ORIGINAL ARTICLE

ETHE1 mutations are specific to ethylmalonic encephalopathy

V Tiranti, E Briem, E Lamantea, R Mineri, E Papaleo, L De Gioia, F Forlani, P Rinaldo, P Dickson, B Abu-Libdeh, L Cindro-Heberle, M Owaidha, R M Jack, E Christensen, A Burlina, M Zeviani

J Med Genet 2006;**43**:340–346. doi: 10.1136/jmg.2005.036210

See end of article for authors' affiliations

Correspondence to: Dr M Zeviani, Unit of Molecular Neurogenetics, National Neurological Institute ''Carlo Besta'', Via Temolo 4, 20126 Milan, Italy; zeviani@istituto-besta.it

Received 23 June 2005
Revised version received 7 September 2005
Accepted for publication 15 September 2005
Published Online First 23 September 2005

Mutations in *ETHE1*, a gene located at chromosome 19q13, have recently been identified in patients affected by ethylmalonic encephalopathy (EE). EE is a devastating infantile metabolic disorder, characterised by widespread lesions in the brain, hyperlactic acidaemia, petechiae, orthostatic acrocyanosis, and high levels of ethylmalonic acid in body fluids. To investigate to what extent *ETHE1* is responsible for EE, we analysed this gene in 29 patients with typical EE and in 11 patients presenting with early onset progressive encephalopathy with ethylmalonic aciduria (non-EE EMA). Frameshift, stop, splice site, and missense mutations of *ETHE1* were detected in all the typical EE patients analysed. Western blot analysis of the ETHE1 protein indicated that some of the missense mutations are associated with the presence of the protein, suggesting that the corresponding wild type amino acid residues have a catalytic function. No *ETHE1* mutations were identified in non-EE EMA patients. Experiments based on two dimensional blue native electrophoresis indicated that ETHE1 protein works as a supramolecular, presumably homodimeric, complex, and a three dimensional model of the protein suggests that it is likely to be a mitochondrial matrix thioesterase acting on a still unknown substrate. Finally, the 625G→A single nucleotide polymorphism in the gene encoding the short chain acyl-coenzyme A dehydrogenase (SCAD) was previously proposed as a co-factor in the aetiology of EE and other EMA syndromes. SNP analysis in our patients ruled out a pathogenic role of SCAD variants in EE, but did show a highly significant prevalence of the 625A alleles in non-EE EMA patients.

Ethylmalonic encephalopathy (EE) (OMIM #602473) is an autosomal recessive disorder originally reported in Italian families and predominantly affecting children of Mediterranean or Arab descent. EE is characterised by psychomotor regression and generalised hypotonia, later evolving into spastic tetraparesis, dystonia, and eventually global neurological failure.[1] MRI examination shows the presence of symmetrical and asymmetrical ''patchy'' lesions, distributed in the deep grey structures of the brain, including the brainstem, thalamus, and corpus striatum.[2 3] The encephalopathy is typically accompanied by widespread lesions of the small blood vessels, causing showers of petechiae, especially during intercurrent infections, easy bruising, and orthostatic acrocyanosis. Chronic diarrhoea is another prominent feature of EE. The course is relentlessly progressive and usually leads to death within the first decade of life. From a biochemical point of view, EE is characterised by persistent lactic acidaemia, a reduction in the activity of mitochondrial respiratory complex IV in skeletal muscle, and markedly elevated excretion of ethylmalonic and methylsuccinic acid in urine.[4] Ethylmalonic acid is believed to derive from the carboxylation of butyryl-coenzyme A (CoA), as a consequence of disorders of the mitochondrial β-oxidation of fatty acids, or from 2-ethylmalonic-semi-aldehyde, as a consequence of the R pathway catabolism of isoleucine.[5] The differential diagnosis of EE includes other known causes of persistent ethylmalonic aciduria (EMA), including deficiency of short chain acyl-CoA dehydrogenase (SCAD), glutaric acidaemia type 2, and Jamaican vomiting sickness. In addition, Gregersen *et al*[6] reported that, in several patients, mild EMA was associated with a predominance of the 625A over the 625G allele due to a single nucleotide polymorphism (SNP), in the gene encoding SCAD.

We present a study on the prevalence of both *ETHE1* (MIM #608451) mutations and SCAD 625G→A SNP alleles in patients with EE and non-EE EMA. To gain insight into the pathogenetic mechanisms of mutations, ETHE1 protein variants were further investigated by immunoassay and three dimensional (3D) modelling.

PATIENTS

Table 1 lists the genetic findings and the clinical information obtained for our patients.

Of the 40 EMA patients enrolled in this study, an initial group of 29 unrelated subjects was classified as having EE. The inclusion criteria for the diagnosis of EE were: (*a*) early onset progressive encephalopathy with symmetrical lesions in the basal ganglia and brainstem; (*b*) vasculopathic petechiae and orthostatic acrocyanosis; (*c*) chronic diarrhoea; and (*d*) EMA. The latter was usually associated with methylsuccinic aciduria, excretion of C4 acylglycines, C4 acylcarnitines, and lactic acidosis. With a few exceptions, most of the 29 EE patients were of Arabic or Mediterranean origin. Seventeen of them have already been reported elsewhere.[3] Unfortunately, the clinical and laboratory information for some of these patients is not complete. For instance, although the available clinical records reported elevated levels of urinary EMA and methyl-succinic acid for each case, the precise values were missing for several patients.

Abbreviations: 3D, three dimensional; 2D-BNE, two dimensional blue native electrophoresis; CRM, cross reacting material; DHPLC, denaturing high performance liquid chromatography; EE, ethylmalonic encephalopathy; EMA, ethylmalonic aciduria; ETF, electron transfer factor; GlyII, glyoxalase II; GSH, glutathione; SCAD, short chain acyl-CoA dehydrogenase; SNP, single nucleotide polymorphism

Table 1 Clinical and molecular features

Patient	Clinical features	EMA/C4 profile	*ETHE1* mutation	Protein	625G/A
EE					
1	EE	EMA = 136/C4 = 2,7	[g.-83delCGCCC]+[c.376-1G→T]	Not translated	A/G
2	EE	High EMA + C4	[g.del ex 4]+[g.del ex 4]	Not translated	A/G
3	EE	High EMA + C4	[g.del ex 4]+[g.del ex 4]	Not translated	A/G
4	EE	EMA = 730/High C4	[c.3G→T]+[c.3G→T]	[p.M1I]+[p.M1I]	A/A
5	EE	EMA = 194	[c.488G→A]/[c.488G→A]	[p,R163Q]+[p.R163Q]	A/G
6	EE	EMA = 150	[c.187C→T]+[c.482G→A]	[p.Q63X]+[p.C161Y]	G/G
7	EE	High EMA/C4 = 1.95	[c.230delA]+[c.230delA]	[p.N77fsX144]+[p.N77fsX144]	G/G
8	EE	EMA = 45	[g.del ex 4]+[g.del ex 4]	Not translated	G/G
9	EE	EMA = 192	[c.406A→G]+[c.488G→A]	[p.T136A]+[p.R163Q]	G/G
10	EE	EMA = 180/High C4	[c.34C→T]+[c.34C→T]	[p.Q12X]+[p.Q12X]	G/G
11	EE	EMA = 446/C4 = 24	[c.375+5G→A]+[c.375+5G→A]	Splice exon-intron 3	G/G
12	EE	EMA = 230/C4 = 18	[c.487C→T]+[c.487C→T]	[p.R163W]+[p.R163W]	G/G
A*	EE	EMA = 320	[c.604_605insG]+[c.604_605insG]	[p.V202fsX220]+[p.V202fsX220]	G/G
B*	EE	High EMA	[c.3G→T]/[c.3G→T]	[p.M1I]+[p.M1I]	A/A
C*	EE	High EMA	[g.del ex 4]+[g.del ex 4]	Not translated	A/G
D*	EE	High EMA	[c.487C→T]+[c.487C→T]	[p.R163W]+[p.R163W]	A/G
E*	EE	EMA = 196	[c.406A→G]+[c.406A→G]	[p.T136A]+[p.T136A]	A/G
F*	EE	EMA = 481	[c.222_223insA]+[440_450del11]	[p.Y74fsX97]+[p.H147fsX176]	A/G
G*	EE	High EMA	[c.222_223insA]+[c.222_223insA]	[p.Y74fsX97]+[p.Y74fsX97]	ND
H*	EE	High EMA	[g.del ex 1–7]+[g.del ex 1–7]	Not translated	G/G
I*	EE	High EMA	[g.del ex 4]+[g.del ex 4]	Not translated	A/A
J*	EE	High EMA	[c.505+1G→T]+[c.505+1G→T]	Splice exon-intron 4	G/G
K*	EE	High EMA	[c.375+5G→A]+[c.375+5G→A]	Splice exon-intron 3	A/G
L*	EE	High EMA	[c.131_132delAG]+[c.488G→A]	[p.E44fsX102]+[p.R163Q]	A/G
M*	EE	High EMA	[c.592_593insC]+[c.592_593insC]	[p.H198fsX220]+[p.H198fsX220]	G/G
N*	EE	High EMA	[c.487C→T]+[c.487C→T]	[p.R163W]+[p.R163W]	G/G
O*	EE	High EMA	[c.113A→G]+[c.554T→G]	[p.Y38C]+[p.L185R]	A/G
P *	EE	High EMA	[c.487C→T]+[c.487C→T]	[p.R163W]+[p.R163W]	ND
Q *	EE	High EMA	[c.505+1G→A]+[c.505+1G→A]	Splice exon-intron 4	A/G
Non-EE					
CE	PD, S, H	High EMA	–		A/A
AK	PD, H, CV	EMA = 70/normal C4	–		A/G
LY	M, LF	High EMA	–		A/A
CR	PD, H, CV	EMA = 102/normal C4	–		A/A
TL	PD, CV	EMA = 239/normal C4	–		A/A
GO	PD	High EMA	–		A/A
SF	PD, H, BA, S	EMA = 80	–		A/A
II	PD	High EMA	–		A/G
IR	PD, DM, M, PE	EMA = 55	–		A/A
DK	PD, PE	High EMA	–		A/A
RE	PE, S, DM	EMA = 141	–		A/A

EMA, ethylmalonic acid expressed as mg/g creatinine (normal values <14.6); C4, C4 acylcarnitine expressed as mmol/l (normal values <1.4); PD, psychomotor delay; S, seizures; H, hypotonia; CV, cyclic vomiting; M, myopathy; LF, liver failure; BA, brain atrophy; DM, dysmorphisms; PE, progressive encephalopathy. DNA mutation numbering based on cDNA sequence starts from the first ATG. *Reported in Tiranti *et al*.[3]

Eleven other patients were referred to us as presenting an infantile encephalopathy or encephalomyopathy consistently associated with EMA. For these patients, urinary excretion of methyl-succinic acid was normal. For three of them, the acylcarnitine profile was available in the clinical records and was also normal (table 1). Again, most of these patients died several years ago and the clinical information that we were able to obtain is incomplete (table 1). The neurological features were more heterogeneous in this group of patients than in the EE group. In addition, we did not find the typical extra-neurological EE hallmark features in any of these patients; for instance, vascular purpura, orthostatic acrocyanosis, and chronic diarrhoea were notably absent. The age of onset varied from birth to childhood. Additional signs were reported in some patients, including facial dysmorphism, agenesis of corpus callosum, mid brain calcifications, and brain atrophy (table 1). They were referred to us as non-typical cases, and were classified by us as having non-EE EMA.

When available, the quantitative values of EMA in both cohorts were extremely variable, ranging from 45 to 730 mg/g creatinine.

Mitochondrial β-oxidation of fatty acids was analysed fluorimetrically in fibroblasts[7] obtained from five patients presenting with typical EE (subjects 1, 2, 5, 8, and 10 in table 1) and in three patients presenting with non-EE EMA (subjects LY, CE, RE in table 1); in all cases no abnormalities were found.

MATERIALS AND METHODS

DNA and western blot analysis

DNA extracted from lymphocyte or fibroblast specimens was used as a template to amplify the seven exons of the *ETHE1* gene using intronic primers, as previously described.[3] Sequence analysis was performed on a 3100 ABI automated sequencer on samples prepared using a commercial kit (BigDye Termination kit; Applied Biosystems). Data were elaborated and analysed using the SeqScape software (Applied Biosystems). DNA mutation numbering was based on the cDNA sequence starting from the first ATG.

Western blot analysis was carried out on patient fibroblasts using a polyclonal antibody (anti-Ethe1p^{C17}) against an oligopeptide encompassing amino acids 190 to 206 in the C-terminus of the ETHE1 protein.[3] Approximately 2×10^6 cells were trypsinised, pelleted, sonicated, and solubilised, as described elsewhere.[8] Electrophoresis on 12% SDS polyacrylamide gel of 100–200 μg protein/lane and western blot analysis were performed using the ECL chemiluminescence kit (Amersham), as described previously.[8]

DHPLC analysis of the 625G→A polymorphisms

The following intronic primers were used to amplify exon 6 for the detection of 625A→G in the SCAD gene: −625 sense, 5′-ttggtgttgggtgtgctggtgt-3′; −625 antisense, 5′-gggctcaccatggctatcttga-3′. PCR amplification conditions were: an

initial denaturation step at 94℃ for 3 minutes, then 32 cycles of 94℃ for 30 seconds, 59℃ for 20 seconds, and 72℃ for 30 seconds, with a final extension at 72℃ for 2 minutes.

Analysis of the SNP was carried out on PCR fragments using a denaturing high performance liquid chromatography (DHPLC) apparatus (Transgenomic). After PCR amplification, the fragment was denatured at 95℃ for 10 minutes, re-annealed at 65℃ for 10 minutes, and slowly cooled to 4℃ to allow the formation of the heteroduplex. To detect homozygous changes, we also performed DHPLC analysis using a mixture composed of a 1:1 ratio of wild type fragment and fragments that showed a single DHPLC elution peak when run alone. Conditions for DHPLC analysis were optimised using the Navigator software provided by Transgenomic.

Western blot analysis of mitochondrial protein fraction resolved by 2D-BNE

For two dimensional blue native electrophoresis (2D-BNE), isolated mitochondria from mouse kidney were obtained as described.[9] 2D-BNE and electroblotting procedures were performed as described previously,[8] except that the first (non-denaturing) BNE run was performed through a 10–20% gradient polyacrylamide gel. The blots were immunostained using rabbit antibodies specific to the ETHE1 protein,[3] or to the α or β subunit of the electron transfer factor (ETF) as previously described.[10] In a second set of experiments, the samples were either treated with 5 mmol/l β-mercaptoethanol (final concentration), or boiled for 10 minutes, prior to be subjected to the first (non-denaturing) BNE run.

Generation of the N-terminus variant of *ETHE1* cDNA and cloning in a prokaryotic inducible expression system

A short variant of human *ETHE1* cDNA lacking the first 10 amino acid residues on the N-terminus was cloned into the prokaryotic expression vector pMW172. The recombinant cDNA contained, on the 3′ end, an in frame DNA region encoding a six histidine epitope.

Purification of recombinant ETHE1 protein

Escherichia coli strain BL21(DE3) was transformed with the plasmid pMW172^{ETHE}11. Fresh recombinant colonies were grown up in 1 litre of Luria broth containing 100mg/l ampicillin. The culture was induced for *ETHE1* gene expression with 0.2 mmol/l isopropyl-D-thiogalactoside (IPTG) to an optical density of 0.7 at 600 nm and cultured for 4 additional hours. Cells were harvested by centrifugation and resuspended in 50 mmol/l NaH_2PO_4, 300 mmol/l NaCl pH 8, plus 1 mmol/l dithiothreitol and a cocktail of protease inhibitors. Cells were then lysed by 5–6 passages of freeze/thawing and sonication, and cell lysates were centrifuged at 4℃ for 20 minutes at 10 000 ***g*** to remove cell debris. After adding 10 mmol/l imidazole (final concentration) the lysate was applied to an affinity column (HIS-Select Nickel Affinity Gel column; Sigma Chemical Co.) pre-equilibrated with 50 mmol/l sodium phosphate pH 8, 0.3 mol/l NaCl, and 10 mmol/l imidazole. Elution of the protein was performed by increasing the concentration of imidazole to 250 mmol/l in the same buffer.

Quantitative gel filtration

The native molecular mass of recombinant ETHE1 protein was determined by gel filtration on a Superdex 75 10/300 GL column (Amersham) using HPLC (Waters Corporation) with 50 mmol/l Tris-HCl pH 8, 300 mmol/l NaCl, a flux of 0.5 ml/min, and a sparge of 20 ml/min. The molecular size standards for gel filtration were: aprotin (6.5 kDa), cytochrome c (12.4 kDa), trypsin inhibitor (20.1 kDa), β-lacto-globulin (36.8 kDa), egg albumin (45 kDa), transferrin (76 kDa), and β-amylase (200 kDa). Results were analysed using Empower software (Waters Corp.).

Three dimensional model of the ETHE1 protein

Homologues of the ETHE1 protein and human glyoxalase II (hGlyIIp) were searched in the non-redundant database of protein sequences at NCBI, using both BLAST and PSI-BLAST.[11] Multiple sequence alignments and phylograms were generated with ClustalW[12] and T-Coffee.[13] Secondary structure was predicted with JPRED[14] and PSI-PRED.[15]

The alignment adopted for the 3-D model generation was obtained using data from multiple sequence alignments, secondary structure prediction, and the threading metaserver 3D-Jury.[16] The three dimensional model of the ETHE1 protein was built by comparative modelling with the hGlyII (PDB code 1QH5) scaffold using the Jackal protein modelling software package (http://trantor.bioc.columbia.edu/~xiang/jackal). The initial model was submitted to molecular mechanics optimisation using the CVFF forcefield.[17] In particular, only the geometry of the protein side chains was initially optimised (1000 steps using the steepest descent algorithm followed by 10 000 steps using the conjugate gradient algorithm). The optimisation was then restarted restraining only the alpha carbon of the peptide chain. In the last step, the structure of the second domain was optimised after removal of all constraints. The quality of the final models was evaluated using the What If suite of software programs.[18]

RESULTS

ETHE1 sequence and western blot analysis

Sequence analysis of the *ETHE1* gene revealed the presence of 22 mutations, which were all found in the 29 subjects affected by EE (table 1). No mutation was found in the 11 non-EE EMA patients. Sixteen mutations predicted nonsense changes, while six mutations predicted missense changes. A few silent polymorphisms were also detected (not shown).

Western blot analysis performed on nine fibroblast cell lines from EE patients showed the absence of ETHE1 protein specific cross reacting material (CRM) in all of the alleles carrying loss of function mutations. However, no CRM was detected in the alleles carrying the missense mutations T136A (fig 1), Y38C, and L185R.[3] In contrast, the homozygous mutation R163W was associated with the persistence of substantial amounts of CRM in the three patients who carried this change. The same result was obtained in patient L,[3] carrying the two allelic mutations E44fs102X and R163Q, and patient 9 (table 1), carrying the two allelic mutations T136A and R163Q. In these two patients, the R163Q mutation was allelic to either a loss of function mutation (E44fs102X), or T136A, both associated with the absence of the corresponding protein. Therefore, the only contribution to

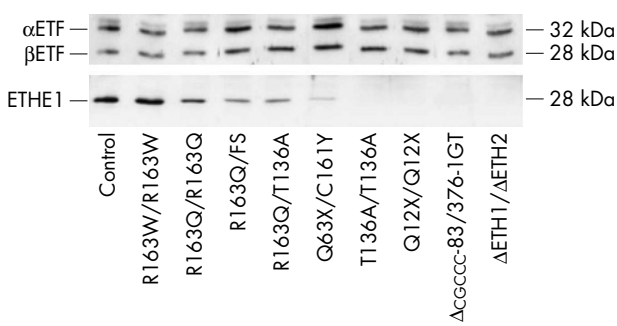

Figure 1 Western blot analysis on patient fibroblasts. The same filter was hybridised with anti-ETHE1 protein specific antibody and, after stripping, with a mixture of anti-α and anti-β ETF subunits.

the ETHE1 protein specific CRM in patients L and 9 was from the R163Q allele. Finally, Western blot analysis on the fibroblast lysate from patient 6 (table 1), carrying the missense mutation C161Y, and the nonsense Q63X mutation, which is presumably associated with the absence of the protein, showed a reduced but clearly detectable amount of ETHE1 protein specific CRM.

In addition, no ETHE1 protein specific CRM was detected in fibroblast lysate from patient 1 (table 1). This patient was a compound heterozygote carrying a splice site mutation in one allele. The second allele carried a 5 bp deletion affecting a CGCCC sequence that is repeated twice and is located about 78–83 bp upstream from the first ATG of the *ETHE1* open reading frame and 12–17 bp upstream from the putative 5′ untranslated regipn of the gene. Using Proscan (http://bimas.dcrt.nih.gov/molbio/proscan/), a software program for the prediction of promoter sequences, we found that the CGCCC deleted sequence is part of a 10 bp putative recognition site for the Sp1 transcription factor (CCCCCGCCCC).

None of these mutations was found in 250 DNA samples from normal Italians and 25 normal individuals of Arabic origin.

SCAD SNP analysis

In order to elucidate the role of susceptibility SCAD variants in the excretion of high level of ethylmalonic acid, we investigated the prevalence of the 625G/A polymorphism in the 29 EE patients (group I), in the 11 non-EE *ETHE1* negative EMA patients (group II), and in 53 control individuals (group III). The distribution of the 625G and 625A alleles in groups I and II is reported in table 1. No significant difference was detected in the allele frequencies between groups I and III, while group II had a significant higher frequency of the 625A allele relative to the 625G allele, compared with each of the other groups(χ^2 test): group II versus group I, $p = 4.25 \times 10^{-5}$; group II versus group III $p = 2.52 \times 10^{-8}$.

Western blot analysis on 2D-BNE

The predicted molecular weight of mature ETHE1 protein is 28 kDa. To understand whether ETHE1 protein works as a monomer or is part of a relatively stable supramolecular complex, we performed 2D-BNE on proteins extracted from mouse kidney mitochondria. After electroblotting, an ETHE1 protein specific rabbit polyclonal antibody, which cross reacts with the human and murine protein species, was used to immunostain the filter replicas of the 2D-BNE gels. As shown (fig 2A–C), the ETHE1 protein specific signal migrated slightly faster than those specific to the α and β subunits of the ETF, an $\alpha_1\beta_1$ 60 kDa heterodimeric mitochondrial enzyme. As shown (fig 2D–F), a spot corresponding to a protein species of smaller size (approximately 28 kDa), presumably the ETHE1 protein monomer, was visualised by boiling the protein extract prior to 2D-BNE, while pre-treatment with the reducing compound β-mercaptoethanol was ineffective (data not shown). Taken together, these results suggest that, in native conditions, the ETHE1 protein is part of a complex, the size of which is between 28 and 60 kDa.

Quantitative gel filtration

The molecular weight of an affinity purified recombinant ETHE1 protein in native conditions was further analysed by size exclusion chromatography. A calibration curve was constructed by calculating the elution volume parameter (K_{av}) for each protein used for the calibration of the column. The molecular weight of recombinant ETHE1 protein,

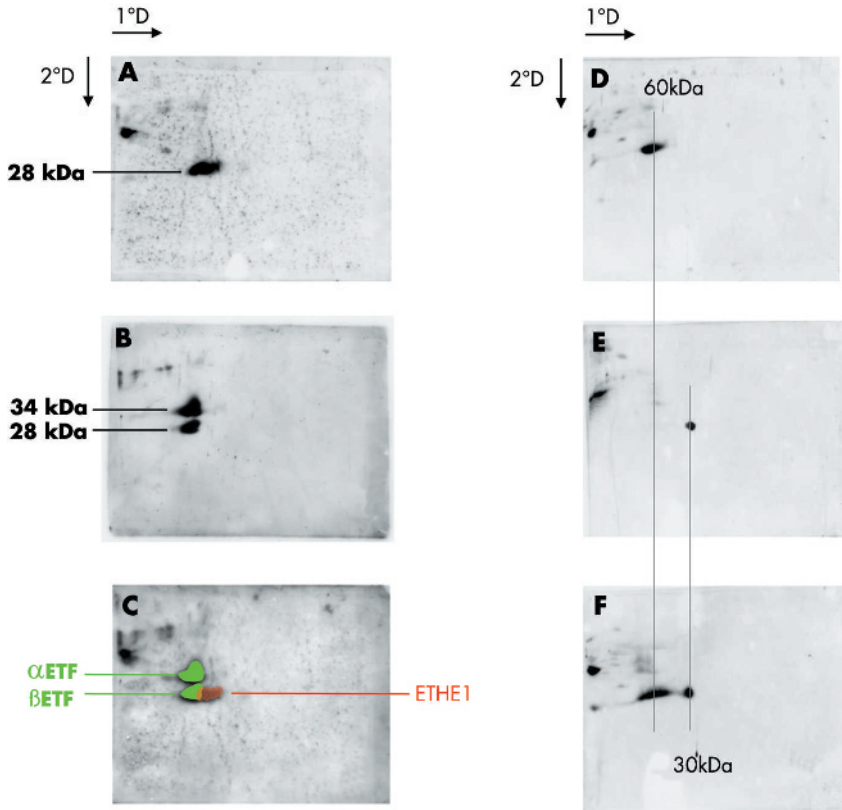

Figure 2 2D-BNE analysis. Western blot analysis of (A) ETHE1 protein; (B) α and β ETF subunits on the same filter. (C) Overlap between (A) and B). Western blot analysis of (D) native ETHE1 protein; (E) sample boiled prior to 2D-BNE. (F) Mixed experiment using a combination of native and pre-boiled samples.

deduced by calculating the elution volume parameter (K_{av}) of the protein using the calibration curve, resulted to be approximately 50–55 kDa. These results confirm those obtained by 2D-BNE analysis, and suggest that in native conditions most of the ETHE1 protein is present as a homodimeric structure.

Three dimensional model of ETHE1 protein

The ETHE1 protein primary sequence shares similarity with both cytosolic and mitochondrial glyoxalase II (GlyII).[3] Both the GlyII and ETHE1 amino acid sequences contain the THXHXD signature, which is typical of the metallo-β-lactamase superfamily.[19] However, the alignment (not shown) and the phylogram (fig 3) obtained comparing homologues of the ETHE1 protein (ETHE1-like sequences) and of human GlyII (GlyII-like sequences) show that these two classes are evolutionarily distinct, indicating that ETHE1-like proteins are representative of a new family of metallo-hydrolase enzymes characterised by the β-lactamase fold.

According to a 3D model of the ETHE1 protein obtained by a comparative modelling technique using the hGlyII as a 3D template, the first domain is characterised by the typical β-lactamase fold and metal binding site, which are both present in hGlyII (fig 4). It is known that five histidine and two aspartate residues interact directly with two zinc ions in hGlyII.[20] All these residues are conserved in ETHE1, strongly suggesting that also this protein can bind two metal ions. In particular, the metal ion ligands are predicted to be H79, H81, and H135 (first metal ion; M1), and D83, H84, and H195 (second metal ion; M2). D154 is predicted to be the residue bridging the two ions.

In addition, the model indicates that the C-terminal region of the ETHE1 protein presents only three α-helices instead of the five predicted for both cytoplasmic and mitochondrial GlyII (fig 4). In particular, helices corresponding to α4 and α5 of hGlyII are missing in ETHE1. Notably, the first helix of the second domain in ETHE1 corresponds to the α6 helix of hGlyII, which is an invariant feature of enzymes with a β-lactamase fold.[21]

DISCUSSION

Genotype to phenotype correlations

In this study, we provide evidence that mutations in the *ETHE1* gene are specifically associated with typical EE syndrome. Contrariwise, no mutation in this gene was found in patients with EMA who did not present EE. Non-EE patients were all characterised by the presence of heterogeneous neurological symptoms and neuropathology, and by the consistent absence of both vascular manifestations and chronic diarrhoea. Taken together, these results indicate that non-EE EMA is due to the involvement of genes other than *ETHE1*, including SCAD. Mutations in this gene leading to impairment of the terminal part of the mitochondrial fatty acid β-oxidation pathway are typically associated with EMA.[6] Our results clearly demonstrate that while EE is a monogenic disorder exclusively caused by mutations in *ETHE1*, the SCAD 625A SNP acts as a susceptibility factor in ''idiopathic'', non-EE EMA patients. A role of SCAD in EMA, at least as a

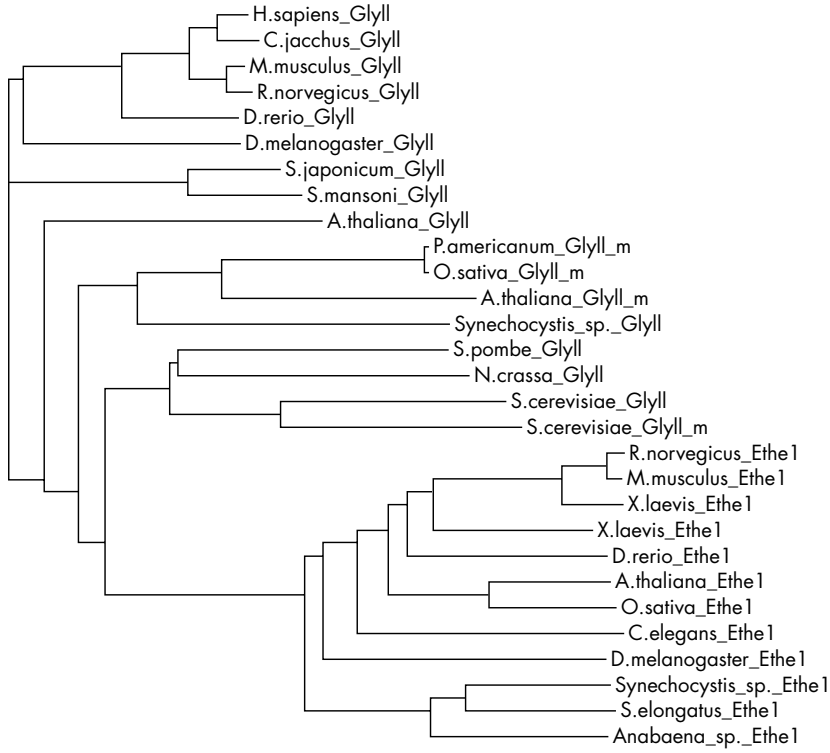

Figure 3. Phylogram tree of GlyII-like and ETHE1-like protein sequences. Accession codes: *H sapiens*_GlyII (SwissProt code Q16775); *Mus musculus*_GlyII (Q99KB8); *Rattus norvegicus*_GlyII (O35952); *Callithrix jacchus*_GlyII (Q28333); *. cerevisiae*_GlyII (Q05584); *Saccharomyces cerevisiae*_GlyII_m (Q12320); *Arabidopsis thaliana*_GlyII (O24496); *A thaliana*_GlyII_m (O24495); *Oryza sativa*_GlyII_m (Q940L0); *Pennisetum americanum*_GlyII_m (Q8LRN0); *Drosophila melanogaster*_GlyII (Q9I7P8); *Schistosoma japonicum*_GlyII (Q86DW6); *Schistosoma mansoni*_GlyII (Q26547); *Schizosaccharomyces pombe*_GlyII (Q9UT36); *Synechocystis*_sp _GlyII (Q26547); *Neurospora crassa*_GlyII (Q9C2G8); *Danio rerio*_GlyII (GenPep code AAH66607. 1). The signature ''_m'' indicates the sequences of mitocondrial GlyII-like proteins. Sequences of ETHE1-like proteins are: *H sapiens*_ETHE1 (Swiss Prot code O95571); *M. musculus*_ETHE1 (Q9DCM0); *A thaliana*_ETHE1 (Q93VR2); *Xenopus laevis*_ETHE1 (Q8AVQ9); *D melanogaster*_ETHE1 (Q86PD3); *Synechocystis_sp.* _ETHE1 (P73799); *Caenorhabditis elegans*_ETHE1 (O17636); *Oryza sativa*_ETHE1 (Q8RZ88); *Anabaena*_sp. _ETHE1 (Q8YQC9); *Synechococcus elongatus*_ETHE1 (Q8DIG2); *Bradyrhizobium japonicum*__ETHE1 (Q89VN8); *R. norvegicus*_ETHE1 (GenPep code XP_214838. 1); *D rerio*_ETHE1 (AAH67574. 1).

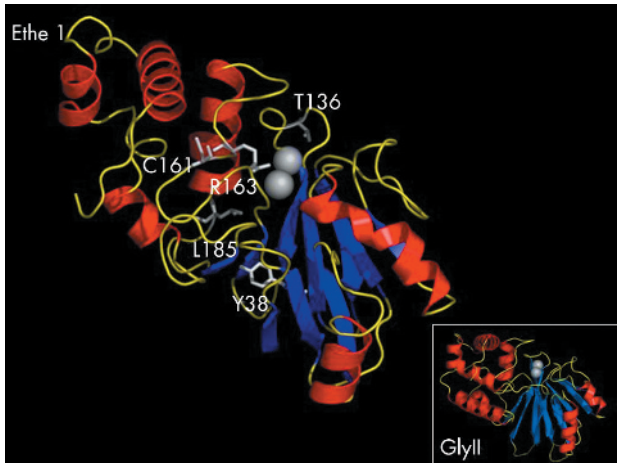

Figure 4 Three dimensional models. The ETHE1 protein predicted using hGlyIIp as a template. The β-strands, α-helices, and loop regions are shown as blue, red, and yellow ribbons, respectively. The two metal ions (shown as grey spheres) are located in the model in the same position that they occupy in the hGlyII 3D structure. The side chain of Y38, T136, C161, R163, and L185 are shown as grey sticks and dots. (Inset) The human GlyII model.[18]

susceptibility gene, was suggested by biochemical studies.[22] Results were controversial, however, possibly because the cohort of EMA patients analysed for the presence of the 625→A SNP contained both *ETHE1* positive EE patients and *ETHE1* negative patients affected by ''idiopathic'' EMA. Biochemical studies on the terminal part of the mitochondrial fatty acids β-oxidation pathway, and sequence analysis of the *SCAD* gene will help clarify the role of SCAD in the pathogenesis of the EMA associated disease in these patients.

The majority of the *ETHE1* mutations so far identified predict the loss of function of the ETHE1 protein. However, we also found missense mutations affecting highly conserved aminoacid residues, which were associated with the presence of either normal or reduced amounts of the ETHE1 protein. Nevertheless, the clinical phenotype was quite homogeneous among our patients, with a few exceptions, and no obvious correlation could be established between the clinical features, including age at onset, severity of symptoms, time of progression, and the position or type of mutation. For instance patient 8 (table 1), who, albeit affected by typical features of EE, is now 20 years old, carries a loss of function homozygous mutation associated with the complete absence of ETHE1 protein specific CRM. Conversely, the clinical features of patients carrying the ''catalytic'' ETHE1 protein mutations R163W and R163Q, which are associated with the presence of normal amounts of CRM, were indistinguishable from the cases in which the protein was absent. Even though EE has been described mainly in families from the Mediterranean basin and the Arabian peninsula, we obtained no evidence for the existence of an ancestral haplotype or a cluster of common mutations in our cohort of patients with EE. A possible explanation could be the presence of a selective advantage of subject carrying heterozygous mutation in the *ETHE1* gene.

Molecular considerations

The results of 2D-BNE based immunoblotting and gel filtration based analysis strongly suggest that the ETHE1 28 kDa mature protein takes part in a supramolecular complex, most likely a homodimer, of ≈55 kDa in size. The individual protein subunits of this complex must interact with each other via non-covalent bonds, as the complex is disrupted by heat. However, an interaction via intermolecular disulphide bridges is unlikely, given the inability of a strongly reducing thiol compound, such as β-mercaptoethanol, to resolve the 28 kDa ETHE1 protein monomer from the native complex.

The comparison between a 3D model of ETHE1 protein, obtained by comparative modelling with the X-ray structure of human GlyII (hGlyII), highlights both relevant similarities and differences which help understand structure function relationships. Three conserved basic residues of hGlyII (R249, K252, and K143) are present in all GlyII proteins identified so far. R249 and K252 play a crucial role in substrate recognition, by interacting with the glycine carboxylate of the glutathione (GSH) moiety. Accordingly, these residues are not conserved in the GlyII-like proteins from *Trypanosoma brucei* and other kinetoplastid organisms, for which the physiological substrates are not GSH thioesters but trypanothione thioesters.[23] In the ETHE1 protein, the basic residues R249 and K252 of hGlyII are substituted by the hydrophobic residues P242 and M245, suggesting a specific substrate different from D-lactoyl-GSH, the main substrate of hGlyII. This is further corroborated by the phylogram obtained comparing ETHE1-like and GlyII-like members, which indicates that the two families are evolutionarily distinct (fig 3).

Other residues relevant for the interaction with the substrate of hGlyII are Y175, N179, Y145, F182 and D253.[20] The Y197 residue, corresponding to Y175 in hGlyII, is the only residue involved in substrate interactions that is conserved in ETHE1. Remarkably, this aminoacid plays a key role in the interaction with the thioester group of the substrate in hGlyII,[24] and is replaced by different residues in non-thioesterase enzymes of the β-lactamase superfamily.[25] The N179 and F182 residues, which are located in helices α4 and α5 in GlyII, are not present in ETHE1, where the helices α4 and α5 are missing. Taken together, these data suggest that the ETHE1 protein specific substrate is a (still unknown) compound, which, albeit clearly different from D-lactoyl-GSH, contains, like the latter, a thioester bond.

These observations and previous studies[3] both rule out the hypothesis that ETHE1-like proteins are mitochondrial GlyII enzymes. This is in keeping with the recent identification of the human mitochondrial GlyII as an alternative splice variant of cytoplasmic GlyII, which determines the synthesis of an N-terminal mitochondrial target peptide II.[26]

Finally, a mutation disrupting an Sp1 site in the promoter region of the gene was associated with the complete absence of the protein, suggesting that the expression of *ETHE1* is under strong transcriptional control.

Catalytic versus structural mutations

Our ETHE1 model makes it possible to analyse the spatial location of the aminoacid changes predicted by missense mutations (table 1). Two types of mutation were identified in our patients. The T136A, Y38C, and L185R changes are ''structural'' mutations associated with the absence of ETHE1 protein specific CRM. By contrast, the R163W, R163Q, and C161Y changes are ''catalytic'' mutations associated with the maintenance of normal or slightly reduced amounts of ETHE1 protein specific CRM.

The T136 residue, which is strictly conserved in both ETHE1 and GlyII-like proteins, belongs to the second coordination sphere of the metal ions, in close proximity to the active site (fig 4). However, in contrast to the ''catalytic'' mutations discussed below, the absence of ETHE1 protein specific CRM in T136A mutant fibroblasts indicates that the T136 residue is crucial in maintaining the structural integrity of the protein. The same is likely to occur for the Y38C and L185R mutations. In particular, the Y38 residue is located at the interface between the first and the second half of the first domain, where it is involved in the formation of a cluster of

DATABASE INFORMATION

- ETHE1 (OMIM#608451)
- EE (OMIM#602473)
- *Homo sapiens Ethe1* (Swiss Prot code O95571)
- *H sapiens* ethylmalonic encephalopathy 1 *ETHE1*, mRNA (NM_014297)
- *H sapiens* acyl-coenzyme A dehydrogenase, short/branched chain mRNA (NM_000017)

hydrophobic residues in the protein core. The L185 residue belongs to the first domain, in close proximity to two positively charged amino acids, K181 and R209. As a consequence, the effect of the L185R mutation, which changes an apolar leucine to a positively charged arginine, could be due to unfavourable electrostatic interactions between the three positively charged residues, leading to destabilisation and degradation of the protein.

The R163 residue in ETHE1 corresponds to the K143 in the hGlyII protein and as shown (fig 4), is located in a loop region, which is part of the putative catalytic pouch, near the binding site of zinc ions. This observation suggests that R163Q and R163W mutations impair the catalytic activity, rather than the structural integrity, of Ethe1p. The C161 residue, which does not interact with the metal binding site (fig 4), belongs to a GCG motif, which is conserved in both the ETHE1-like and the GlyII-like protein families. In hGlyII, the GCG motif forms an alpha hairpin that interacts with helices α4 and α6, by hydrophobic contacts and hydrogen bonds.[20] Although the helix corresponding to hGlyII α4 is missing in ETHE1, the GCG motif present in ETHE1 can still play a key structural role in the stabilisation of the protein. This consideration can explain the low amount of C161Y ETHE1 protein specific CRM detected by Western blot analysis (fig 1).

Taken together. these findings confirm that the active site of ETHE1 is similar to that of hGlyII, and corroborate the hypothesis that ETHE1, like hGlyII, is a thioesterase enzyme, which is part of a novel metabolic pathway specific to mitochondria.

ACKNOWLEDGEMENTS

We are indebted to Ms B Geehan for revising the manuscript.

Authors' affiliations

V Tiranti, E Briem, E Lamantea, R Mineri, M Zeviani, Unit of Molecular Neurogenetics, Pierfranco and Luisa Mariani Center for the Study of Children's Mitochondrial Disorders, National Neurological Institute ''C. Besta'', Milan, Italy
E Papaleo, L De Gioia, Department of Biotechnology and Biosciences. University of Milan-Bicocca, Italy
F Forlani, Dipartimento di Scienze Molecolari Agro-Alimentari, Facoltà di Agraria, State University of Milan, Italy
P Rinaldo, Laboratory Medicine and Pathology, Mayo Clinic and Foundation, Biochemical Genetics Laboratory, Rochester, MN, USA
P Dickson, Harbor-UCLA Medical Center, Division of Medical Genetics, Torrance, CA, USA
B Abu-Libdeh, Makassed Hospital, Mount of Olives, Jerusalem, Israel
L Cindro-Heberle, Paediatric Neurology Unit, Al Sabah Hospital, Kuwait
M Owaidha, Al-Jahra Hospital, Kuwait
R M Jack, Biochemical Genetics lab, Children's Hospital, Sand Point Way NE, Seattle, WA, USA
E Christensen, Department of Clinical Genetics, Rigshospitalet, Copenhagen, Denmark
A Burlina, Division of Inborn Metabolic Diseases, Department of Pediatrics, University of Padua, Italy

Supported by Fondazione Telethon-Italy (grant no. GGP030039), Fondazione Pierfranco e Luisa Mariani, MITOCIRCLE and EUMITOCOMBAT network grants from the European Union Framework Program 6.

Competing interests: there are no competing interests.

The first two authors contributed equally to this work.

REFERENCES

1 **Burlina A**, Zacchello F, Dionisi-Vici C, Bertini E, Sabetta G, Bennet MJ, Hale DE, Schmidt-Sommerfeld E, Rinaldo P. New clinical phenotype of branched-chain acyl-CoA oxidation defect. *Lancet* 1991;**14**:1522–3.
2 **Grosso S**, Balestri P, Mostardini R, Federico A, De Stefano N. Brain mitochondrial impairment in ethylmalonic encephalopathy. *J Neurol* 2004;**251**:755–6.
3 **Tiranti V**, D'Adamo P, Briem E, Ferrari G, Mineri R, Lamantea E, Mandel H, Balestri P, Garcia-Silva MT, Vollmer B, Rinaldo P, Hahn SH, Leonard J, Rahman S, Dionisi-Vici C, Garavaglia B, Gasparini P, Zeviani M. Ethylmalonic encephalopathy is caused by mutations in ETHE1, a gene encoding a mitochondrial matrix protein. *Am J Hum Genet* 2004;**74**:239–52.
4 **Garcia-Silva MT**, Ribes A, Campos Y, Garavaglia B, Arenas J. Syndrome of encephalopathy, petechiae, and ethylmalonic aciduria. *Pediatr Neurol* 1997;**17**:165–70.
5 **Nowaczyk MJ**, Lehotay DC, Platt BA, Fisher L, Tan R, Phillips H, Clarke JT. Ethylmalonic and methylsuccinic aciduria in ethylmalonic encephalopathy arise from abnormal isoleucine metabolism. *Metabolism* 1998;**47**:836–9.
6 **Gregersen N**, Winter VS, Corydon MJ, Corydon TJ, Rinaldo P, Ribes A, Martinez G, Bennett MJ, Vianey-Saban C, Bhala A, Hale DE, Lehnert W, Kmoch S, Roig M, Riudor E, Eiberg H, Andresen BS, Bross P, Bolund LA, Kolvraa S. Identification of four new mutations in the short-chain acyl-CoA dehydrogenase (SCAD) gene in two patients: one of the variant alleles, 511C→T, is present at an unexpectedly high frequency in the general population, as was the case for 625G→A, together conferring susceptibility to ethylmalonic aciduria. *Hum Mol Genet* 1998;**7**:619–27.
7 **Freeman FE**, Goodman SI. Fluorimetric assay of acyl-CoA dehydrogenases in normal and mutant human fibroblasts. *Biochem Med* 1985;**33**:38–44.
8 **Tiranti V**, Galimberti C, Nijtmans L, Bovolenta S, Perini MP, Zeviani M. Characterization of SURF-1 expression and Surf-1p function in normal and disease conditions. *Hum Mol Genet* 1999;**8**:2533–40.
9 **Nijtmans LG**, Henderson NS, Holt IJ. Blue Native electrophoresis to study mitochondrial and other protein complexes. *Methods* 2002;**26**:327–34.
10 **Parini R**, Vegni C, Martini J, Romeo A, Garavaglia B. Sudden infant death and multiple acyl-coA dehydrogenation disorders. *Eur J Ped* 1995;**154**:421–22.
11 **Altschul SF**, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;**25**:3389–402.
12 **Higgins D**, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibsor TJ. CLUSTAL-W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;**22**:4673–80.
13 **Notredame C**, Higgins DG, Heringa J. T-Coffee: a novel method for multiple sequence alignments. *J M B* 2000;**302**:205–17.
14 **Cuff JA**, Clamp ME, Siddiqui AS, Finlay M, Barton GJ. Jpred: a consensus secondary structure prediction server. *Bioinformatics* 1998;**14**:892–3.
15 **McGuffin LJ**, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics* 2000;**16**:404–5.
16 **Ginalski K**, Elofsson A, Fischer D, Rychlewski L. 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics* 2003;**19**:1015–18.
17 **Dauber-Osguthorpe P**, Roberts VA, Osguthorpe DJ, Wolff J, Genest M, Hagler AT. Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins* 1998;**4**:31–47.
18 **Vriend G**. WHAT IF: a molecular modelling and drug design program. *J Mol Graph* 1990;**8**:52–6.
19 **Aravind L**. An evolutionary classification of the metallo beta-lactamase fold proteins. *In Silico Biol* 1999;**1**:69–91.
20 **Cameron AD**, Ridderstrom M, Olin B, Mannervik B. Crystal structure of human glyoxalase II and its complex with a glutathione thioester substrate analogue. *Structure* 1999;**7**:1067–78.
21 **Gomes CM**, Frazao C, Xavier AV, Legall J, Teixeira M. Functional control of the binuclear metal site in the metallo-beta-lactamase-like fold by subtle amino acid replacements. *Protein Science* 2002;**11**:707–12.
22 **Pedersen CB**, Bross P, Winter VS, Corydon TJ, Bolund L, Bartlett K, Vockley J, Gregersen N. Misfolding, degradation, and aggregation of variant proteins. The molecular pathogenesis of short chain acyl-CoA dehydrogenase (SCAD) deficiency. *J Biol Chem* 2003;**28**:47449–58.
23 **Irsch T**, Krauth-Siegel RL. Glyoxalase II of African trypanosomes is trypanothione-dependent. *J Biol Chem* 2004;**279**:22209–17.
24 **Ridderstrom M**, Jemth P, Cameron AD, Mannervik B. The active site residue Tyr-175 in human glyoxalase II contributes to binding of glutathione derivatives. *Biochim Biophys Acta* 2000;**1481**:344–8.
25 **Daiyasu H**, Osaka K, Ishino Y, Toh H. Expansion of the zinc metallo hydrolase family of the beta-lactamase fold. *FEBS Lett* 2001;**10**:1–6.
26 **Cordell PA**, Futers TS, Grant PJ, Pease RJ. The Human hydroxyacylglutathione hydrolase (HAGH) gene encodes both cytosolic and mitochondrial forms of glyoxalase II. *J Biol Chem* 2004;**279**:28653–661.