allow one to apply statistical methods to determine the probability that all of the related members have been examined. However, the methods of identifying family members must be carefully examined before ruling out the existance of new genes.

The PCR+1 method developed for the analysis of the α_1-PI genes should be directly applicable to the study of other gene families. The PCR+1 method followed by the characterization of cloned products in bacteria permits a rapid and sensitive method to sort out the complexity of multigene families and to yield gene-specific probes. It is important to emphasize that PCR is fraught with numerous systematic artifacts that can best be controlled by deriving PCR independent confirmation of any conclusions. To this end, gene-specific probes can be used in northern and southern analyses as well as in the characterization of genomic and cDNA clones. Given the large number of multigene families currently under study and the many instances where co-amplification of multiple related template sequences is a problem, PCR+1 provides not only a solution to a potential artifact but a useful way to determine the sequence of individual members rapidly. The method is compatible with a large scale screening effort to obtain preliminary sequence data on the variability of any multi-template system making this method one of general applicability, limited only by the requisite preliminary sequence data for amplimer design and Taq polymerase fidelity.

ACKNOWLEDGEMENTS

We thank Diane McGovern and Cathy Rapelje for prompt and expert synthesis of the oligonucleotides. We thank Graham Walker at MIT for the DNA mismatch repair mutant, mutS:tn10. This work was supported in part by National Institutes of Health Grants CA39553 and CA31855 (to K. K.) and 5T32GM7288 (to F. B. who is an MSTP student).

REFERENCES

1. Carrell, R.W., Pemberton, P. A. and Boswell, D.R. (1987) The Serpins: Evolution and Adaptation in a Family of Protease Inhibitors. Cold Spring Harbor Symposia on Quantitative Biology, Vol II, pp. 527–535.
2. Laskowski, M., Kato, I., Kohr, W.J., Park, S.J., Tashiro, M. and Whatley, H.E. (1987) Positive Darwinian Selection in Evolution of Protein Inhibitors of Serine Proteinases ibid, pp. 545–553.
3. Hill, R.E. and Hastie, N.D. (1987) *Nature* 326, 96–99.
4. Laskowski, M and Kato, I (1980) *Biochem.* **49**, 593–626.
5. Cohen, A. B. (1973) *J. Biol. Chem.* **248**, 7055–7059
6. Travis, J. and Salvesen, G. S. (1983) *Ann. Rev. Biochem.* **52**, 655–709.
7. Gadek, J.E. and Crystal, R.G. (1982) in J. B. Stanbury, J. B. Wyngaarden, D. S. Frederickson, J. L. Goldstein, and M. S. Brown (ed.), The Metabolic Basis of Inherited Disease. 5th Edition, pp. 1450–1467.
8. Bao, J.J., Sifers, R.N., Kidd, V.J., Ledley, F.D., Woo, S.L. (1987) *Biochemistry* **26**, 7755–7759.
9. Krauter, K.S., Citron, B.A., Hsu, M.-T., Powell, D., Darnell, J.E. (1986) *DNA* **5**, 29–36
10. Carrell, R.W., Jeppson, J.-O., Laurell, C.-B., Brennan, S.O., Owen, M.C. Vaughan, L., and Boswell, D.R. (1982) *Nature* **298**, 329–334.
11. Hill, R.E., Shaw, P.H., Barth, R.K., and Hastie, N.D. (1985) *Molec. and Cellular Biol.* **5**, 2114–2122.

12. Blank, R. E., Campbell, G. R., Calabro, A. and D'Eustachio (1988) *Genetics* **120**, 1073–1083.
13. Takahara, H and Sinohara, H (1983) *J. Biochem* 93, 1411–1419.
14. Johnson, D., Travis, J. (1978) *J. Biol. Chem.* **253**, 7142–7144.
15. Hill, R.E., Shaw, P.H., Boyd, P.A., Baumann, H. and Hastie, N.D. (1984) *Nature* **311**, 175–177.
16. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory. .pp 309–362..
17. Kraft, R., Tardiff, J., Krauter, K.S., and Leinwand, L.A. (1988) *Biotechniques* 6, 544–547
18. Frohman, M. A. and Martin, G. R.. (1990) in Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (eds.) PCR Protocols: A Guide to Methods and Applications
19. Tindall, K.S., Kunkel, T.A. (1988) *Biochemistry* 27, 6008–6013.
20. Radman, M., Wagner, R. (1986) *Ann. Rev. Genet.* **20**, 523–538.