Am. J. Hum. Genet. 60:342–351, 1997

Identification of Mutations in *Cystatin B*, the Gene Responsible for the Unverricht-Lundborg Type of Progressive Myoclonus Epilepsy (EPM1)

Maria D. Lalioti,[1] Maria Mirotsou,[1] Catherine Buresi,[2] Manuel Claude Peitsch,[3] Colette Rossier,[1] Réda Ouazzani,[5] Michel Baldy-Moulinier,[6] Armand Bottani,[4] Alain Malafosse,[2] and Stylianos E. Antonarakis[1,4]

[1]Laboratory of Human Molecular Genetics, Department of Genetics and Microbiology, University of Geneva Medical School, [3]Division de Neuropsychiatrie, Hôpital Belle-Idée, [3]Glaxo Institute for Molecular Biology, and [4]Division of Medical Genetics, University Hospital of Geneva, Geneva; [5]Service de Neurologie et de Neurophysiologie Clinique, Hôpital de Spécialités, Rabat, Morocco; and [6]Laboratoire de Médecine Expérimentale, CNRS UPR 9008, INSERM U249, Montpellier, France

Summary

Progressive myoclonus epilepsy (EPM1) is an autosomal recessive disorder, characterized by severe, stimulus-sensitive myoclonus and tonic-clonic seizures. The *EPM1* locus was mapped to within 0.3 cM from *PFKL* in chromosome 21q22.3. The gene for the proteinase inhibitor cystatin B was recently localized in the EPM1 critical region, and mutations were identified in two EPM1 families. We have identified six nucleotide changes in the *cystatin B* gene of non-Finnish EPM1 families from northern Africa and Europe. The 426G→C change in exon 1 results in a Gly4Arg substitution and is the first missense mutation described that is associated with EPM1. Molecular modeling predicts that this substitution severely affects the contact of *cystatin B* with papain. Mutations in the invariant AG dinucleotides of the acceptor sites of introns 1 and 2 probably result in abnormal splicing. A deletion of two nucleotides in exon 3 produces a frameshift and truncates the protein. Therefore, these four mutations are all predicted to impair the production of functional protein. These mutations were found in 7 of the 29 unrelated EPM1 patients analyzed, in homozygosity in 1, and in heterozygosity in the others. The remaining two sequence changes, 431G→T and 2575A→G, probably represent polymorphic variants. In addition, a tandem repeat in the 5′ UTR (CCCCGCCCCGCG) is present two or three times in normal alleles. It is peculiar that in the majority of patients no mutations exist within the exons and splice sites of the *cystatin B* gene.

Received August 13, 1996; accepted for publication November 4, 1996.

Address for correspondence and reprints: Prof. Stylianos E. Antonarakis, Division de Génétique Médicale, Centre Medical Universitaire, 1 rue Michel-Servet, CH-1211 Geneva 4, Switzerland. E-mail: sea@medsun.unige.ch

0002-9297/97/6002-0013$02.00

Introduction

Progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1) is an autosomal recessive disorder characterized by severe, stimulus-sensitive myoclonus and tonic-clonic seizures (MIM 254800). The age at onset of the disease is between 6 and 13 years, and its severity and progression are variable. Mental deterioration and, finally, cerebellar ataxia develop. Histologically, the brain shows degenerative changes without Lafora bodies and marked loss of Purkinje cells of the cerebellum (Norio and Koskiniemi 1979). EPM1 is a rare disorder in the general population but relatively more common in Finland (1 in 20,000; Norio and Koskiniemi 1979) and the western Mediterranean (Genton et al. 1990).

Linkage analysis and disequilibrium studies in families from Finland and other parts of the world have mapped the *EPM1* locus to the terminal part of human chromosome 21q (Malafosse et al. 1992; Cochius et al. 1993; Lehesjoki et al. 1993*a*, 1993*b*). Refined mapping limited the candidate region to a 175-kb interval between markers *D21S2040* and *D21S1259* (Virtaneva et al. 1996). Since the biochemical defect was unknown, positional cloning strategies were used to identify the gene. The EPM1 critical region was poorly represented by YACs in the CEPH contig (Chumakov et al. 1992), and contigs were subsequently constructed with cosmids, P1 clones, P1 artificial chromosomes, and bacterial artificial chromosomes (Lafrenière et al. 1995; Stone et al. 1996). At least two genes have been proposed as candidates for *EPM1*, but mutation analysis have excluded them from being responsible for the disorder (Yamakawa et al. 1995; Lalioti et al. 1996). The gene has been cloned by cDNA selection and proved to be *cystatin B*, a previously cloned but unmapped gene (Pennacchio et al. 1996). A reduced amount of *cystatin B* mRNA was observed in affected individuals, compared to normal controls and carriers. Two mutations were identified: a G→C transversion at the last nucleotide of intron 1, which was likely to affect the splicing of this intron, and a CGA to TGA mutation, which changes the codon for

Arg68 to a termination codon. However, no sequence alteration has been identified in the Finnish families, which, according to the haplotype data, have a single common mutation (Pennacchio et al. 1996).

Cystatin B (also known as stefin B) is a member of a large family of inhibitors of cysteine proteinases (reviewed by Turk and Bode 1991). They bind tightly to and inhibit the action of papain and papain-like proteinases. These inhibitors might protect the cells from inappropriate endogenous or external proteolysis and/or could be involved in the control of intracellular or extracellular protein breakdown.

We have used SSCA (single-stranded conformation analysis) and direct PCR sequencing, in an attempt to identify the mutations in EPM1 patients from Switzerland, France, and north Africa. We have identified four disease-producing mutations and three polymorphic variants. In the majority of patients, no mutations were detected in the coding regions and exon boundaries of the *cystatin B* gene, as was the case in the study of Pennacchio et al. (1996). The mutations in these families are presumed to lie outside the coding regions, in currently unknown regulatory elements, thereby disrupting the normal expression of the gene.

Patients, Material, and Methods

Patient Selection and Strategy

A total of 29 apparently unrelated patients were studied. Eighteen were of north African origin, 9 were from France, 1 was of Finnish (paternal side) and French (maternal side) origins, and 1 was Swiss. Consanguinity was documented in 14 of the north African families and in the Swiss family, in all of whom linkage and haplotype analysis confirmed linkage to 21q22.3; one common haplotype was shared by 6 of the North African families (Lalioti et al. 1995; P. Labauge, R. Quazzani, A. Mrabet, D. Grid, P. Genton, C. Dravet, T. Chkili, et al., unpublished data). These data suggest that more than one mutation should exist in the group of patients examined.

All patients were diagnosed after neurological examination using the following criteria: age at onset between 6 and 16 years; tonic-clonic seizures, often presenting as the first symptom; stimulus-sensitive myoclonus; characteristic EEG findings with generalized spike-wave and polyspike-wave paroxysms; sensitivity to photic stimulation; and progressive course. Skin biopsy material of at least one patient in each family was studied; no Lafora bodies were detected. The origin and the clinical characteristics of the affected individuals in which *cystatin B* mutations were found are summarized in table 1.

DNA and RNA Isolation

Lymphoblastoid cell lines from EPM1 patients were cultured in RPMI supplemented with 10% fetal bovine serum (FBS). Fibroblasts were cultured in DMEM supplemented with 10% FBS. DNA from cultured cells was prepared using standard techniques. In brief, cells were harvested by trypsinization (for fibroblasts) or centrifugation (for lymphoblasts) and incubated in 100 mM NaCl, 25 mM EDTA pH 8, 0.1mg/ml proteinase K, and 0.2% SDS overnight at 50°C. DNA was then extracted with phenol/chloroform, precipitated with ethanol and 0.4 M NaCl, and resuspended in Tris-EDTA. Extraction of DNA from peripheral blood lymphocytes was performed using the same procedure as described for the cultured cells. Total RNA from cultured cells was obtained after cell lysis with Trizol (Life Technologies, Inc.), and poly(A)$^+$ RNA was isolated using the FastTrack® 2.0 kit (Invitrogen).

Northern Blot Analysis

Northern blot analysis of poly(A)+ RNA from lymphoblastoid lines from two unrelated patients (GVA05 and GVA07), the father of GVA07, and from three normal individuals was performed by standard techniques. The filter was hybridized with the entire *cystatin B* cDNA, stripped, and rehybridized with a β-actin probe to permit normalization of the amount of RNA in each lane. The intensities of the *cystatin B* and β-actin signals were quantified with a Molecular Dynamics phosphorimager. The mean values from two independent experiments are reported here.

PCR and SSCA

The *cystatin B* gene (GenBank accession no. U46692) was screened by SSCA as shown in figure 1, with the PCR primers shown in table 2. Exon 1 was amplified for SSCA with primers CSTB.pF4 and CSTB.4R and for sequencing using CSTB.pF4 and CSTB.R14. Amplification was carried out in a 30-cycle PCR, in which the initial 5-min denaturation of template DNA at 94°C was followed by a "touch-down" program for 10 cycles: 94°C/30 s, 65°C/20 s (−1°C/cycle), and 72°C/2 min and then 20 cycles: 94°C/30 s, 55°C/20 s, and 72°C/2 min, in a volume of 20 μl containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 0.2 mM of each dNTP (where half of the dGTP was substituted with 7-deaza-2′-dGTP), 1.5 mM $MgCl_2$, 0.5 μM of each primer, 10% dimethylsulfoxide (DMSO), and 0.5 units of *Taq* polymerase. Exon 2 was amplified for SSCA with primers CSTB.F11 and CSTB.5R, and exon 3 with primers CSTB.4L and CSTB.dR3. For sequencing, exons 2 and 3 and intron 2 were amplified in a single PCR with CSTB.F11 and CSTB.1R. PCR amplifications were done as above but with 0.05 mM of each dNTP, 5% DMSO, and a shorter extension time (1 min).

For the SSCA, primers were labeled by T4 polynucleotide kinase (Pharmacia) with γ-^{32}P-ATP. After PCR, two volumes of loading buffer (95% formamide, 10 mM NaOH, 0.25% bromophenol blue, 0.25% xylene cya-

Table 1

Patients, Mutations, and Sequence Variants Detected

Patient	Allele 1	Allele 2[a]	Origin	Handicap[b]
GVA01	426G→G	426G→C	Morocco	2
GVA02	1924G→C	ni	Southern France	2
GVA03	1924G→C	2575A→G	Southern France	3
GVA04	1924G→C	{ ...	France (maternal side)	3
		{ ni	Finland (paternal side)	...
GVA05	1924G→C	ni	Central France	2
GVA06	2352A→G	ni	Western France	3
GVA07	del2399TC	ni	Western France	3
GVA08[c]	2575A→G	2575A→G	Switzerland	3–4
GVA09[c]	2575A→G	2575A→G	Central France	4
GVA10[c]	431G	431T	Spain	...

[a] ni = not identified.
[b] 1 = normal; 2 = moderate, autonomous; 3 = need help; 4 = severely affected.
[c] The changes in the last three GVA08–10 individuals are most likely nonpathogenic sequence variants.

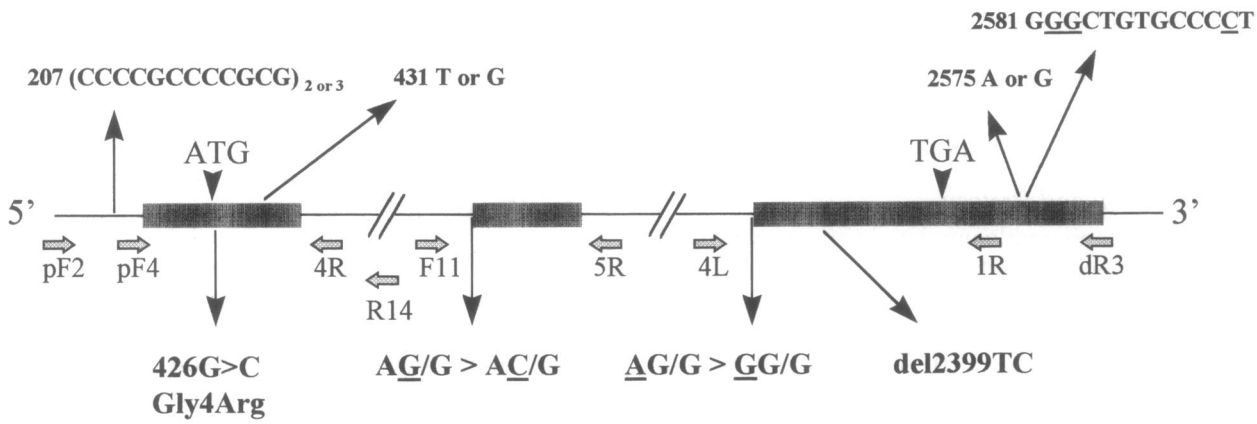

Figure 1 Schematic representation of the *cystatin B* gene (exons are depicted as filled boxes; the translation initiation and termination codons are shown in exons 1 and 3, respectively). The oligonucleotide primers used for PCR amplification, SSCA, and sequencing (see Patients, Material, and Methods) are shown below the gene. The disease-causing mutations found in this study are below the primers. The polymorphisms identified in this study are depicted above the gene.

Table 2

Oligonucleotide Primers Used for PCR, SSCA, and Sequencing

Name	Sequence	Strand	Position
CSTB.pF2	ctc cga ctg ccc ctt ccc tat c	+	5′ UTR (18–39)
CSTB.pF4	cgc agg gga ctc cga agc caa agt g	+	5′ UTR (272–296)
CSTB.4R	caa agc ggc ttc ttt cgc ctc	−	Intron 1 (572–553)
CSTB.R14	cca ctt tta gca aga agc ct	−	Intron 1 (940–918)
CSTB.F11	cca ccg tac cca gct gga act g	+	Intron 1 (1728–1749)
CSTB.5R	cct gca cag gcg gtt tcc tac	−	Intron 2 (2089–2070)
CSTB.4L	ggc gca gca agg tga ctt gg	+	Intron 2 (2299–2318)
CSTB.1R	gat cac agg tgc acg ctc tg	−	3′ UTR (2560–2541)
CSTB.dR3	gtc taa aag caa agg agg caa tc	−	3′ UTR (2756–2734)

Table 3

Mutations Identified and Enzyme Detection

Mutation and Consequence	Position[a]	Enzyme Detection
426G→C Gly4Arg	Exon1	*Hpa*II (mutant)[b]; *Aci*I (normal)
1924G→C aberrant splicing	ivs1—acceptor splice site	*Bfa*I (normal)
2352A→G aberrant splicing	ivs2—acceptor splice site	*Bst*NI (normal)
del2399TC fs; truncated protein[c]	Exon 3	...
2575A→G	3′UTR	*Alw*26I (mutant)
431G or T—no amino acid change	Exon 1	*Kas*I (normal)
$(CCCGCCCGCG)_{2-3}$	5′UTR	...
2581GGGCTGTGCCCCT[d]	3′UTR	...

[a] ivs = intervening sequence (intron).
[b] Allele cut by the enzyme given in parentheses.
[c] fs = frameshift.
[d] The underlined nucleotides are present only in some individuals.

nol) were added, and the samples were denatured at 95°C for 10 min. PCR products were electrophoresed in a nondenaturing gel with two different sets of conditions for each exon. Condition 1: 0.5 × MDE gel (FMC BioProducts), 0.6 × TBE, 10% glycerol, and electrophoresis at room temperature. Condition 2: 0.5 × MDE gel, 0.6 × TBE, no glycerol, and electrophoresis at 4°C. Both types of gels were electrophoresed for 20,000 V/h at 20 W. The gel was subsequently dried and exposed overnight at −80°C. The PCR products were directly sequenced with primers CSTB.4R for exon 1, CSTB.5R for exon 2, and CSTB.4L for exon 3 and their splice junctions, in an automated sequencer ABI373 using standard protocols.

Mutant alleles of compound heterozygous patients were cloned using the Original TA cloning kit® (Invitrogen) after PCR amplification. Single colonies were picked and tested for the presence of the mutant allele with the enzymes presented in table 3. DNA from these clones was prepared with Qiagen columns and sequenced as before.

Results

Strategy for Mutation Identification

Because the *cystatin B* gene contains only 3 exons and the entire genomic nucleotide sequence is known, we decided to undertake SSCA to identify variations in the migration pattern of PCR-amplification products. We also used direct sequencing of genomic PCR products of all exons and their intron-exon junctions in all 29 patients (Patients, Material, and Methods; fig. 1; table 2).

A Gly4Arg Missense Mutation: 426G→C

By use of SSCA of exon 1, a variant band was apparent in the DNA of patient GVA01 (fig. 2*A*, lane 3). Sequencing revealed a homozygous G→C tranversion (GGG to CGG) at nt 426 in *cystatin B* (figs. 1 and 3*A*), leading to substitution of Gly4 by Arg in the protein. The Gly4 is highly conserved in all known cystatins from all species (fig. 4) (Turk and Bode 1991). The crystal structure of the *cystatin B*–papain complex has been resolved (Stubbs et al. 1990), revealing that the amino-terminal part of cystatin B that includes Gly4 interacts with papain (fig. 5*A*). The Gly4Arg substitution would modify the binding pocket by bringing in an amino acid with a long and charged side chain. Three-dimensional modeling suggests that the Gly4Arg mutation is likely to jeopardize and reduce or abolish the interaction of the two proteins (fig. 5*B*, *C*).

Splice-Site Mutations: 2352A→G and 1924G→C

SSCA and sequence analysis revealed two different splice-site mutations. The first was an A→G change (2352A→G) at the invariant AG dinucleotide of the acceptor splice site of intron 2 (AG/G to GG/G [figs. 1;

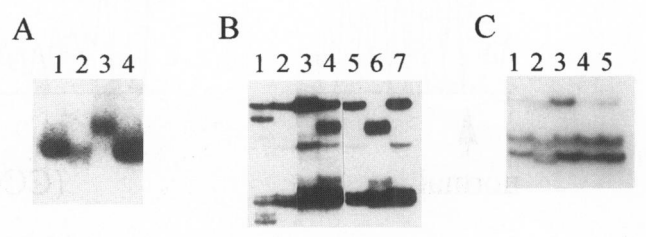

Figure 2 SSCA. For experimental details, see Patients, Material, and Methods. *A*, SSCA of exon 1 using condition 1. Lanes 1 and 4, Normal control DNAs. Lane 2, Patient heterozygous for the 431G/T polymorphism. Lane 3, Patient homozygous for the 426G→C mutation. *B*, SSCA of exon 3, using condition 2. Lane 1, Patient heterozygous for the 2352A→G mutation in the acceptor splice site AG dinucleotide. Lanes 2, 3, 5, and 7, Normal controls. Lane 4, Patient heterozygous for the 2575A→G 3′ UTR variant. Lane 6, Patient homozygous for the 2575A→G *C*, SSCA of exon 3, using condition 1. Lane 2, Patient heterozygous for the del2399TC frameshift mutation. Lanes 1 and 3–5, Normal controls.

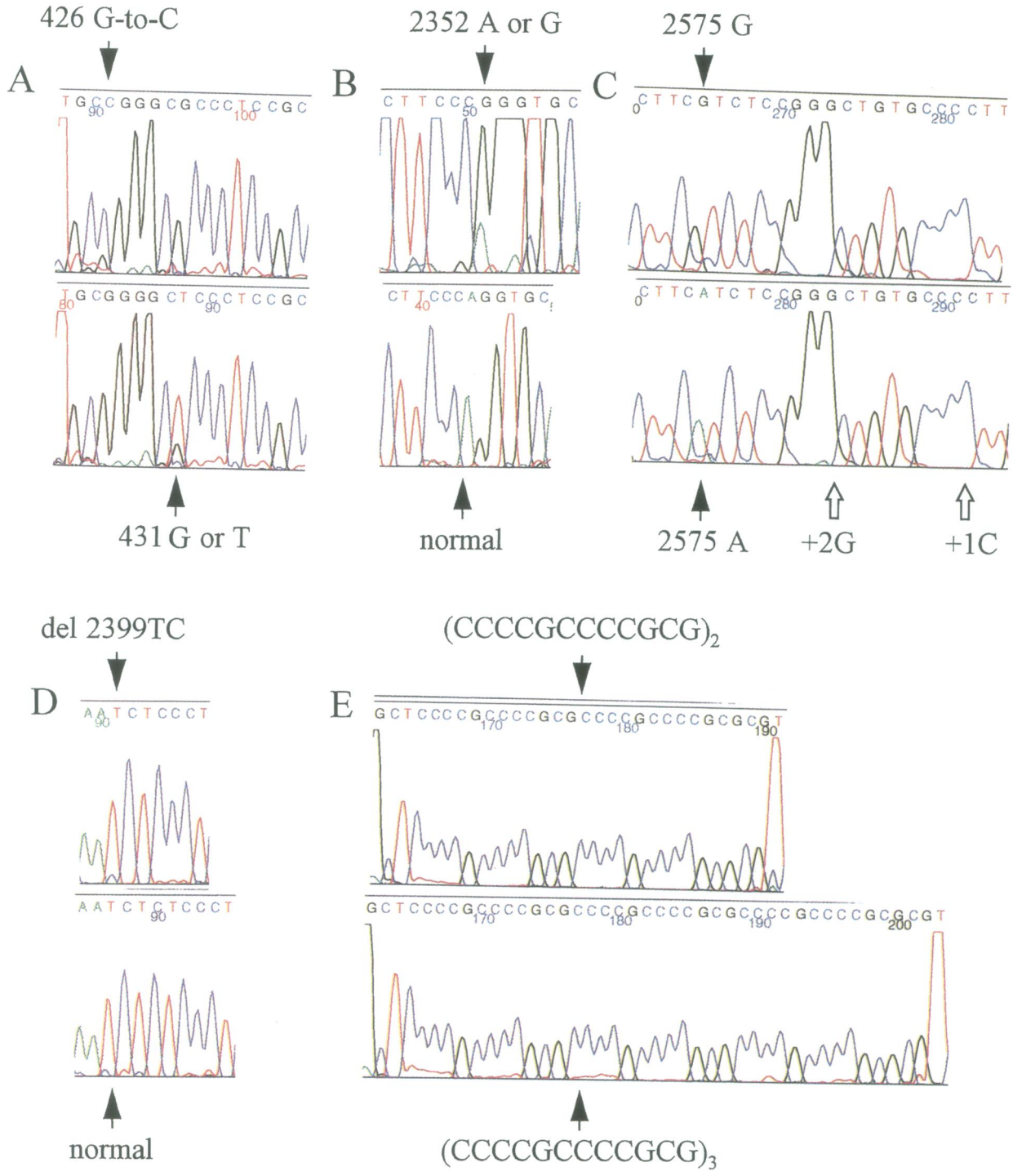

Figure 3 Nucleotide sequences of the newly identified mutations and polymorphisms. *A*, Direct genomic sequencing of the 426G→C missense mutation and the 431G/T polymorphism. *B*, Direct genomic sequencing of the 2352A→G splice-site mutation. *C*, Direct genomic sequencing of the 2575A→G 3′UTR variant. Open arrows indicate the further changes detected compared to the published sequence (see text). *D*, Nucleotide sequence of the cloned alleles from a patient heterozygous for the del2399TC frameshift mutation. *E*, Direct genomic sequencing from two normal individuals, showing that the sequence (CCCCGCCCCGCG)n is polymorphic in the general population and repeated two or three times.

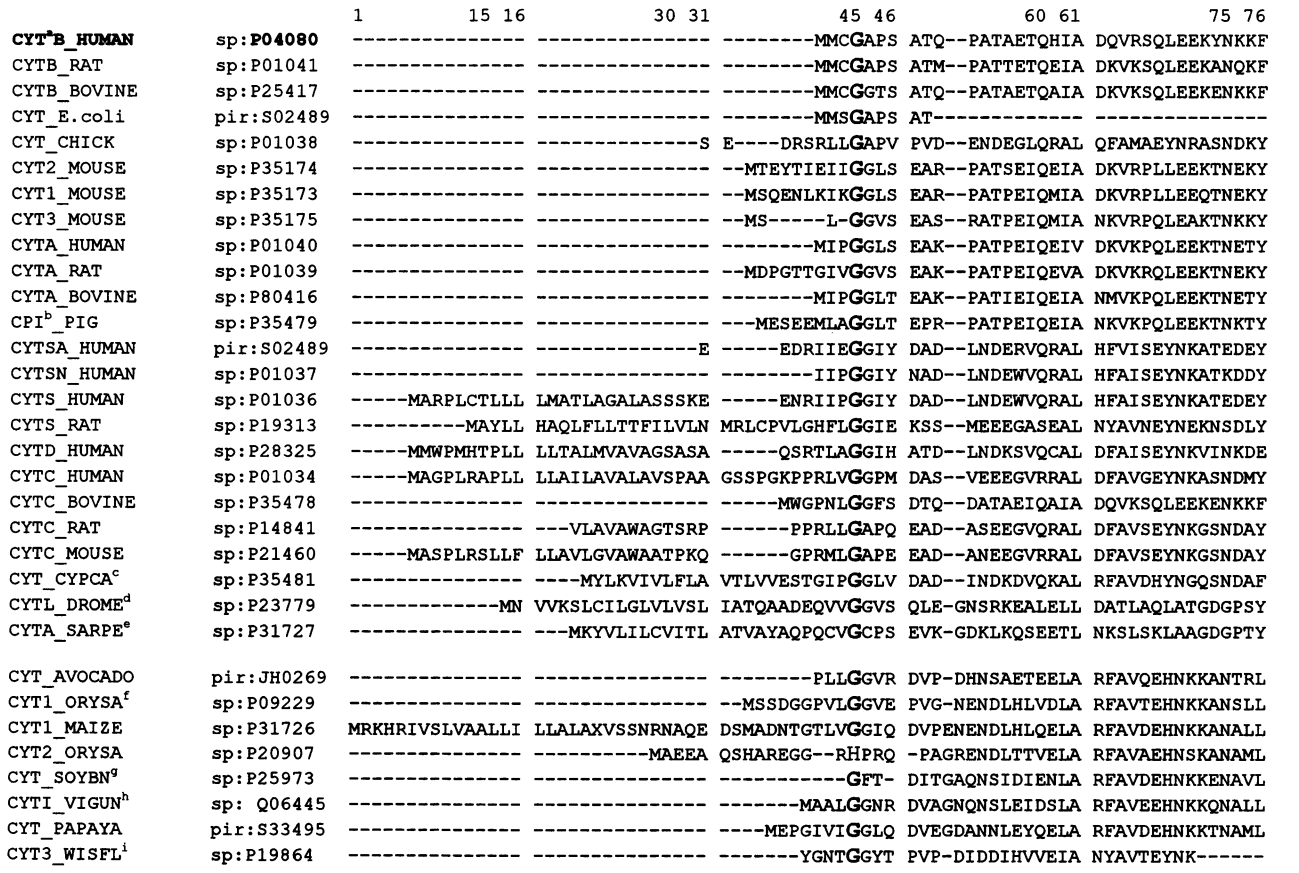

Figure 4 Amino acid sequence alignment of the amino-terminal end of different cystatin polypeptides, showing the conservation of Gly4. Sequences have been extracted from Swiss-prot (sp) or PIR (pir), and alignment was performed using the Clustal W program. The Swiss-prot and PIR numbers follow the names of the sequences. a = cystatin; b = cysteine-proteinase inhibitor; c = cyprinus; d = drosophila; e = flesh fly; f = rice; g = soybean; h = cowpea; i = wysteria (flower).

2*B*, lane 1, and 3*B*]). This mutation was found in one allele of patient GVA06 and is likely to result in abnormal splicing. It destroys a *Bst*NI restriction site, which was used to confirm the mutation in the patient's genomic DNA.

The second splice-site mutation was identical to that reported by Pennacchio et al. (1996), a G→C change (1924G→C) in the conserved AG dinucleotide of the 3′-splice site of intron 1 (fig. 1). This transversion probably also results in abnormal splicing. The mutation was found in heterozygosity in three apparently unrelated French patients (GVA02, GVA03, and GVA05) and in one patient of French-Finnish origin (GVA04). The mutation destroys a *Bfa*I site (table 3). The use of this diagnostic enzyme in the parents of the patient revealed that the splicing mutation originated from the French side of the family. Northern blot analysis of mRNA from a lymphoblastoid cell line from patient GVA05, who has the 1924G→C mutation in addition to an unidentified second mutation, indicated that the level of *cystatin B* mRNA was ~28% of the mean value of normal controls.

Frameshift Mutation: del2399TC

The DNA of patient GVA07 showed, after SSCA and sequencing, a deletion of two nucleotides in one allele (figs. 1; 2*C*, lane 2; and 3*D*). This deletion in exon 3 involves the TC dinucleotide starting at position 2399 in codon 72 of the coding region and occurs in the context of a short dinucleotide repeat AATCTCTCCC. It creates a frameshift that introduces a stop codon three codons downstream of Ser72, resulting in a polypeptide that is 23 amino acids shorter than the normal 98–amino acid cystatin B. Northern blot analysis of mRNA from a lymphoblastoid cell line from this patient (in whom the second mutation is unknown) indicated that the level of *cystatin B* mRNA was ~29% of the mean of normal controls. Furthermore, the father of this patient, who does not have the del2399TC mutation and therefore is presumed to carry the unidentified mutation, expressed only 53% of the normal level of *cystatin B* mRNA.

Other Nucleotide Changes: Predicted Polymorphisms

We have found a substitution, 2575A→G, in the 3′ UTR (figs. 1; 2*B*, lanes 4 and 6; and 3*C*). This alteration

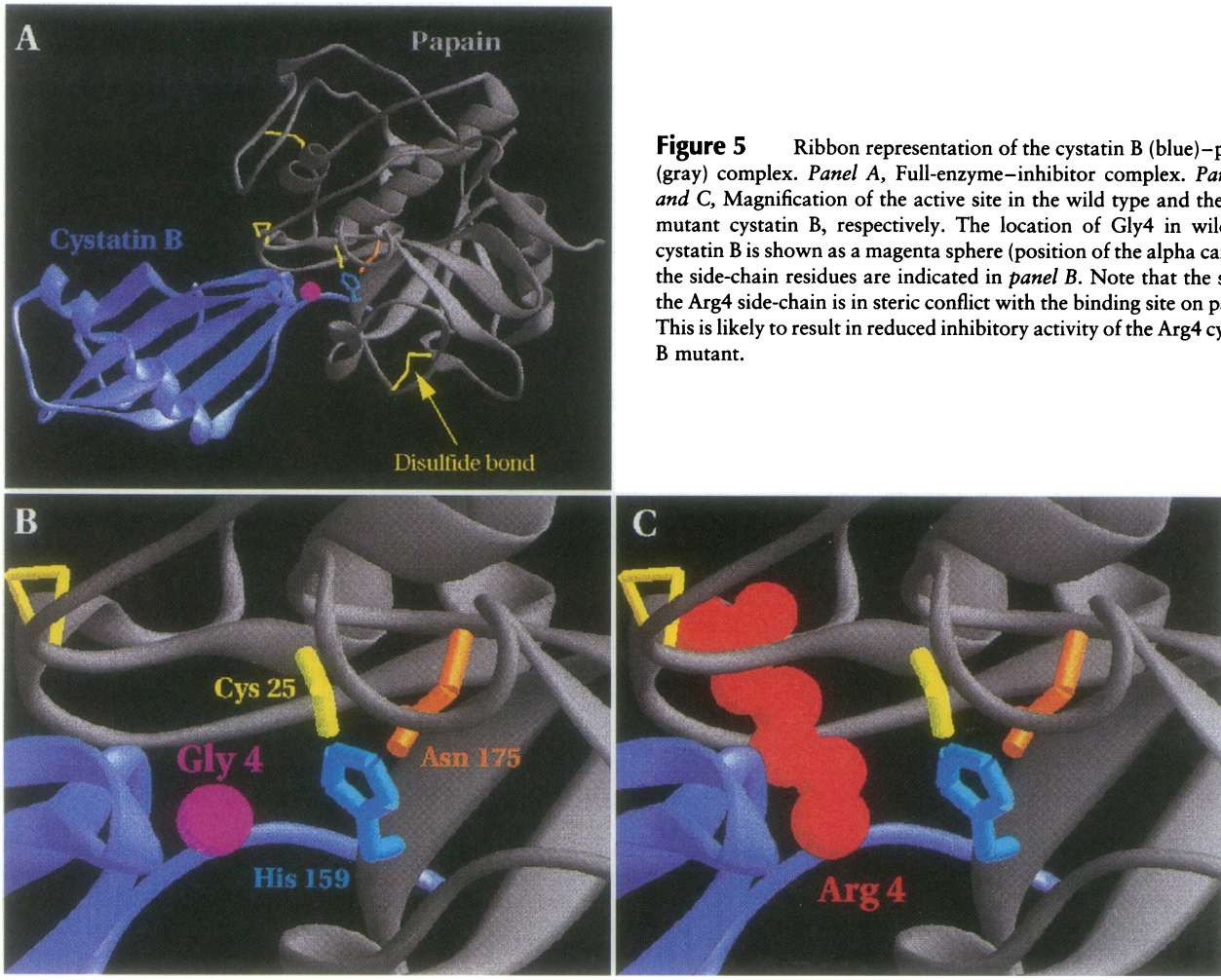

Figure 5 Ribbon representation of the cystatin B (blue)–papain (gray) complex. *Panel A,* Full-enzyme–inhibitor complex. *Panels B and C,* Magnification of the active site in the wild type and the Arg4 mutant cystatin B, respectively. The location of Gly4 in wild-type cystatin B is shown as a magenta sphere (position of the alpha carbon); the side-chain residues are indicated in *panel B.* Note that the size of the Arg4 side-chain is in steric conflict with the binding site on papain. This is likely to result in reduced inhibitory activity of the Arg4 cystatin B mutant.

was detected by SSCA and was present in homozygosity in two patients, of French and Swiss origin, respectively, and in heterozygosity in an additional French patient (table 1). The substitution creates a new restriction site for the enzyme *Alw*26I, which was used to test 100 *cystatin B* alleles from normal individuals from the same canton as the Swiss family. No further alleles with this nucleotide change were found. In addition, in the three families with this change, no unaffected individual is homozygous for 2575A→G. In patient GVA03, who is heterozygous for this variant as well as for the 1924 G→C splice-site mutation, the two nucleotide changes lie on different alleles (data not shown). However, the fact that this substitution occurs in the middle of the 3′ UTR, 95 nt from the stop codon and 160 nt from the polyadenylation site, implies that it represents a nonpathogenic sequence variant rather than a disease-causing mutation. A silent polymorphism in exon 1, 431G→T, was detected in one Spanish family (figs. 1; 2*A,* lane 2; and 3*A*).

Sequencing of the 3′ UTR has detected some variation from the previously published sequence, both in affected and normal individuals (figs. 1 and 3*C*). nts 2581–2590 differ as follows: GGGCTGTGCCCCT, the underlined nucleotides being missing from the available sequence. By screening the dbEST database, we confirmed that both versions are present in different cDNAs. We consequently consider this variation as a common polymorphism. Sequencing of the 5′ UTR revealed that the sequence (CCCCGCCCCGCG)n (starting at nt 207) is polymorphic (figs. 1 and 3*E*); there are two allelic variants, with two or three tandem copies of this dodecamer. The frequency of the three-copy allele was 66% in the normal Caucasian population.

Discussion

We have performed SSCA and sequencing to detect mutations in the *cystatin B* gene of all three exons, the four splice junctions, plus 120 bp 5′ to the initiator ATG and 60 bp 3′ to the termination codon in 29 unrelated patients with EPM1. The patients were of

diverse geographic origins, different from those described by Pennacchio et al. (1996). In retrospect, all the sequence variants found by direct sequencing of PCR products were also detected by SSCA with the two different sets of conditions described in Patients, Material, and Methods.

The 426G→C mutation identified in exon 1, which results in a Gly4Arg substitution, is the first missense substitution described associated with EPM1. The two previously identified mutations (Pennacchio et al. 1996) and all other mutations described in this paper are nonsense, frameshift, or severe splicing defects. The most amino-terminal part of the proteinase inhibitor cystatin B (Met1 to Pro6), as well as the loops connecting the strands B to C and D to E, were shown to interact with the proteinase papain by crystallographic studies of the enzyme-inhibitor complex (PDB entry 1STF; Stubbs et al. 1990). Gly4 is located within 4 Å of the closest papain residue, Asp158, and also within <5 Å of Cys25 in the active site. The Gly4Arg substitution is likely to cause major steric hindrance and strongly affect the papain-binding capacity of cystatin B (fig. 5*B*, *C*). Moreover, Gly4 is conserved in cystatin proteins from many species (Rawlings and Barrett 1990; Turk and Bode 1991). Truncation experiments with chicken and human cystatin B proteins showed that N-terminal deletions, which included Gly4 (or the equivalent Gly9 in chicken), result in a 5,000-fold decrease in binding affinity to papain (Abrahamson et al. 1987; Machleidt et al. 1989; Thiele et al. 1990). We conclude that in all probability the Gly4Arg substitution creates a cystatin B molecule that does not bind papain efficiently and therefore does not have inhibitory activity. However, it is unclear whether the pathogenesis of the disease in this patient is a result of reduced inhibitory activity or/and decreased stability of mutant cystatin B.

The splice-site mutations detected in the conserved dinucleotides of the consensus sequences of the acceptor splice sites of introns 1 and 2, 1924G→C and 2352A→G, respectively, are likely to result in abnormal splicing. The former mutation has been observed in five apparently unrelated patients from both France and the United States. It is possible that there is more than one independent origin of this mutation, since the four French patients do not share a common haplotype for polymorphic sites *PFKL* to *D21S171*. This mutation could result in either skipping of the middle exon or utilization of alternative cryptic splice sites. RNA from a lymphoblastoid cell line from patient GVA05, who is heterozygous for this mutation, was used for RT-PCR to detect abnormally spliced products, but no such products were detected. However, we cannot exclude aberrant splicing on the basis of this RT-PCR reaction, since a previously studied patient had almost undetectable *cystatin B* mRNA in northern blot analysis, implying instability of RNA derived from this allele (Pennacchio et al. 1996). Patient GVA04 also carries this mutation. His father, who transmitted a presently undefined allele, is of Finnish origin. Haplotypes for markers *PFKL, D21S154,* and *D21S171* could not be established; father, mother, and affected had alleles 2 and 3 for *PFKL* and alleles 1 and 2 for *D21S154,* while allele 5 for *D21S171* was linked to the paternal EPM1 allele. It is possible that the paternal EPM1 allele has the common 2-2-5 haplotype for these three markers (Lehesjoki et al 1993*b*) and therefore contains the common but still unknown Finnish mutation.

The 2575A→G variant identified in the 3′ UTR is intriguing. There is no obvious explanation as to how this sequence variant could be responsible for the disease. However, in the three families we have studied, this sequence is always found on the disease alleles. It is remarkable that this change in the *cystatin B* gene is found in one Swiss family from an isolated area and in two families from different regions of France. No A→G change was found in 100 alleles of *cystatin B* from individuals originating from the same canton as the affected Swiss family, suggesting that this nucleotide substitution is not a common DNA polymorphism.

Despite the identification of the new disease-causing mutations described here, the majority of mutations still remain unidentified. Nucleotide sequence analysis of 58 *cystatin B* alleles from apparently unrelated EPM1 patients has revealed deleterious mutations in only 8 alleles (13%). The remainder presumably carry mutations outside of the coding regions and intron-exon junctions, which silence or significantly impair expression of the *cystatin B* gene. This hypothesis is supported by the observation that *cystatin B* mRNA levels are markedly reduced in all patients studied, to date (Pennacchio et al. 1996; present study). Point mutations in *cis*-acting regulatory elements of the *cystatin B* gene could alter its expression. Mutations in the middle of the introns could produce novel splice sites and alter mRNA processing. Deletions, insertions, or rearrangements in the 5′ or 3′ flanking regions of the *cystatin B* gene might eliminate important regulatory regions or bring distal silencers into the vicinity of the gene.

The pathogenesis of the disease in EPM1 patients remains a mystery, since the link between the absence of cystatin B function and the dramatic neuronal deterioration of the patients is unclear. Cystatin B is a ubiquitously expressed, cytosolic cysteine proteinase inhibitor. Cysteine proteinases are small lysosomal proteins that catalyze the destruction of diverse "unwanted" polypeptides (Henskens et al. 1996). The difference in the localization of the proteinase and its inhibitor suggests that cystatin B protects the cell from cysteine proteinases leaking from the lysosomes (McNamara and Puranam

1996). It is unknown whether all brain cells (in particular those affected in EPM1) are normally expressing *cystatin B* and where this proteinase inhibitor is localized in expressing cells. In addition, the exact proteinase(s) inhibited by cystatin B in the affected subpopulations of brain cells have not yet been identified. Furthermore, it is possible that cystatin B may have functions other than proteinase inhibition (McNamara and Puranam 1996).

The mutations detected so far presumably result in either absence of a full-length cystatin B protein or, in the case of the missense Gly4Arg mutation, a polypeptide that does not bind to its target proteinases and is therefore nonfunctional. Other proteinase inhibitors of the cystatin family could partially compensate for the absence of cystatin B (at least for some of its functions) and contribute to the relatively late onset of the phenotype. Imbalance of unknown substances from the continuous absence of cystatin B is compatible with the progressive nature of the disease. The construction of a mouse model for EPM1 with a targeted disruption of the *cystatin B* gene should contribute to the elucidation of the pathophysiology of this human disorder.

Acknowledgments

We thank L. A. Pennacchio, D. Cox, and R. Myers for communicating data prior to publication and M. A. Morris and H. S. Scott for critical reading of the manuscript, valuable suggestions, and assistance in preparing the manuscript. We also thank the genetic collaborative group of the French League Against Epilepsy for providing patients and families, and le Centre de Transfusion, Croix-Rouge, Sion, Switzerland, for blood samples. A. Schönbörner, M.-P. Papasavvas, and P. Hutter are thanked for assistance during the study. The study was supported by grants 31.40500.94 from the Swiss Fonds Nationale de Recherche Scientifique, PL930015 from the Office Federal de l'Education et Recherche/European Union, and from the University and Cantonal Hospital of Geneva, the French INSERM, the French Association against Myopathies, and "le groupement de recherche et d'etude des genomes." M.D.L. is a trainee of the Molecular and Cellular Biology graduate program of the Medical School of the University of Geneva, and M.M. was supported by an Erasmus studentship. We also thank the patients and their families for their collaboration.

References

Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. J Biol Chem 262: 9688–9694

Chumakov I, Rigault P, Guillou S, Ougen P, Billaut A, Guasconi G, Gervy P, et al (1992) Continuum of overlapping clones spanning the entire human chromosome 21q. Nature 359:380–387

Cochius JI, Figlewicz DA, Kälviäinen R, Nousiainen U, Farrell K, Patry G, Söderfeldt, et al (1993) Unverricht-Lundborg disease: absence of nonallelic genetic heterogeneity. Ann Neurol 34:739–741

Genton P, Michelucci R, Tassinari CA, Roger J (1990) The Ramsay-Hunt syndrome revisited: Mediterranean myoclonus versus mitochondrial encephalomyopathy with ragged-red fibers and Baltic myoclonus. Acta Neurol Scand 81:8–15

Henskens YMC, Veerman ECI, Amerongen AVN (1996) Cystatins in health and disease. Biol Chem Hoppe Seyler 377: 71–86

Lafrenière RG, De Jong P, Rouleau GA (1995) A 405-kb cosmid contig and *Hin*dIII restriction map of the progressive myoclonus epilepsy type 1 (EPM1) candidate region in 21q22.3. Genomics 29:288–290

Lalioti MD, Bottani A, Morris MA, Antonarakis SE (1995) The Swiss type of Progressive myoclonus epilepsy (EPM1) maps to 21q22.3. Med Genet 2:249A

Lalioti MD, Chen HM, Rossier C, Shafaatian R, Reid JD, Antonarakis SE (1996) Cloning the cDNA of human PWP2, which encodes a protein with WD repeats and maps to 21q22.3. Genomics 35:321–327

Lehesjoki AE, Eldridge R, Eldridge J, Wilder BJ, de la Chapelle A (1993*a*) Progressive myoclonus epilepsy of Unverricht-Lundborg type: a clinical and molecular genetic study of a family from the United States with four affected sibs. Neurology 43:2384–2386

Lehesjoki AE, Koskiniemi M, Norio R, Tirrito S, Sistonen P, Lander E, de la Chapelle A (1993*b*) Localisation of the EPM1 gene for progressive myoclonus epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. Hum Mol Genet 2:1229–1234

Machleidt W, Thiele U, Laber B, Assfalg-Machleidt I, Esterl A, Wiegand G, Kos J, et al (1989) Mechanisms of inhibition of papain by chicken egg white cystatin. FEBS Lett 243: 234–238

Malafosse A, Lehesjoki AE, Genton P, Labauge P, Durand G, Tassinari CA, Dravet C, et al (1992) Identical genetic locus for Baltic and Mediterranean Myoclonus. Lancet 339: 1080–1081

McNamara J, Puranam RS (1996) Protease inhibitor implicated. Nature 381:26–27

Norio R, Koskiniemi M (1979) Progressive myoclonus epilepsy: genetic and nosological aspects with special reference to 107 Finnish patients. Clin Genet 15:382–398

Pennacchio LA, Lehesjoki A-E, Stone NE, Willour VL, Virtaneva K, Miao J, D'Amato E, et al (1996) Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). Science 271:1731–1734

Rawlings ND, Barrett AJ (1990) Evolution of the cystatin superfamily. J Mol Evol 30:60–71

Stone NE, Fan JB, Willour N, Pennacchio LA, Warrington JA, Hu A, de la Chapelle A, et al (1996) Construction of a 750-kb bacterial clone contig and restriction map in the region of human chromosome 21 containing the progressive myoclonus epilepsy gene. Genome Res 6:218–225

Stubbs MT, Laber B, Bode W, Huber R, Jerala R, Lenarcic B, Turk V (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine

proteinase papain: a novel type of proteinase inhibitor interaction. EMBO J 9:1939–1947

Thiele U, Assfalg-Machleidt I, Machleidt W, Auerswald E-A (1990) N-terminal variants of recombinant stefin B: effect on affinity for papain and cathepsin B. Biol Chem Hoppe Seyler 371:125–136

Turk V, Bode W (1991) The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett 285:213–219

Virtaneva K, Miao J, Träskelin A-L, Stone N, Warrington JA, Weissenbach J, Myers RM, et al (1996) Progressive myoclonus epilepsy EPM1 locus maps to a 175-kb interval in distal 21q. Am J Hum Genet 58:1247–1253

Yamakawa K, Mitchell S, Hubert R, Chen X-N, Colbern S, Huo Y-K, Gadomski, et al (1995) Isolation and characterization of a candidate gene for progressive myoclonus epilepsy on 21q22.3. Hum Mol Genet 4:709–716