Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis

Heiko Witt[1], Werner Luck[1], Hans Christian Hennies[2], Martin Claβen[3], Andreas Kage[4], Ulrich Laβ[5], Olfert Landt[5] & Michael Becker[1]

Chronic pancreatitis (CP) is a continuing or relapsing inflammatory disease of the pancreas. In approximately one-third of all cases, no aetiological factor can be found, and these patients are classified as having idiopathic disease. Pathophysiologically, autodigestion and inflammation may be caused by either increased proteolytic activity or decreased protease inhibition. Several studies have demonstrated mutations in the cationic trypsinogen gene (*PRSS1*) in patients with hereditary[1–3] or idiopathic[4] CP. It is thought that these mutations result in increased trypsin activity within the pancreatic parenchyma. Most patients with idiopathic or hereditary CP, however, do not have mutations in *PRSS1* (ref. 4). Here we analysed 96 unrelated children and adolescents with CP for mutations in the gene encoding the serine protease inhibitor, Kazal type 1 (*SPINK1*), a pancreatic trypsin inhibitor. We found mutations in 23% of the patients. In 18 patients, 6 of whom were homozygous, we detected a missense mutation of codon 34 (N34S). We also found four other sequence variants. Our results indicate that mutations in *SPINK1* are associated with chronic pancreatitis.

Serine protease inhibitor, Kazal type 1, also known as pancreatic secretory trypsin inhibitor (PSTI), is a potent protease inhibitor that is thought to be an inactivation factor of intrapancreatic trypsin activity. Human *SPINK1* is approximately 7.5 kb, has 4 exons and is located on chromosome 5 (ref. 5). The gene product consists of 79 amino acids, which include a 23 amino acid signal peptide. SPINK1 possesses a reactive site (a lysine residue at position 41 and an isoleucine residue at position 42, corresponding to positions 18 and 19 of the mature peptide) that serves as a specific target substrate for trypsin[6]. Incubation of equimolar quantities of trypsin and the inhibitor produces a complex with a covalent bond between the catalytic serine residue of the enzyme and the lysine carboxyl group of the reactive site of SPINK1.

We found a mutation in *PRSS1* in only 11 of 96 unrelated patients with CP (5×A16V, 1×N29I, 5×R122H), suggesting that mutations in other genes might be involved in CP. Because SPINK1 is an intrapancreatic trypsin inhibitor, we analysed *SPINK1* for mutations. We investigated the coding exonic and flanking intronic sequences by direct DNA sequencing, and detected mutations in 22 of 96 patients. In 18 patients with CP, we found an A→G transition resulting in a substitution of asparagine by serine at codon 34 in exon 3 (N34S, corresponding to codon 11 of the mature peptide). Six patients were homozygous for the N34S mutation, which in all cases was inherited from heterozygous parents. None of the patients showed a mutation in *PRSS1*. We found 12 patients with CP to be heterozygous for this mutation. We observed no other mutation in *SPINK1* in these patients.

There were no phenotypic differences between heterozygous and homozygous N34S patients. We examined 279 healthy controls for the N34S mutation by direct sequencing or melting curve analysis, and detected no homozygous individuals and only one heterozygous control individual. To exclude compound heterozygosity in the 12 individuals heterozygous for the N34S mutation, we sequenced the complete intronic sequence of *SPINK1* after long-range PCR using five different sets of oligonucleotides (Fig. 1). Analysis of the introns revealed that the N34S mutation was in complete linkage disequilibrium with four other sequence variants (IVS1-37 T→C, IVS2+268 A→G, IVS3-604 G→A, IVS3-69 insTTTT). Because we detected one of these variants in the heterozygous state in all of the five amplified fragments in each of the 12 patients, we were able to rule out gross deletions or insertions of any part of the coding or intronic sequence of *SPINK1*. The N34S mutation located near the reactive Lys-Ile site of SPINK1 might lead to a decreased inhibitory capacity. The asparagine at position 34 is not fully conserved between various species[7]. The porcine trypsin inhibitor, which shows a 70% identity in the amino acid sequence, also bears a serine residue at this position, but does not inhibit human trypsin effectively[8].

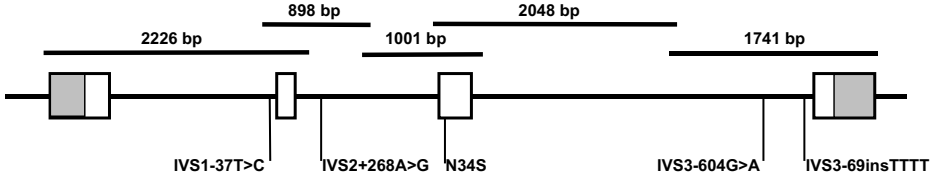

Fig. 1 Schematic presentation of *SPINK1* genomic structure, and length and position of the five amplified fragments for analysis of the intronic sequences. Sequence variations in introns associated with N34S are depicted below the line. Exons are represented by boxes.

Analysis of *SPINK1* in the available parents (n=29) demonstrated a strong transmission disequilibrium for the N34S mutation (one-sided *P* value=0.0000038). Analysis of

[1]*Department of Pediatrics, Charité, Campus Virchow-Klinikum, Humboldt University, Berlin, Germany.* [2]*Department of Molecular Genetics and Gene Mapping Center, Max-Delbrück-Centrum, Berlin, Germany.* [3]*Department of Pediatrics, Zentralkrankenhaus 'Links der Weser', Bremen, Germany.* [4]*Institute of Laboratory Medicine and Pathobiochemistry, Charité, Campus Virchow-Klinikum, Humboldt University, Berlin, Germany.* [5]*TIB MOLBIOL, Berlin, Germany. Correspondence should be addressed to H.W. (e-mail: heiko.witt@charite.de).*

nucleotide databases by sequence similarity searches showed that *SPINK1* is contained in the genomic BAC clones CIT-HSPC_327F10 and CIT978SKB_38B5, which also include *D5S2090*. We analysed 9 microsatellites in an interval of 12.3 cM around *D5S2090* in the CP patients with the N34S mutation, and their parents, by the transmission disequilibrium test (TDT) (ref. 9) and multi-locus haplotype relative risk (HRR) test[10,11]. Both TDT (data not shown) and the HRR test (Fig. 2) revealed linkage disequilibrium between marker alleles and the disease. A multi-locus HRR lod score greater than 3 was obtained in a small interval of 1 cM on either side of *D5S2090* with a maximum lod score of 5.4 (Fig. 2). These findings indicate that a gene involved in CP must reside very close to *D5S2090* in these patients. *SPINK1* is less than 200 kb from *D5S2090* and no other candidate gene maps to this region, indicating that *SPINK1* mutations are associated with CP.

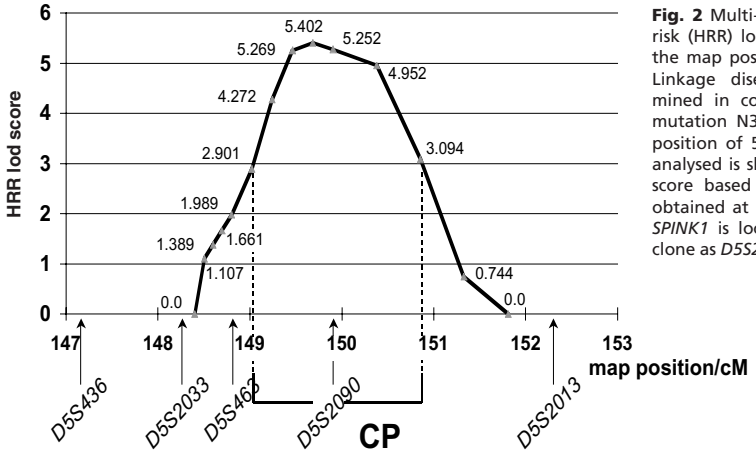

Fig. 2 Multi-locus haplotype relative risk (HRR) lod scores plotted against the map position on chromosome 5. Linkage disequilibrium was determined in core families segregating mutation N34S in *SPINK1*. The map position of 5 of 9 microsatellite loci analysed is shown[16]. A maximum lod score based on a χ^2 HRR test was obtained at 149.7 cM near *D5S2090*. *SPINK1* is located on the same BAC clone as *D5S2090*.

We detected a heterozygous mutation destroying the start codon of *SPINK1* (2T→C, M1T) in one patient with hereditary CP. This mutation was found in the affected grandfather and in the unaffected father. The M1T mutation destroys the only translation initiation codon of *SPINK1*. Another patient has a mutation affecting the T at position 2 of the splice donor site, which is highly conserved in eukaryotes (IVS3+2T→C). In one patient we detected a T→C transition in exon 1, leading to a substitution of leucine by proline in the signal peptide (41T→C, L14P). In another patient we found a C→T transition in the 5′ UTR at position –53 relative to the ATG initiation codon. This alteration may lead to a new initiation codon (ACG→ATG) with premature chain termination due to a TAA stop codon located two codons downstream. We cannot rule out that these four alterations are rare variants, but we found none of these four mutations when sequencing 52 unrelated healthy controls. We found eight other nucleotide alterations in non-coding sequences in patients as well as in controls that probably represent polymorphisms (Fig. 3 and Table 1). Moreover, in one control subject, we detected a C→T transition in exon 3 leading to a proline-to-serine substitution at amino acid residue 55 (position 32 of the mature peptide).

It is generally believed that inappropriate intrapancreatic activation of zymogens by trypsin leads to autodigestion and pancreatitis. Two different protective mechanisms prevent activation of the pancreatic digestive enzyme cascade. First, SPINK1 inhibits up to 20% of potential trypsin activity[8,12]. Second, trypsin itself activates trypsin-like enzymes, such as mesotrypsin, readily degrading trypsinogen and other zymogens[13] (Fig. 4*a*). Pancreatitis may be the result of an imbalance of proteases and their inhibitors within the pancreatic parenchyma (Fig. 4*b*). We have demonstrated for the first time the association of a mutated inhibitor with CP. CP may be classified as a disease secondary to well-recognized factors (for example, alcohol abuse) and as hereditary and idiopathic disease. Previous results indicated an association between idiopathic CP and mutations in the cystic fibrosis transmembrane conductance regulator gene[14,15] (*CFTR*) or *PRSS1* (ref. 4), suggesting a genetic basis for the so-called idiopathic CP. The presence of *SPINK1* mutations in patients with idiopathic CP further challenges the concept of idiopathic CP as a non-genetic disorder and the differentiation between hereditary and idiopathic CP. Our data suggest that patients with CP may suffer from autosomal recessive disease as a result of mutations of *SPINK1* or as a result of a combination of genetic defects of *SPINK1* and yet unidentified genes. Finally, our study may have implications for developing therapeutic strategies to prevent or control CP by substitution or augmentation of *SPINK1*.

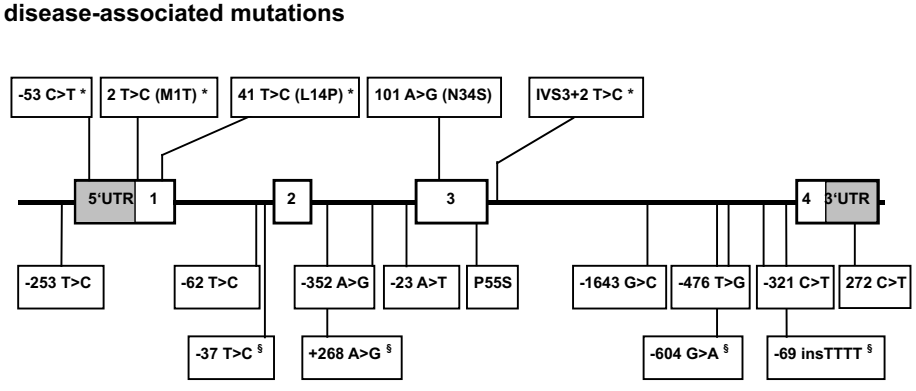

Fig. 3 Schematic presentation of the *SPINK1* genomic structure and sequence variations found in CP patients. Exons are represented by numbered boxes. Numbering of nucleotides begins with the ATG initiation codon. *This variation was found in only one patient. Its relevance remains unclear. §These sequence alterations are in complete linkage disequilibrium with N34S.

Table 1 • Mutations and polymorphisms in *SPINK1*

Sequence alterations	Amino acid	Location	Frequency	Allele status
Mutations associated with CP				
–53C→T	–	5′ UTR	1 p	ht
2T→C	M1T	exon 1	1 p	ht
41T→C	L14P	exon 1	1 p	ht
101A→G	N34S	exon 3	18 p (0.13)	hm & ht
IVS3+2T→C	–	intron 3	1 p	ht
Polymorphisms				
–253T→C	–	promotor	p (0.15) & c (0.11)	hm & ht
IVS1-62T→C[a]	–	intron 1	1 p	ht
IVS1-37T→C[b]	–	intron 1	18 p (0.13)	hm & ht
IVS2+268A→G[b]	–	intron 2	18 p (0.13)	hm & ht
IVS2-352A→G	–	intron 2	p (0.08) & c (0.41)	hm & ht
IVS2-23A→T	–	intron 2	2 c & 1 p	ht
163C→T	P55S	exon 3	1 c	ht
IVS3-1643G→C	–	intron 3	p (0.13) & c (0.27)	hm & ht
IVS3-604G→A[b]	–	intron 3	18 p (0.13)	hm & ht
IVS3-476T→G	–	intron 3	p (0.04) & c (0.05)	hm & ht
IVS3-321C→T	–	intron 3	p (0.13) & c (0.39)	hm & ht
IVS3-69insTTTT[b]	–	intron 3	18 p (0.13)	hm & ht
272C→T	–	3′ UTR	p (0.03) & c (0.05)	hm & ht

ht, Heterozygous; hm, homozygous; p, patients; c, controls. Allele frequencies are given in parentheses. [a]The IVS1-62 T→C variation was found in a patient heterozygous for the N34S mutation on the same allele. [b]These alterations are in complete linkage disequilibrium with N34S.

Methods

Patients. We enrolled 96 unrelated children and adolescents with chronic pancreatitis for genetic analysis. Because alcohol abuse is uncommon in children and adolescents, this age group was thought to be suitable for a prevalence study of genetic defects in *SPINK1*. Patients originated from Germany and Austria. Chronic pancreatitis was arbitrarily defined as a condition characterized by at least two episodes of recurring abdominal pain, increase in serum lipase and/or amylase levels, and pathological sonographic findings. We classified 68 patients (71%) as having idiopathic disease after exclusion of cystic fibrosis and metabolic, anatomic, traumatic, toxic or infectious causes. There were 28 patients (29%) with a family history of chronic pancreatitis. Fifty-three patients were female (55%) and 43 male (45%). Blood for genetic analysis was available from both parents of 41 patients and from one parent for 12 patients. Informed consent was obtained. Fifty-two unrelated medical students or staff of German origin served as controls for sequencing analysis after giving informed consent. For melting curve analysis, another 227 healthy Germans served as controls.

Mutation screening. We extracted genomic DNA from peripheral blood leukocytes according to standard protocols (Qiagen). We designed primers from intronic sequences flanking the four *SPINK1* exons based on the published nucleotide sequence[5] or results of sequencing the intronic sequences after intron-spanning long-range PCR. Primer sequences used for PCR amplification were as follows: promotor region, 5′–TTTGAGTTCATCT TACAGGTGAG–3′ and 5′–TATGGCAGATGGCAGCAAGG–3′; 5′ UTR and exon 1, 5′–CCAGGCTATGACACAGAGTC–3′ and 5′–GTGCTTCA CAAAGCAACAGGTC–3′; exon 2, 5′–GGGTGGGGAATGAAAGAGCC–3′ and 5′–AAAGGTGACAGCAAGGCTGC–3′; exon 3, 5′–CCAATCACAGT TATTCCCCAGAG–3′ and 5′–GTTTGCTTTTCTCGGGGTGAG–3′; exon 4, 5′–CCCTGTTTTTCTCCCATAGTCAC–3′ and 5′–CCAAAGTCCCCT-GACCCTGG–3′. We performed PCR using AmpliTaq Gold polymerase (0.5 U; Perkin Elmer), deoxynucleoside triphosphates (400 µmol/l) and primer (0.1 µmol/l each) in a total volume of 25 µl. Cycle conditions were as follows: initial denaturation for 12 min at 95 °C; 36 cycles of 15 s denaturation at 95 °C, 30 s of annealing at 56 °C and 30 s of primer extension at 72 °C; and a final extension for 2 min at 72 °C. We purified PCR products with spin columns and performed cycle sequencing using purified template (1.5 µl), BigDye terminator mix (3 µl; Applied Biosystems) and 3 µl forward or reverse primer (0.3 µmol/l). The reaction products were purified with spin columns and loaded onto an ABI 373A fluorescence sequencer (Applied Biosystems). We analysed DNA sequences by sequencing both strands of the entire coding region of all patients and all controls. To exclude mutations in

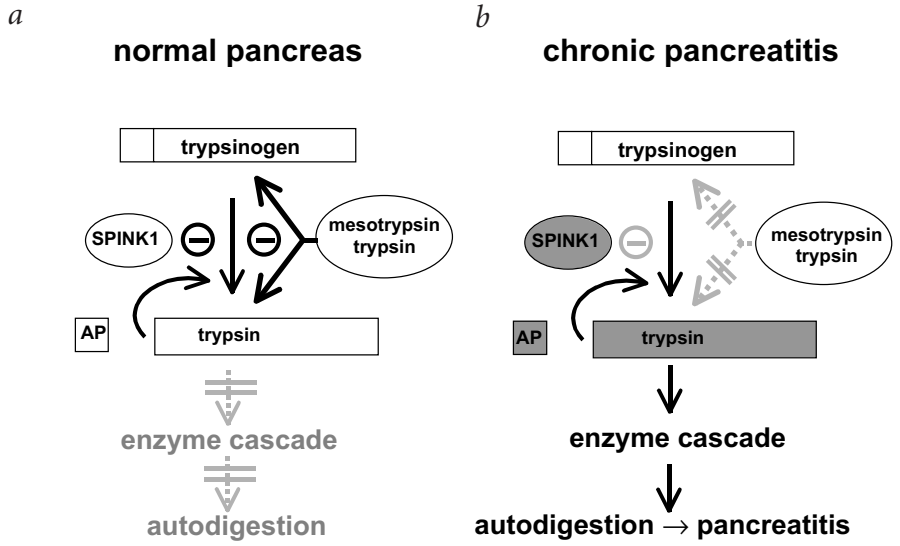

Fig. 4 Model of chronic pancreatitis. ***a***, Normal pancreas. Trypsin resulting from autoactivation of trypsinogen in the pancreatic parenchyma is inhibited by *SPINK1* and by mesotrypsin or trypsin. This defence mechanism prevents the pancreas from activation of the pancreatic enzyme cascade and autodigestion. ***b***, Chronic pancreatitis. Mutations in *SPINK1* or in *PRSS1* lead to an imbalance of proteases and their inhibitors within the pancreatic parenchyma, resulting in an inappropriate conversion of pancreatic zymogens to active enzymes with autodigestion and inflammation. Filled boxes represent products of mutated genes (modified from ref. 1; AP, activation peptide).

the intervening sequences or gross deletions or insertions, we analysed the intronic sequences of the 12 individuals heterozygous for the N34S mutation by direct DNA sequencing after intron-spanning, long-range PCR.

Melting curve analysis. We performed melting curve analysis for the N34S mutation using a pair of fluorescence resonance energy transfer (FRET) probes and the LightCycler (Roche Diagnostics). Primers for PCR of exon 3 are described above. The donor probe was 5′–CCAAATGTTACAATGAACTTAATGGATGC-FL–3′ (FL, 5,6-carboxyfluorescein attached to 3′-O-ribose) and the acceptor probe was 5′–LC-CCAAGATATATGACCCTGTCTGTGGGAC-ph–3′ (LC, LightCycler Red 640 attached to 5′ terminus; ph, 3′ phosphate). The donor probe was complementary to the wild-type sequence. During melting curve analysis, a more stable duplex was formed with the wild-type allele than with the mutant allele, resulting in an allele-specific melting curve (65 °C versus 60 °C).

Microsatellite analysis. Microsatellites at *D5S2010*, *D5S2017*, *D5S436*, *D5S2033*, *D5S463*, *D5S2090*, *D5S2013*, *D5S640* and *D5S410*, which cover an interval of 12.3 cM around *SPINK1*, were chosen from the Généthon human linkage map[16]. We performed PCR in a final reaction volume of 10 µl with dNTPs (100 µM each), DNA polymerase (0.4 U; Perkin Elmer), primer (0.7 µM each) and genomic DNA (20 ng). One primer of each primer pair was end-labelled with fluorescent dye. Single-plex reactions were pooled and electrophoresed on an ABI 377 automatic sequencer. For data analysis we used the computer programs Genescan v3.0 and Genotyper v2.5 (Perkin Elmer).

Linkage disequilibrium analysis. Transmission disequilibrium[9] was assessed for each marker using the computer program TDTLIKE developed by J. Terwilliger[11]. We analysed the multi-locus haplotype relative risk[10] (HRR), which compares alleles transmitted and non-transmitted at each locus jointly in a multi-locus test, with the program HRRMULT (ref. 11). Results are presented as lod scores based on a χ^2 HRR test.

Acknowledgements

We thank the families for participation; the members of the Gesellschaft für Pädiatrische Gastroenterologie und Ernährung for providing clinical data and blood samples; C. Güldner and I. Liebner for technical assistance; and F. Rüschendorf for support in computer science.

Received 7 September 1999; accepted 24 March 2000.

1. Whitcomb, D.C. *et al*. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nature Genet.* **14**, 141–145 (1996).
2. Gorry, M.C. *et al*. Mutations in the cationic trypsinogen gene are associated with recurrent acute and chronic pancreatitis. *Gastroenterology* **113**, 1063–1068 (1997).
3. Ferec, C. *et al*. Mutations in the cationic trypsinogen gene and evidence for genetic heterogeneity in hereditary pancreatitis. *J. Med. Genet.* **36**, 228–232 (1999).
4. Witt, H., Luck, W. & Becker, M. A signal peptide cleavage site mutation in the cationic trypsinogen gene is strongly associated with chronic pancreatitis. *Gastroenterology* **117**, 7–10 (1999).
5. Horii, A. *et al.* Primary structure of human pancreatic secretory trypsin inhibitor (PSTI) gene. *Biochem. Biophys. Res. Commun.* **149**, 635–641 (1987).
6. Bartelt, D.C., Shapanka, R. & Greene, L.J. The primary structure of the human pancreatic secretory trypsin inhibitor. Amino acid sequence of the reduced S-aminoethylated protein. *Arch. Biochem. Biophys.* **179**, 189–199 (1977).
7. Marchbank, T., Freeman, T.C. & Playford, R.J. Human pancreatic secretory trypsin inhibitor. *Digestion* **59**, 167–174 (1998).
8. Rinderknecht, H. Pancreatic secretory enzymes. *in The Pancreas: Biology, Pathobiology, and Disease* (eds Go, V.L.W. *et al.*) 219–251 (Raven, New York, 1993).
9. Spielman, R.S., McGinnis, R.E. & Ewens, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52**, 506–516 (1993).
10. Falk, C.T. & Rubinstein, P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann. Hum. Genet.* **51**, 227–233 (1987).
11. Terwilliger, J.D. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56**, 777–787 (1995).
12. Laskowski, M. Jr & Kato, I. Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593–626 (1980).
13. Rinderknecht, H. Activation of pancreatic zymogens. Normal activation, premature intrapancreatic activation, protective mechanisms against inappropriate activation. *Dig. Dis. Sci.* **31**, 314–321 (1986).
14. Sharer, N. *et al.* Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. *N. Engl. J. Med.* **339**, 645–652 (1998).
15. Cohn, J.A. *et al.* Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. *N. Engl. J. Med.* **339**, 653–658 (1998).
16. Dib, C. *et al.* A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**, 152–154 (1996).