Journal of Medical Genetics 1987, **24**, 47–51

DNA polymorphisms of the human α_1 antitrypsin gene in normal subjects and in patients with pulmonary emphysema

I HODGSON AND N KALSHEKER

From the Department of Medical Biochemistry, University of Wales College of Medicine, Cardiff Royal Infirmary, Cardiff CF2 1SZ.

SUMMARY Alpha$_1$ antitrypsin deficiency predisposes subjects to developing pulmonary emphysema and childhood liver cirrhosis. We have studied restriction fragment length polymorphisms (RFLPs) of the α_1 antitrypsin gene in a normal population and a group of patients with pulmonary emphysema. We have identified five RFLPs with eight restriction enzymes. The most frequent polymorphisms have been detected with the enzymes *Msp*I, *Pst*I, and *Taq*I at frequencies of 46·8%, 6·4%, and 5·0% respectively in the group of normal subjects.

Alpha$_1$ antitrypsin (AAT) is a major serine protease inhibitor in the blood.[1] Physiologically it is particularly effective against leucocyte elastase, but it also inhibits other proteases including trypsin, chymotrypsin, and cathepsin G.[2] Over 30 different AAT variants have been described on the basis of isoelectric focusing.[3] The Z variant of the protein, which is due to a single amino acid change, predisposes subjects to pulmonary emphysema[4] and liver disease.[5] The homozygous ZZ subject is deficient in AAT with 15% of the normal mean plasma AAT concentration. AAT deficiency is thought to result in uninhibited leucocyte elastase, and the enzyme has the propensity to digest the connective tissue of the lung.[6]

Only about 5% of those with pulmonary emphysema have abnormal AAT as detected by serum AAT concentrations and isoelectric focusing.[7] We anticipated that there would be protein variants which could not be detected by isoelectric focusing as well as possible variants which affect transcription of the genes or translation of the protein. Therefore, we have used specific gene probes to look for sequence variation in and around the gene for AAT.

We describe here the results of screening a normal South Wales Caucasian population and a number of patients with pulmonary emphysema for restriction fragment length polymorphisms (RFLPs) of the human AAT gene. We have estimated the degree of heterozygosity in the gene.

Received for publication 15 October 1985.
Revised version accepted for publication 24 March 1986.

Methods

PATIENTS

We screened 101 healthy white Caucasian adult blood donors (mean age 31 years) and 24 patients with clinically well defined pulmonary emphysema (mean age 56 years). Criteria used for the evaluation of pulmonary emphysema included clinical, physiological, and radiological features. The radiological criteria used to diagnose pulmonary emphysema included hyperinflated lung fields, a paucity of vascular markings at the lung periphery, and enlarged pulmonary arteries, and in three patients thin walled bullae were evident.

AAT PHENOTYPING

AAT or proteinase inhibitor (Pi) phenotyping was performed by isoelectric focusing in agarose[8] or ultrathin polyacrylamide gels.[9]

DNA POLYMORPHISMS

DNA was extracted from 10 ml unclotted blood by standard procedures.[10] About 10 µg of DNA was digested to completion with restriction enzymes according to the manufacturer's instructions (usually for 12 to 16 hours). The DNA fragments were separated on a 1% agarose gel by electrophoresis at 40 V for 16 hours and then transferred onto

nitrocellulose by Southern blotting.[11] The nitrocellulose strips were prehybridised overnight at 55°C in 6 × concentrated SSC, 5 × Denhardt's,[12] 0·5 SDS, 100 µg/ml herring sperm DNA, and 0·05% sodium pyrophosphate.

Two different DNA probes were used: a cDNA probe of 250 bp at the 3′ end of the AAT gene including most of exon V (pNJ),[13] and a genomic clone containing a 6·5 kb *Bam*HI fragment of the AAT gene (pAT 6·5), a generous gift from Dr S Woo, Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas.[14 15] These were labelled by nick translation using α [^{32}P] dCTP.[16]

Hybridisation was carried out at 55°C (pNJ) or 58°C (pAT 6·5) for 48 to 72 hours in 6 × concentrated SSC, 1 × Denhardt's, 100 µg/ml poly A, and 0·05% sodium pyrophosphate. Filters were washed at room temperature in 2 × concentrated SSC, 0·5 SDS for five minutes, 2 × concentrated SSC, 0·1 SDS for 15 minutes, and finally 0·5 × concentrated SSC, 0·5 SDS for one hour at hybridisation temperature. Autoradiography was carried out at −70°C for three to seven days.

Often fainter bands are seen on autoradiographs which are presumably due to the homologous sequence previously described,[17] and they can be removed by more stringent washes.

Results

A number of polymorphisms have been found in samples obtained from apparently normal healthy blood donors and patients with pulmonary emphysema (table 1). Polymorphisms were detected most frequently with the enzymes *Msp*I, *Pst*I, and *Taq*I at frequencies of 46·8%, 6·4%, and 5·0% respectively in the normal group (figs 1 and 2). There are four *Msp*I fragments screened by pAT 6·5 as predicted by the genomic sequence of AAT.[15] The *Msp*I fragments screened range in size from 0·79 to 4·1 kb but only two fragments are large enough to be seen easily.

TABLE 1 *Restriction fragment length polymorphisms of the AAT gene.*

Restriction enzyme	*Fragment screened (kb)*	*No of polymorphisms/total screened*	
*Taq*I	1·5	Normal	5/101
		Affected	5/24
*Bam*HI	6·5	Normal	1/101
		Affected	1/24
*Hinc*II	3·2,2·9	Normal	1/38
		Affected	0/22
*Hind*III	2·7	Normal	0/43
		Affected	0/18
*Eco*RI	9·6	Normal	0/35
		Affected	0/15
*Xho*I	2·6	Normal	0/27
		Affected	0/15
*Pst*I	5·7, 2·4	Normal	3/47
		Affected	0/18
*Msp*I	3·3, 4·1	Normal	22/47
		Affected	12/21*

*Three subjects were homozygous for the polymorphism.

These are fragments of 3·3 and 4·1 kb. The *Msp*I polymorphism is due to deletion of a recognition site for the enzyme in the 3′ non-coding region (fig 1) which could be part of the gene or a flanking sequence; the precise 3′ end of transcription of the gene has not been established.[15] There are five *Pst*I fragments screened by pAT 6·5 as predicted by the genomic sequence of AAT,[15] which range from 0·37 to 5·7 kb. Three fragments are easily seen at 1·9, 2·4, and 5·7 kb. Other fainter bands are often seen which are presumably related to the homologous sequence and the most consistent of these are present at 1·6, 1·4, and 1·3 kb (fig 2). The *Pst*I polymorphism appears to be due to the creation of a new site in the gene close to the site of the *Msp*I polymorphism (fig 1).

There are eight *Taq*I sites in and around the AAT gene which range in size from 0·09 to 5·7 kb. The cDNA probe which spans exon V detects an invariant allele at 1·5 kb and also detects a 2·4 kb variant allele (fig 2) confirming the location of the polymorphism (fig 1). Therefore, the *Taq*I poly-

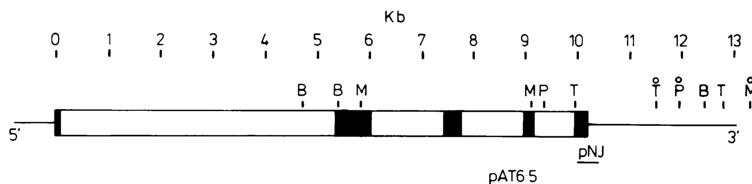

FIG 1 *Schematic representation of the human α₁ antitrypsin gene. Shaded areas represent coding sequences and non-shaped areas are introns. Only the relevant sites for some restriction enzymes are shown. B=BamHI, T=TaqI, P=PstI, M=MspI. pNJ and pAT 6·5 represent the cDNA and genomic clones respectively, M, P, T represent the approximate locations of the most frequent polymorphisms. The MspI and TaqI polymorphisms are due to loss of a restriction enzyme site and the PstI polymorphism due to creation of a new site.*

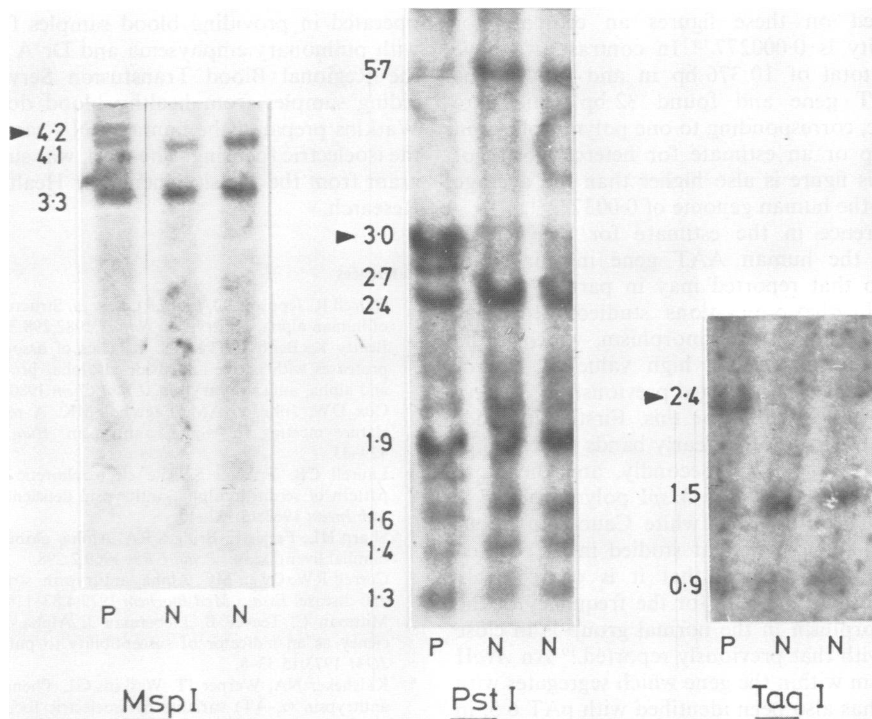

FIG 2 *Autoradiographs showing three common RFLPs of the AAT gene detected with the restriction enzymes MspI, PstI, and TaqI. P=subjects with DNA polymorphism, N='normal' subjects. The arrowheads show the position of the RFLPs. The MspI and PstI digests were probed with pAT 6·5 and the TaqI digest was probed with pNJ. The faint bands seen in the PstI digest are related to a homologous sequence.*

morphism is almost certainly due to loss of a recognition sequence in the 3′ end of the gene (fig 1). There is a small but significant increase in the frequency of the polymorphism in the emphysema group compared to the normal population (two tailed Fisher's exact probability test, $p<0·05$) in contrast to other polymorphisms which do not show significant differences in frequencies in the two groups. A study of a seven member family extending over two generations demonstrated that the polymorphism shows Mendelian inheritance and in this particular family is confirmed by the fact that the polymorphism segregates with the F allele, which is a fast moving isoelectric focusing variant. The Pi types of the subjects with the *Taq*I polymorphism include FM, MM (two subjects), MS, MZ in the blood donor group and FM, MM (two subjects), MZ, and ZZ in the patient group. The RFLP does not appear to have an association with any particular protein variant. The Pi types for the healthy blood donors and the patient group are summarised in table 2. The *Bam*HI and *Hinc*II polymorphisms are isolated examples.

We have calculated an estimate for heterozygosity by the formula of Nei[18]

$$H=1-[a/b)^2+(b-a/b)^2]$$

where a = number of polymorphic base pairs found with all the enzymes and b = total number of base pairs screened. These figures give an estimate of 0·00615 for heterozygosity of the AAT gene.

Discussion

In a previous study using a cDNA probe, only six polymorphisms out of a total of 43 232 bp were

TABLE 2 *Pi phenotypes.*

	Normal subjects (n=101)	*Patients (n=24)*
MM	82	16
MS	11	4
MZ	4	2
FM	2	1
ZZ	0	1

found. Based on these figures an estimate for heterozygosity is 0·000277.[19] In contrast, we have screened a total of 10 376 bp in and around the human AAT gene and found 32 bp which are polymorphic, corresponding to one polymorphism in every 325 bp or an estimate for heterozygosity of 0·00615. This figure is also higher than the average estimate for the human genome of 0·0037.[20 21]

The difference in the estimate for the heterozygosity of the human AAT gene in our study compared to that reported may in part be due to differences in the populations studied. However, significantly, the *Msp*I polymorphism, which is the major contribution to our high value of heterozygosity, has not been reported previously.[19 22] There are two possible reasons for this. Firstly, it is not always possible to resolve clearly bands which differ in size by about 100 bp. Secondly, and the more likely possibility, is that the *Msp*I polymorphism is found in our South Wales white Caucasian population group and not in that studied in the United States, indicating perhaps that it is of relatively recent origin. Our estimate for the frequency of the *Taq*I polymorphism in the normal group is in close agreement with that previously reported.[19] An *Ava*II polymorphism within the gene which segregates with the Z allele has also been identified with pAT 6·5, in addition to two *Taq*I polymorphisms whose frequencies and locations are not given.[22] Three other polymorphisms have been identified with a 5′ AAT probe which includes the first untranslated exon and first intron.[22]

It is clear that the AAT gene is highly polymorphic, but many of the polymorphisms occur at low frequency. The DNA polymorphisms are not related to any particular protein variants and are not associated with AAT deficiency. The high frequencies of *Msp*I, *Pst*I, and *Taq*I RFLPs contribute significantly to our estimate for heterozygosity and they all occur in close proximity to one another (fig 1). *Msp*I and *Taq*I detect polymorphisms due to the CpG dimer occurring in their recognition sequence.[23] Methylated cytosine is thought to deaminate to thymine resulting in loss of the recognition site.[24]

In our relatively small group of patients the *Taq*I polymorphism is present at a significantly higher frequency (20·8%) compared to the normal group (5·0%). The polymorphism is clearly also present in apparently healthy blood donors who have no evidence of lung disease. Consequently, it may be a coincidental finding which reflects some subtle population difference, though we have not been able to observe any obvious or consistent differences in the populations studied.

We would like to thank the physicians who cooperated in providing blood samples from patients with pulmonary emphysema and Dr A Napier from the Regional Blood Transfusion Service for providing samples from healthy blood donors. Mr G Watkins prepared the human DNA samples and did the isoelectric focusing. The work was supported by a grant from the Welsh Scheme for Health and Social Research.

References

1. Carrell R, Jeppson JO, Laurell CB, *et al.* Structure and variation of human alpha$_1$-antitrypsin. *Nature* 1982;**298**:329–34.
2. Beatty K, Beith J, Travis J. Kinetics of association of serine proteases with native and oxidised alpha$_1$ proteinase inhibitor and alpha$_1$ antichymotrypsin. *J Biol Chem* 1980;**255**:3931–4.
3. Cox DW, Johnson AM, Fagerhol MK. A report of nomenclature meeting for alpha$_1$ antitrypsin. *Hum Genet* 1980;**53**: 429–33.
4. Laurell CB, Erikson S. The electrophoretic alpha$_1$ globulin pattern of serum in alpha$_1$-antitrypsin deficiency. *Scand J Clin Lab Invest* 1963;**15**:132–40.
5. Sharp HL, Freier E, Bridges RA. Alpha$_1$ globulin deficiency in familial liver disease. *Pediatr Res* 1969;**2**:298.
6. Carrell RW, Owen MC. Alpha$_1$-antitrypsin: structure, variation and disease. *Essays Med Biochem* 1979;**4**:83–119.
7. Mittman C, Teevee B, Lieberman J. Alpha$_1$-antitrypsin deficiency as an indicator of susceptibility to pulmonary disease. *JOM* 1973;**15**:33–8.
8. Kalsheker NA, Warner JT, Watkins GL. Phenotyping alpha-1-antitrypsin (α_1-AT) variants by isoelectric focusing in agarose and immunoblotting. *Clin Chim Acta* 1985;**148**:157–60.
9. Jeppson JO, Franzen B. Typing of genetic variants of α_1-antitrypsin by electrofocusing. *Clin Chem* 1982;**28**:219–25.
10. Kunkel LM, Smith DK, Boyer SH, *et al.* Analysis of Y chromosome specific reiterated DNA in chromosome variants. *Proc Natl Acad Sci USA* 1977;**74**:1245–9.
11. Southern E. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975;**98**: 503–17.
12. Denhardt DT. A membrane-filter technique for the detection of complementary DNA. *Biochem Biophys Res Commun* 1966;**23**: 641–6.
13. Rogers J, Kalsheker N, Wallis S, *et al.* The isolation of a clone for human alpha$_1$-antitrypsin and the detection of alpha$_1$-antitrypsin in mRNA from liver and leukocytes. *Biochem Biophys Res Comm* 1983;**116**:375–82.
14. Leicht M, Long GL, Chandra T, *et al.* Sequence homology and structural comparison between the chromosomal human alpha$_1$-antitrypsin and chicken ovalbumin genes. *Nature* 1982;**297**: 655–9.
15. Long GL, Chandra T, Woo SLC, Davie EW, Kurachi K. Complete sequence of the cDNA for human alpha$_1$-antitrypsin and the gene for the S variant. *Biochemistry* 1984;**23**:4828–37.
16. Rigby PWJ, Dieckmann M, Rhodes C, Berg P. Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J Mol Biol* 1977;**113**: 237–51.
17. Lai EC, Kao FT, Law ML, Woo SLC. Assignment of the alpha$_1$-antitrypsin and a sequence related gene to human chromosome 14 by molecular hybridisation. *Am J Hum Genet* 1983;**35**:385–92.
18. Nei M. *Molecular population genetics and evolution.* Amsterdam: North Holland, 1975.
19. Matteson KJ, Ostrer H, Chakravarti A, *et al.* A study of restriction fragment length polymorphisms at the human alpha$_1$-antitrypsin locus. *Hum Genet* 1985;**69**:263–7.

[20] Cooper DN, Schmidtke J. DNA restriction fragment length polymorphisms and heterozygosity in the human genome. *Hum Genet* 1984;**66**:1–16.

[21] Cooper DN, Smith BA, Cooke HJ, Niemann S, Schmidtke J. An estimate of unique DNA sequence heterozygosity in the human genome. *Hum Genet* 1985;**69**:201–5.

[22] Cox DW, Woo SLC, Mansfield T. DNA restriction fragments associated with α_1-antitrypsin indicate a single origin for deficiency allele, Pi Z. *Nature* 1985;**316**:79–81.

[23] Barker D, Schafer M, White R. Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. *Cell* 1984;**36**:131–8.

[24] Bird AP. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* 1980;**8**:1499–504.

Correspondence and requests for reprints to Dr N A Kalsheker, Department of Medical Biochemistry, University of Wales College of Medicine, Cardiff Royal Infirmary, Cardiff CF2 1SZ.