Indian Journal of Biochemistry & Biophysics
Vol. 44, June 2007, pp. 176-178

NOTE

An alpha-1 antitrypsin genetic variant from India

Ajit K Shasany*, Ashutosh K Shukla, Mahendra P Darokar, Megha Saraiya, Nidarshana Chaturvedi, Lankeshwar Tewari[+] and Suman P S Khanuja

Genetic Resources and Biotechnology Division, Central Institute of Medicinal and Aromatic Plants, Lucknow 226 015, India
[+]Department of Biochemistry, King George's Medical University, Lucknow 226 003, India

Received 11 October 2006; revised 10 May 2007

The highly polymorphic human alpha-1 antitrypsin (AAT) gene codes for the most abundant circulating plasma serine protease inhibitor. Previously, genetic variants of the AAT gene were reported from different regions of the world. In the present study, the AAT gene was characterized in an Indian sample. The AAT gene was isolated and cloned from a liver biopsy sample through RT-PCR and the full-length gene was sequenced. Nucleotide sequence comparison with the human genome and the AAT sequences available in the GenBank (NCBI) demonstrated four unique variations—(i) an A to G variation at position 286 (Thr96Ala), (ii) an A to G variation at position 839 (Asp280Gly), (iii) a T to C variation at position 1182 that did not result in any change in the protein sequence (TTT to TTC both code for Phe) and (iv) an A to C variation at position 1200 (Glu400Asp) that resulted in replacement by an amino acid of similar nature. Other variations found were T to C at position 710 (Val237Ala) and T to A at position 863 (Val288Glu), which were also reported earlier. In conclusion, this study reports the entire 1257 bp nucleotide sequence of protein coding region of the human AAT gene from an Indian sample. This preliminary finding is significant, as it reports for the first time the AAT gene sequence in the Indian sample.

Keywords: Human alpha-1 antitrypsin, RT-PCR, Gene variants, Gene sequence

Human alpha-1 antitrypsin (AAT) is the most abundant serine protease inhibitor in the blood circulation. It inhibits elastase, a potent protease that hydrolyzes structural proteins, released from activated or disintegrating neutrophils[1]. Its deficiency causes pulmonary emphysema and childhood liver disease[2]. The pulmonary damage is probably due to the failure to adequately control the activity of neutrophil elastase, either due to low level of circulating AAT or lack of activity of the secreted AAT, or a combination of both[3]. The AAT gene is located on the long arm of human chromosome 14[4] and is organized into seven exons and six introns, of which the first three exons (1A to 1C) are non-coding. The open-reading-frame is located in the last four exons (2 to 5) and codes for 418 amino acids. The first 24 amino acids form a signal peptide, which is cleaved during intracellular processing, resulting in a 394 amino acid long mature protein[5-7].

Although a number of genetic variants of AAT have been described, most individuals express the normal M allele, which is the most common and corresponds to a normal amount of AAT in the plasma[8]. The Z allele is most commonly associated with a deficiency state (circulating level of AAT is well below the 57 mg/dl threshold needed to protect the lungs) severe enough to cause disease[2]. Individuals homozygous for Z allele or those who express it in combination with other deficiency alleles are at risk of developing emphysema. The S allele is another deficiency allele found in a small fraction of the population, effecting lower concentration of AAT in plasma. AAT is a highly polymorphic protein with more than 75 variants determined at the protein (based on electrophoretic mobility and isoelectric focusing) and/or gene level that can be categorized into four groups, according to their serum AAT level and function — normal, deficient, dysfunctional, and absent[8].

Previously, genetic variants of the highly polymorphic AAT gene have been reported from different regions of the world[7,9]. Here, we report an AAT genetic variant from the liver biopsy sample of an Indian subject, having six nucleotide variations including four unique ones.

Materials and Methods

Total RNA isolation

Total RNA was isolated from a human liver biopsy sample of a normal Indian subject using the RNeasy kit (QIAGEN Inc., USA).

Cloning and sequencing of human AAT gene

Gene-specific primers were used for RT-PCR amplification of the human AAT gene using the

*Author for correspondence:
Tel: +91-522-2359632/ext. 434
Fax: +91-522-2342666
E-mail: akshasany@yahoo.com

ThermoScript™ RT-PCR System (Invitrogen™ Life Technologies, USA). The sense and antisense primer sequences were 5′-TGGAATTCAATGCCGTCTT-CTGTCTCGTGGGG-3′ and 5′-AGCTCGAGTTAT-TTTTGGGTGGGATTCACCAC-3′, respectively and were synthesized on an ABI392 DNA/RNA synthesizer (Applied Biosystems, USA). The RT-PCR product was cloned on to pCR®2.1-TOPO® cloning vector (Invitrogen Life Technologies, USA) and transformed into *E. coli* competent cells. A recombinant clone was selected using blue/white screening and the AAT gene insert on the recombinant plasmid was sequenced using the M13 forward sequencing primer and Big Dye® Terminator *v*3.1 cycle sequencing kit (Applied Biosystems, USA) on a 3130xl Genetic Analyzer (Applied Biosystems, USA). The nucleotide variations were confirmed through three independent sequencing reactions. Sequence analysis was carried out using BLAST[10] and CLUSTALW[11].

Results and Discussion

Although many AAT variants have been reported, the majority of studies have been focused on Caucasian populations[7]. This study reports an AAT genetic variant from India for the first time. The AAT gene sequence obtained was compared with previously reported AAT sequences in the database (accession nos. K02212, BC015642, J02619, BT019455, X01683) using ClustalW (Table 1). Six nucleotide variations were found in the AAT gene sequence, which was submitted to GenBank (NCBI) under accession no. DQ682455 (Fig. 1). Four of these were unique and not found in other reported AAT sequences – (i) an A to G variation at position 286 (Thr96Ala), (ii) an A to G variation at position 839 (Asp280Gly), (iii) a T to C variation at position 1182 that did not result in any change in the protein

```
   1 M  P  S  S  V  S  W  G  I  L  L  L  A  G  L  C  C  L  V  P
   1 ATGCCGTCTTCTGTCTCGTGGGGCATCCTCCTGCTGGCAGGCCTGTGCTGCCTGGTCCCT
  21 V  S  L  A  E  D  P  Q  G  D  A  A  Q  K  T  D  T  S  H  H
  61 GTCTCCCTGGCTGAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCAT
  41 D  Q  D  H  P  T  F  N  K  I  T  P  N  L  A  E  F  A  F  S
 121 GATCAGGATCACCCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGC
  61 L  Y  R  Q  L  A  H  Q  S  N  S  T  N  I  F  F  S  P  V  S
 181 CTATACCGCCAGCTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGC
  81 I  A  T  A  F  A  M  L  S  L  G  T  K  A  D  A  H  D  E  I
 241 ATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACCAAGGCTGACGCTCACGATGAAATC
 101 L  E  G  L  N  F  N  L  T  E  I  P  E  A  Q  I  H  E  G  F
 301 CTGGAGGGCCTGAATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTC
 121 Q  E  L  L  R  T  L  N  Q  P  D  S  Q  L  Q  L  T  T  G  N
 361 CAGGAACTCCTCCGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAAT
 141 G  L  F  L  S  E  G  L  K  L  V  D  K  F  L  E  D  V  K  K
 421 GGCCTGTTCCTCAGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAG
 161 L  Y  H  S  E  A  F  T  V  N  F  G  D  T  E  E  A  K  K  Q
 481 TTGTACCACTCAGAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAG
 181 I  N  D  Y  V  E  K  G  T  Q  G  K  I  V  D  L  V  K  E  L
 541 ATCAACGATTACGTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTT
 201 D  R  D  T  V  F  A  L  V  N  Y  I  F  F  K  G  K  W  E  R
 601 GACAGAGACACAGTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGA
 221 P  F  E  V  K  D  T  E  E  E  D  F  H  V  D  Q  A  T  T  V
 661 CCCTTTGAAGTCAAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGCGACCACCGTG
 241 K  V  P  M  M  K  R  L  G  M  F  N  I  Q  H  C  K  K  L  S
 721 AAGGTGCCTATGATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCC
 261 S  W  V  L  L  M  K  Y  L  G  N  A  T  A  I  F  F  L  P  G
 781 AGCTGGGTGCTGCTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGGT
 281 E  G  K  L  Q  H  L  E  N  E  L  T  H  D  I  I  T  K  F  L
 841 GAGGGGAAACTACAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTG
 301 E  N  E  D  R  R  S  A  S  L  H  L  P  K  L  S  I  T  G  T
 901 GAAAATGAAGACAGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACC
 321 Y  D  L  K  S  V  L  G  Q  L  G  I  T  K  V  F  S  N  G  A
 961 TATGATCTGAAGAGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCT
 341 D  L  S  G  V  T  E  E  A  P  L  K  L  S  K  A  V  H  K  A
1021 GACCTCTCCGGGGTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCT
 361 V  L  T  I  D  E  K  G  T  E  A  A  G  A  M  F  L  E  A  I
1081 GTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATA
 381 P  M  S  I  P  P  E  V  K  F  N  K  P  F  V  F  L  M  I  D
1141 CCCATGTCTATCCCCCCCGAGGTCAAGTTCAACAAACCCTTCGTCTTCTTAATGATTGAC
 401 Q  N  T  K  S  P  L  F  M  G  K  V  V  N  P  T  Q  K  Stop
1201 CAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAAATAA
```

Fig. 1—Human AAT gene sequence from an Indian subject depicting the unique (bold and underlined) and other variations (bold)

Table 1—Comparison of the Indian AAT gene variant with the other human AAT sequences in the GenBank

Nucleotide Position	DQ682455 Indian variant [a]	K02212 (S variant)	J02619 (Z variant) [b]	BC015642	BT019455	X01683
286	G	A	A	A	A	A
710	C	T	C	C	C	T
839	G	A	A	A	A	A
863	A	T	A	A	A	A
1182	C	T	T	T	T	T
1200	C	A	A	A	A	A

[a]The sequence data has been submitted to GenBank (NCBI) under accession no. DQ682455; [b]The Z allele has additional variation at position 459 (A in Z allele; G in Indian variant and normal alleles) and 1096 (A in Z allele; G in Indian variant and normal alleles)

sequence (TTT to TTC both code for Phe), and (iv) an A to C variation at position 1200 (Glu400Asp) that resulted in replacement by an amino acid of similar nature. Other variations found were T to C at position 710 (Val237Ala) and T to A at position 863 (Val288Glu), which were also reported earlier.

Although this study reports a genetic variant of human AAT from India and extends the sequence data already available on AAT, further work is required to determine the role of these variations on AAT gene expression via their effect on the structural gene itself, as well as the influence on its regulatory elements like hepatocyte nuclear factors 1alpha and 4 (HNF-1alpha and HNF-4)[12]. No apparent disease phenotype could be related to the nucleotide variations in the AAT gene from the normal Indian subject. However, the clinical relevance of these nucleotide variations needs to be studied in detail by analyzing a larger sample size from the Indian population, which would also help to compute the allele frequency of the Indian genetic variant.

Acknowledgement

We are grateful to the Council of Scientific and Industrial Research, India and the Department of Biotechnology, India for the financial support.

References

1 Travis J & Salvesen G S (1983) *Annu Rev Biochem* 52, 655-709

2 Blank C A & Brantly M (1994) *Ann Allergy* 72, 105-120; quiz 120-122

3 Ray S, Mickleborough T D & Brown J L (2005) *Biochim Biophys Acta* 1740, 390-402

4 Billingsley G D, Walter M A, Hammond G L & Cox D W (1993) *Am J Hum Genet* 52, 343-353

5 Long G L, Chandra T, Woo S L, Davie E W & Kurachi K (1984) *Biochemistry* 23, 4828-4837

6 Perlino E, Cortese R & Ciliberto G (1987) *EMBO J* 6, 2767-2771

7 Hayes V M (2003) *Hum Mutat* 22, 59-66

8 Brantly M, Nukiwa T & Crystal R G (1988) *Am J Med* 84 (6A), 13-31

9 Zorzetto M, Ferrarotti I, Campo I, Balestrino A, Nava S, Gorrini M, Scabini R, Mazzola P & Luisetti M (2005) *Diagn Mol Pathol* 14, 121-124

10 Altschul S F, Madden T L, Schäffer A A, Zhang J, Zhang Z, Miller W & Lipman D J (1997) *Nucleic Acids Res* 25, 3389-3402

11 Thompson J D, Higgins D G & Gibson T J (1994) *Nucleic Acids Res* 22, 4673-4680

12 Rollini P & Fournier R E (1999) *Proc Natl Acad Sci* (USA) 96, 10308-10313