J Inherit Metab Dis (2007) 30:217–226
DOI 10.1007/s10545-007-0429-x

ORIGINAL ARTICLE

Estimation of the total number of disease-causing mutations in ornithine transcarbamylase (OTC) deficiency. Value of the OTC structure in predicting a mutation pathogenic potential

J. A. Arranz · E. Riudor · C. Marco-Marín · V. Rubio

Received: 13 July 2006 / Submitted in revised form: 4 January 2007 / Accepted: 29 January 2007 / Published online: 1 March 2007

Summary Ornithine transcarbamylase deficiency (OTCD), the X-linked, most frequent urea cycle error, results from mutations in the *OTC* gene, encoding a 354-residue polypeptide. To date 341 OTCD clinical mutations, including 222 missense single nucleotide changes (mSNCs), have been compiled (*Hum Mutat* 2006;27:626). OTCD mutation detection might be simplified if the entire repertoire of OTCD-causing mutations were known. We estimate the size of this repertoire from 23 new OTCD patients exhibiting 22 different mutations, of which 9, including 4 mSNCs, are novel. The complete repertoire of OTCD-causing mutations is estimated as 560 mutations (95% confidence interval, 422–833 mutations), including 290 mSNCs (95% confidence interval, 230–394 mSNCs). Thus, OTCD diagnosis based on the screening for known mutations might attain ∼90% sensitivity in <5 years. Since disease-causing mSNCs represent <20% of the 2064 possible OTC mSNCs, simple approaches are essential for discrimination between causative and trivial mSNCs. Observation of the OTC structure appears a simple approach for such discrimination, comparing favourably in our sample with three formalized structure-based and/or sequence-based *in silico* assessment methods, and supporting the causation of complete deficiency by the mutations p.Pro305Arg and p.Ser96Phe, and of partial deficiency by p.Asp41Gly, p.Glu122Gly, p.Leu179Phe, p.Pro220Thr and p.Glu273del. Five non-mSNC novel mutations (p.Gly71X, a 7-nucleotide and a 10-nucleotide duplication and deletion in exon 5, G>A transitions at bases +1 and +5 of introns 4 and 9, respectively) are obviously pathogenic. The previously reported mSNCs p.Arg26Gln, p.Arg40His, p.Glu52Lys, pLys88Asn, p.Arg129His, p.Asn161Ser, p.Thr178Met, p.His202Tyr, p.Ala208Thr and p.His302Arg, found in our cohort, are also discussed.

Communicating editor: Alberto Burlina

Competing interests: None declared

References to electronic databases: OMIM #311250; EC 2.1.3.3; HUGO-approved gene symbol, *OTC*; GenBank reference sequence for human *OTC*: NP_000522 for the amino acid sequence, NM_000531 for the mRNA sequence, NC_000023 for the gene sequence; PDB entry (http://www.rcsb.org/pdb) 1OTH for the human OTC structure.

J. A. Arranz · E. Riudor
Unitat de Metabolopaties, Hospital Materno-Infantil Vall d'Hebron, Barcelona, Spain

C. Marco-Marín · V. Rubio (✉)
Instituto de Biomedicina de Valencia and CIBER de Enfermedades Raras, Valencia, Spain
e-mail: rubio@ibv.csic.es

Abbreviations

CI	confidence interval
mSNC	missense single nucleotide change
nSNC	nonsense single nucleotide change
OTC	ornithine transcarbamylase
OTCD	ornithine transcarbamylase deficiency
PALO	phosphonoacetyl-L-ornithine
SNC	single nucleotide change
SSCP	single-strand conformational polymorphism

Introduction

Ornithine transcarbamylase (OTC; EC 2.1.3.3) deficiency (OTCD; OMIM #311250) is X-linked (gene location Xp21.1) (Lindgren et al 1984) and is the most frequent inborn error of the urea cycle; in males it causes either neonatal hyperammonaemia or late-onset presentations depending, respectively, on whether the deficiency is complete or partial.

Female carriers can exhibit variable degrees of clinical deficiency (Brusilow and Horwich 2001). The *OTC* gene, of 73 kb, comprising 10 exons and 9 introns (Hata et al 1988), is expressed only in the liver and intestinal mucosa, yielding a 354-residue polypeptide which, upon passage to the mitochondrial matrix, matures by cleavage of the 32 residues at its N-terminus to a 322-amino acid polypeptide (Fig. 1A) that associates into the active trimer (Fig. 1B) (Brusilow and Horwich 2001), with involvement of residues from adjacent subunits in each active centre (Shi et al 1998, 2000, 2001).

Mutation identification is a key component of OTCD clinical diagnosis and is the gold standard for prenatal and female carrier diagnosis. However, since there is little recurrence of the numerous mutations identified thus far (341 compiled as at early 2006; Yamaguchi et al 2006), genetic diagnosis requires extensive DNA sequencing and is laborious and expensive. It would appear important to estimate how many more OTCD-causing mutations remain to be identified, since if a reasonably complete clinical mutation database were obtained, OTCD genetic diagnosis might be simplified by screening for known mutations, possibly using a DNA microchip strategy similar to those for other inborn diseases (Bhardwaj et al 2003; Galvin et al 2004; Green and Pass 2005; Tejedor et al 2005). Here we exploit the investigation of the causative mutation in 23 new, unrelated Spanish OTCD patients to estimate how large a fraction of the total mutation repertoire is represented in the recent compilation (Yamaguchi et al 2006) of 341 mutations, including 222 missense mutations due to single nucleotide changes (mSNCs). This database contains the mutations found in 14 of our 23 patients, including 13 of the 17 mSNCs represented in our sample, leading to estimates of the total number of mutations of any type and of mSNCs that cause OTCD that are not too far from the present numbers in the compilation, raising hopes that genetic diagnosis of OTCD based on screening for known mutations may be feasible soon. In addition, the much lower number of mSNCs causing OTCD than the total number of possible mSNCs for the *OTC* gene raises the urgent need for easy discrimination between disease-causing and trivial missense changes, by methods that should be very simple and accessible to every laboratory looking for mutations. In this context, we stress here the value of observing the OTC crystal structure (Shi et al 1998) by exemplifying the use of this approach with the novel or the less-characterized mutations found in our patients. Three formalized structure-based and/or sequence-based *in silico* assessment methods are also utilized, and give generally coincident results with the observational approach, although the latter appears more informative. In any case, these approaches should be used by all in the first line of assessment of the pathogenic potential of mutations found in OTCD patients.

Patients, materials and methods

Twenty-three unrelated Spanish patients (two of them, patients19 and 23, having a father of central or northern African ancestry, respectively), 15 males and 8 females, were referred for mutational study because of enzymatically proven and/or laboratory proven (high blood ammonia and glutamine, high orotate excretion, low plasma levels of citrulline) OTC deficiency, or because the clinical history and laboratory data (generally including a positive allopurinol test) were highly suggestive of this disease. Seven male patients presented early-onset deficiency, with hyperammonaemic coma on days 2–3 of life, and two additional male patients developed coma on days 7–8. The remaining 6 male and 8 female patients had characteristic late-onset presentations at ages ranging from 1.5 months to 9 years, and generally in these cases the presence of an elevated orotate excretion, either spontaneous or after allopurinol challenge, or the finding of decreased OTC activity (assayed colorimetrically at high substrate concentrations) (Nuzum and Snodgrass 1976) were key diagnostic elements.

Genomic DNA was extracted from peripheral leukocytes using a commercial kit (QIAamp DNA blood, from Qiagen Inc., Valencia, CA, USA) and the ten exons of the *OTC* gene and the flanking intron–exon boundaries were PCR-amplified (Garcia Perez et al 1995) and subjected to single-strand conformational polymorphism (SSCP) analysis at 15°C on the Genephor EF system with Genegel Clean 15/24 precast gels (both from Amersham Biosciences, Barcelona, Spain) following the manufacturer instructions; the amplified DNA fragments yielding abnormal SSCPs were sequenced using the ABI Prism 3100 automated sequencer (from Applied Biosystems, Madrid, Spain). When no abnormal SSCP was found or when a mutation was novel at the time of discovery, the ten *OTC* exons and flanking intronic regions were sequenced. Mutations were identified by comparison with the GenBank reference sequence for human *OTC* (GenBank entries: NP_000522 for the amino acid sequence; NM_000531 for the mRNA sequence; NC_000023 for exon–intron boundaries). The presence in a patient of a given mutation was confirmed by repeating the analysis in the product of an independent PCR reaction from the patient DNA. Except in the case of patient 4, maternal DNA was studied in all cases for the presence of the mutations found in the patients.

The 95% limits of the confidence interval (95% CI) for the proportion of the OTCD-causing mutations that are already included in the more recent OTCD mutation database (Yamaguchi et al 2006) were estimated from the expressions (Doménech i Massons 1980):

$$\text{Lower limit} = p_0 - 1.96\sqrt{(p_0 q_0 / n)}$$
$$\text{Upper limit} = p_0 + 1.96\sqrt{(p_0 q_0 / n)}$$

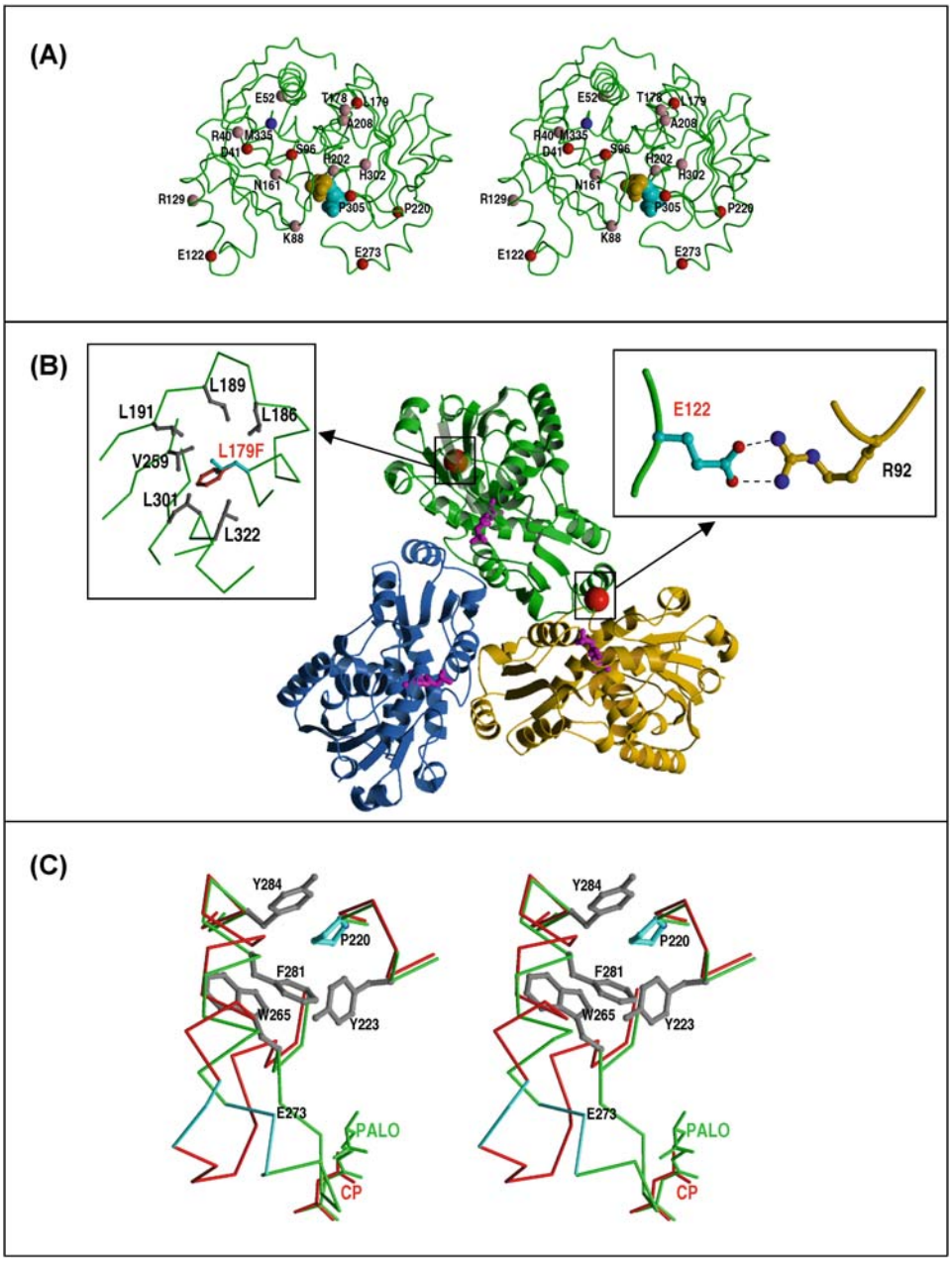

Fig. 1 *OTC* mutations found in the present patient cohort. Amino acids and mutations are given in single letter notation. (A) Stereo diagram showing the backbone trace of an OTC protein subunit with the transition-state analogue inhibitor PALO bound at the active centre of the enzyme. PALO is shown in space-filling representation, with the moieties corresponding to carbamylphosphate and ornithine coloured yellow and aqua, respectively. The coloured spheres mark the positions of the residues for which mutations are reported here; they are shown red if the impact of the mutation on structural bases is discussed here, blue for Met355 where a 28-residue insertion should occur in patient 5, and pink for other mutations. (B) Ribbon representation of the 3D structure of the OTC trimer containing three molecules of bound PALO (represented in ball-and-stick form and coloured magenta). Each subunit is coloured differently. Residues Glu122 and Leu179 are marked in one subunit as red spheres, and are boxed. The inset on the left expands (in main chain representation) the region around Leu179. The polypeptide chain is coloured green, the indicated amino acid side-chains are grey, the side-chain of Leu179 is aqua, and the side-chain of the mutant Leu179F is coloured red. The inset on the right details the Glu122-Arg92 intersubunit connection, which cannot be retained in the Glu122Gly mutant. One subunit is yellow and the other green. O and N atoms of the amino acid side-chains are coloured red and blue, respectively, and the side-chain of Glu122 is coloured aqua. The broken lines indicate hydrogen bonds within the salt bridge. (C) Detailed stereo view of the movement of the SMG loop upon binding of the two substrates (mimicked by PALO). The backbone representations of the loop in the complex with the bisubstrate analogue PALO (PDB file 1OTH) is coloured green and is superimposed with the red-coloured loop for the complex (PDB file 1EP9) with carbamoyl phosphate (abbreviated CP). PALO and CP are in stick representation, in green and red, respectively. The Pro220 ring and the portion of the main chain corresponding to Glu273 are coloured aqua, whereas the side chains of the indicated residues are grey

where p_0 is the fraction of the patients in the cohort who carry already reported mutations, $q_0 = 1 - p_0$, and n is the number of patients in the cohort. The corresponding upper and lower limits of the 95% CI for the absolute number of clinical mutations to be expected in OTCD are obtained by dividing the number of already reported mutations (Yamaguchi et al 2006) by, respectively, the lower or upper limit of the 95% CI for the proportion of reported mutations. The total number of possible single-nucleotide changes causing missense mutations (mSNCs) or nonsense mutations (nSNCs) for the *OTC* gene was estimated manually by changing one by one base in the OTC coding sequence for each one of the other three possible bases at the same position, counting all the single base changes that caused amino acid substitutions or that introduced stop codons, respectively, yielding values of 2064 possible mSNCs and 145 nSNCs.

Statements about conservation of residues are based on the alignment of OTC sequences from different species given at the Human Ornithine Transcarbamylase web page (http://www.cnmcresearch.org/otc/sequence_alignment.asp). This alignment was used for sequence-based *in silico* prediction of the effects of the mutations using the program SIFT (http://blocks.fhcrc.org/sift/SIFT.html) (Ng and Henikoff 2003). Unless indicated, all references to human OTC structure are based on the coordinates of entry 1OTH of the Protein Data Bank (http://www.rcsb.org/pdb) (Shi et al 1998). Visualization and analysis of the structures and of the consequences of the mutations was carried out with the programs RASMOL (Bernstein 2000) and O (Jones et al 1991). We also used a previously reported set of quantitative rules (Wang and Moult 2001) for predicting the effects of the mutations on the bases of 3D structure. In addition, the PolyPhen server (http://www.bork.embl-heidelberg.de/PolyPhen) (Ramensky et al 2002), which combines structure- and sequence-based prediction, was also used for prediction of the phenotypic consequences of the mSNCs. Figures have been prepared using Molscript (Kraulis 1991) and Raster 3D (Merritt and Murphy 1994).

Results and discussion

Mutations found (Tables 1 and 2)

Using an unbiased mutation search, 22 different mutations were identified in the 23 OTCD patients, where one mutation (p.Lys88Asn) recurred in patients 13 and 14 and none of the mutations was detected in any of other 100 unrelated X chromosomes. Nine mutations were novel (patients 1–9, Table 1), and 15 (68%) yielded an SSCP, supporting the current trend towards direct sequencing without SSCP analysis. In accordance with previous findings (Tuchman et al 2002), the mother was not a carrier and thus the mutation appeared to have arisen *de novo* (although the rare possibility of germline mosaicism was not excluded) with a higher frequency in females (3 of 8 females, or 38%) than in male patients (1 of 15 males, or ∼7%).

Three patients (13%) presented small insertions or deletions, and the other 20 (87%) exhibited SNCs. Among the latter, 4 (17%) had mutations at splice sites, two of them within the exon, which also caused a missense change, and two in intronic bases of the splice site. Only one patient (4%) presented a SNC causing a nonsense change. In the remaining 15 patients (65%) having SNCs, the mutation appears to cause a pure amino acid substitution, without any predicted splicing aberration (judged by searching for splice sites with the program SpliceSiteFinder (http://violin.genet.sickkids.on.ca/∼ali/splicesitefinder.html) and searching for splicing enhancers with the program ESEfinder (http://rulai.cshl.edu/tools/ESE)). Interestingly, 12 of the 14 missense mutations, or 86% of these mutations, occurred in exons 2 to 6, which represent only ∼55% of the coding sequence, and thus changes in the region encoded by exons 2 to 6 appear more frequently detrimental than changes occurring elsewhere (Figs. 1A and 1B localize the missense mutations in the OTC structure). The proportions of the patients exhibiting missense changes, splice site mutations and small insertions or deletions agrees grossly with the proportion of these mutations (70–80%, 10–20% and ∼10%, respectively) in previous OTCD mutation compilations (Tuchman et al 2002; Yamaguchi et al 2006). Of the 19 SNCs, 15 are transitions (seven G>A, and four of each of A>G and C>T) and only 4 are transversions (one of each of G>T, A>C, C>G and C>A). Although five SNCs affect one or other of the 19 CpG hotspots found in the OTC cDNA and splice sites, the p.Lys88Asn mutation recurring in two patients does not affect a CpG.

Estimation of the total number of expected clinical mutations in OTCD. Implications for the genetic diagnosis of OTCD

Fourteen (60.9%) patients, 13 of them presenting mSNCs (76.5% of the 17 patients having mSNCs), carry mutations that are reported in the most recent OTCD mutation database (Yamaguchi et al 2006) of 341 mutations of all types including 222 mSNCs. Therefore, ∼61% of mutations of all types (95% CI, 41–81%, see Materials and Methods) and ∼76% of mSNCs (95% CI, 56–97%) appear to be included already in this database, and the complete mutational repertoire is thus estimated to consist of 560 mutations of all types (95% CI, 422–833 mutations), including 290 mSNCs (95% CI, 230–394 mSNCs). Additional support for these estimates is provided by the similarity of the figures obtained for another cohort consisting of 26 Korean OTCD patients reported

Table 1 Novel mutations found in patients with ornithine transcarbamylase deficiency[a]

Patient	Sex	Age of onset[b]	Nucleotide change/ codon change[c]	Nature of change	Exon/Intron	Restriction site alteration[d]	SSCP[e]	Inheritance[f]
1	M	13 months	c.365A>G 122GAA→GGA	p.Glu122Gly	E4		+	Inherited
2[g]	M	8 days	c.386+5G>A	Donor splice site error E4/I4	I4	+ *Psi*I	−	Inherited
3[g]	M	Neonatal	c.516_525 delCCTGGCTGAT	p.Leu173ThrfsX11	E5	− *Mva*I − *Eco*RII	+	Inherited
4	M	7 days	c.535C>T 179CTC→TTC	p.Leu179Phe	E5	− *Bpm*I	−	(Inherited)
5[h]	M	Neonatal	c.1005+1G>A	Donor splice site error E9/I9	I9	+ *Bsp*HI	−	Inherited
6	F	9 months	c.211G>T 71GGA→TGA	p.Gly71X	E2		−	*de novo*
7	F	18 months	c.287C>T 96TCT→TTT	p.Ser96Phe	E3	− *Bst*SFI − *Sfc*I	−	Inherited (*de novo* in mother)
8	F	22 months	c.391_397dupTTGTCTA	p.Ser133IlefX3	E5		+	Inherited
9	F	2 years	c.658C>A 220CCA→ACA	p.Pro220Thr	E6	− *Bsc*4I − *Bsi*YI	+	Inherited

[a]None of the nucleotide changes was found in any of 100 unrelated chromosomes
[b]"Neonatal" indicates the development of hyperammonaemic coma at days 2–3 of life
[c]The figures before the triplets correspond to codon numbers
[d]A − or + sign before the restriction enzyme indicates the respective elimination or creation of the indicated site
[e]A positive sign indicates that the mutation gave an abnormal SSCP
[f]Inheritance of the mutation was ascertained by maternal DNA analysis, except in the case in parentheses, in which no maternal DNA could be obtained but in which the inherited nature of the deficiency was deduced from the positivity of the allopurinol test in the mother. In the cases in which the mutation is considered to have arisen *de novo*, the rare possibility that the mutations were inherited from a progenitor exhibiting germline mosaicism has not been excluded
[g]The deficiency was confirmed by a very low or undetectable ornithine transcarbamylase activity in the liver or in intestinal mucosa
[h]Liver ornithine transcarbamylase activity was 14% of normal

as this paper was under revision (Kim et al 2006): 61.5% of the mutations found in these patients, including 68.4% of the mSNCs, are present already in the database. Using both cohorts combined, closely similar figures but narrower confidence intervals are estimated for the complete repertoire of mutations causing OTCD: 557 mutations (95% CI 455–716 mutations), including 307 mSNCs (95% CI 256–385 mSNCs). Therefore, we are reasonably close, certainly within <5 years (for a rate of detection of ∼175 OTCD mutations per year in the western world and Japan, with only 1/5 of the mutations being novel mSNCs), to compiling a complete mutational repertoire, particularly concerning mSNCs; thus, to reach this objective it is important to report all novel mutations. Diagnosis based on the screening for known mutations might attain ∼90% sensitivity (excluding from diagnosis insertions and deletions, believed to represent ∼10% of the mutations causing OTCD; Tuchman et al 2002) by searching for <1000 mutations: <385 mSNCs (the upper limit of the narrower 95% CI), supplemented with 145 nonsense SNCs (this number of nSNCs was determined as indicated in Materials and Methods) and the 378 possible SNCs affecting the seven bases at exon–intron junctions and thus likely triggering splicing errors.

Assessment of the disease-causing potential of the novel mutations (Fig. 1)

Since there are 2064 possible mSNCs in the OTC gene (calculated as indicated in Materials and Methods), even the upper limit of the 95% CI for disease-causing mSNCs only represents a small fraction (<1/5) of all possible mSNCs, and thus most mSNCs appear to be trivial amino acid substitutions that should not cause OTCD. Therefore, it is crucial to distinguish between clinical mutations and mere polymorphisms. *In vitro* expression studies (Morizono et al 1997), although valuable for this purpose, appear impractical for very extensive mutation repertoires as in OTCD. Structure-based analysis is simpler, it has been used already for OTCD (Shi et al 1998), and although of a conjectural nature, it is based on our understanding of the forces that guide protein folding and assembly, substrate binding and catalysis, providing a rationale for the effects of the mutations, as is illustrated here for the present novel or most recently described, poorly characterized, single-amino-acid mutations (Fig. 1A, in red). For example, Pro305 belongs to the constant ornithine binding signature HCLP, it forms a rare *cis*-peptidic bond that is exclusively found with proline, and

Table 2 Previously reported mutations found in our patients with ornithine transcarbamylase deficiency[a]

Patient	Sex	Age of onset[b]	Nucleotide change /codon change[c]	Nature of change	Exon/Intron	SSCP[d]	Inheritance[e]	First description
10	M	Neonatal	c.77G>A 26CGG→CAG Last base of exon 1 CpG dinucleotide	p.Arg26Gln Donor splice site error at E1/I1	E1	+	Inherited	Grompe et al (1989)
11[f]	M	9 years	c.119G>A 40CGT→CAT CpG dinucleotide	p.Arg40His	E2	+	Inherited	Tuchman et al (1994)
12	M	Neonatal surviving	c.122A>G 41GAC→GGC	p.Asp41Gly	E2	+	*de novo*	Yamaguchi et al (2006)
13[f]	M	16 months	c.264A>C 88AAA→AAC	p.Lys88Asn	E3	+	Inherited	Reish et al (1993)
14	M	28 months	c.264A>C 88AAA→AAC	p.Lys88Asn	E3	+	Inherited	Reish et al (1993)
15	M	Neonatal	c.482A>G 161AAT→AGT	p.Asn161Ser	E5	+	Inherited	Tuchman and Plante (1995)
16	M	16 months	c.604C>T 202CAC→TAC	p.His202Tyr	E6	+	Inherited	Tuchman et al (1997)
17[f]	M	3.5 years	c.622G>A 208GCA→ACA CpG dinucleotide	p.Ala208Thr	E6	+	Inherited	Van Diggelen et al (1996)
18	M	Neonatal	c.905A>G 302CAC→CGC	p.His302Arg	E9	−	Inherited	Genet et al (2000)
19[f]	M	Neonatal	c.914C>G 305CCC→CGC	p.Pro305Arg	E9	+	Inherited	Yamaguchi et al (2006)
20	F	14 months	c.154G>A 52 GAA→AAA	p.Glu52Lys	E2	+	*de novo*	McCullough et al (2000)
21	F	6 years	c.386G>A 129CGG→CAG Last base of exon 4 CpG dinucleotide	p.Arg129His Donor splice site error at E4/I4	E4	+	Inherited	Garcia-Pérez et al (1995)
22	F	4.5 years	c.533C>T 178ACG→ATG CpG dinucleotide	p.Thr178Met	E5	−	Inherited	Oppliger Leibundgut et al (1995)
23	F	1.5 months	c.817_819delGAG	p.Glu273del	E8	+	*de novo*	Segues et al (1996)

[a]None of the nucleotide changes was found in any of 100 unrelated chromosomes
[b]"Neonatal" indicates the development of hyperammonaemic coma at days 2–3 of life
[c]The figures before the triplets correspond to codon numbers
[d]A positive sign indicates that the mutation gave an abnormal SSCP
[e]Inheritance of the mutation was ascertained by maternal DNA analysis. In the cases in which the mutation is considered to have arisen *de novo*, the rare possibility that the mutations were inherited from a progenitor exhibiting germline mosaicism has not been excluded
[f]Ornithine transcarbamylase activity of intestinal mucosa was borderline in patient 11, 14% and 16% of normal in patients 13 and 17, respectively, and undetectable in patient 19

largely determines the shape of the ornithine site. Consequently, mutations at Pro305 should be highly deleterious, as observed in the complete deficiency presented by patient 19 (mutation p.Pro305Arg) and by a previous male patient carrying the p.Pro305His mutation (Climent and Rubio 2002). In contrast, the mutations p.Glu122Gly, p.Asp41Gly and p.Leu179Phe are supported as causing the partial deficiencies observed in male patients 1, 12 and 4, since these mutations should merely decrease enzyme stability or trigger modest structural distortions. p.Glu122Gly should eliminate the intersubunit interactions between the side-chain carboxylate of Glu122 and the guanidinium group of Arg92 of another subunit (Fig. 1B), rendering less stable the OTC trimer that is the active form of the enzyme. The mutation p.Asp41Gly eliminates one of the three hydrogen bonds that interlink strands 2 and 5 of the carbamoyl phosphate domain central β-sheet, and should decrease the stability of this domain. For mutation p.Leu179Phe, the bulky and rigid phenylalanine side-chain should cause some mild steric distortion in the densely packed and hydrophobic core of the ornithine

domain (Fig. 1B) without triggering complete misfolding and total inactivation. In contrast, the mutation p.Ser96Phe, observed here in a female (patient 7), should in males cause complete deficiency because of gross misfolding of the OTC subunit. Ser96 is at the fringe between helix 2 and helix 11, one of the helices that interconnect the two domains of the OTC subunit (Fig. 1A), and the residue corresponding to human Ser96 is consistently small in all OTCs. Therefore, the replacement of Ser96 by the large phenylalanine residue in the p.Ser96Phe mutation may be expected to cause mispositioning of helix 11, distorting the OTC subunit domain architecture. The mutations p.Pro220Thr and p.Glu273del, found also in females (patients 9 and 23), should, in contrast, in males cause partial deficiency, because they are expected to disturb the SMG loop (Fig. 1C), a superficial mobile loop that caps the bound ornithine in the ternary complex of the enzyme with its two substrates (Shi et al 2001). Since this loop does not constitute the core of the ornithine binding site, its alterations may be expected to hamper but not to abolish ornithine binding, causing partial deficiency. The mutation p.Pro220Thr may abolish the anchoring of both ends of the SMG loop on the enzyme body, since this anchoring is mediated by the interactions of the ring of Pro220 with Trp284 and Phe281, and of Phe281 with Trp265. A male patient carrying another mutations at this proline, p.Pro220Ala, exhibited partial deficiency (Oppliger Leibundgut et al 1996). The mutation p.Glu273del shortens the SMG loop and was observed in a previous male patient exhibiting partial OTCD (Segues et al 1996).

The application of a previously published (Wang and Moult 2001) set of quantitative rules for structure-based assessment of the consequences of amino acid substitutions yields predictions that are coincident, although frequently less specific, than with our observational analysis for all but one of these mutations (p.Pro220Thr) (Table 3). Use of the SIFT prediction server (Ng and Henikoff 2003) (Table 3), which is based exclusively on the conservation of the amino acid sequence, or of the PolyPhen server (Ramensky et al 2002), which combines sequence and 3D structure criteria, resulted in predictions of negative effects by the p.P220T mutation, supporting our conclusion that the p.Pro220Thr mutation is disease-causing. It is interesting that the exclusive use of sequence conservation data by the SIFT server is enough for prediction of a negative effect for all but two mutations (p.Asp41Gly and p.Leu179Phe), although this server as well as the PolyPhen server fails to provide any information on the nature of effect expected. Overall, the results of these three *in silico* methods (Table 3) are mostly coincident with our analysis. Only in the case of the p.Leu179Phe mutation did two of the methods, those provided by the SIFT and the PolyPhen servers, fail to predict an effect. However, the application of the quantitative rules of the structure-based method (Wang and Moult 2001) also predicts for this mutation effects on protein stability, as in our analysis.

The effects of the five novel mutations that do not cause pure single amino acid replacements are self-evident. Thus, the mutation p.Gly71X, found in patient 6, and the 10-base deletion and 7-base duplication in exon 5 found in patients 3 and 8, causing frameshifts and premature termination, truncate large regions of the enzyme and are obviously inactivating, as illustrated by the complete deficiency observed in patient 3 (the only male among these three patients). The other two novel mutations are G>A transitions at the +5 and +1 positions of introns 4 and 9 (male patients 2 and 5; Table 1), two mutations that are well-represented among the intronic changes reported in clinical mutation databases (Antonarakis et al 2001), thus being expected to be disease-causing. Since the complete abolition of intron 4 splicing causes enzyme truncation and total OTC inactivation (Garcia-Pérez et al 1995), the delayed (day 8) clinical presentation with ∼1% intestinal OTC activity in patient 2 (Table 1) indicates that the G>A transition at base +5 does not completely abolish normal intron 4 splicing. The dramatic neonatal clinical presentation observed in patient 5 would appear inconsistent with the relatively high (∼14% of normal) OTC activity found in the liver of this patient (Table 1), strongly suggesting that the OTC is in this case active but is kinetically abnormal and is inefficient *in vivo* (Raijman 1976). In fact, lack of normal intron 9 splicing should cause aberrant splicing at the cryptic internal donor splice site CAA/GTGAGA (Shapiro and Senapathy 1987), resulting in the introduction of an in-frame 28-amino-acid insertion after Met335 (Met335 is shown as a blue sphere in the OTC subunit structure, Fig. 1A). Since this insertion is not localized either in the active centre, at the intersubunit surface or at the core of the enzyme, the aberrant OTC may be active but abnormal, as suggested by the observations in patient 5. Overall, simply the inspection of the mutation or, for mSNCs, more in-depth structural analysis, supports the disease-causing nature of all the novel mutations found here, illustrating the potential of these simple approaches, alone or combined with the utilization of the formalized *in silico* approaches (Ng and Henikoff 2003; Ramensky et al 2002; Wang and Moult 2001) shown in Table 3. In addition, the pathogenic nature of the mutations p.Asp41Gly, p.Gly71X and p.Glu273del is supported by the *de novo* appearance of these mutations in the corresponding patient (although, as indicated above, the rare possibility of germline mosaicism has not been excluded), who in these three cases was the first member of the family to manifest clinical OTCD.

Further observations on previously reported single nucleotide changes (Fig. 1A)

In 2 of the 13 patients presenting mSNCs, the mutation affects the last base of either exon 1 (patient 10) or exon

Table 3 Comparison of the results of the present assessment of the effects of mutations causing single amino acid changes with the predictions of three published *in silico* methods

		Wang and Moult method[a]		SIFT server[b]		
Amino acid change	Present structure-based analysis	Structural effect	Molecular function effect	Score	Effect	PolyPhen server[c]
Asp41Gly	Decreased stability of CP domain	Loss of hydrogen bonds	Protein stability	0.46	Tolerated	Possibly damaging
Glu122Gly	Decreased stability of OTC trimer resulting from loss of intersubunit salt bridge	Loss of salt bridge Destabilization of protein multimer	Protein stability	0.00	Affected	Probably damaging
Leu179Phe	Steric distortion in the densely packed and hydrophobic core of the ornithine domain	Overpacking	Protein stability	0.08	Tolerated	Benign
Pro305Arg	Abolished ornithine binding	Backbone strain	Protein stability Ligand binding	0.00	Affected	Probably damaging
Ser96Phe	Mispositioning of helix 11 and gross misfolding	Loss of hydrogen bonds Overpacking	Protein stability	0.00	Affected	Probably damaging
Pro220Thr	Defective SMG loop function, hampering ornithine binding	No prediction of effect	No prediction of effect	0.00	Affected	Probably damaging
Glu273del[d]	Defective SMG loop function, hampering ornithine binding	Loss of hydrogen bonds Loss of salt bridge	Protein stability	0.01	Affected	Probably damaging

[a]Structure-based method. A fixed quantitative set of rules is used (Wang and Moult 1991)
[b]Server-supported (http://blocks.fhcrc.org/sift/SIFT.html) sequence conservation-based method (Ng and Henikoff 2003). Scores >0.05 are considered to correspond to tolerated changes, whereas scores <0.05 are predicted to affect protein function
[c]Server-supported (http://www. bork.embl-heidelberg.de/PolyPhen) method. It makes use of both the 3D structure and sequence conservation data (Ramensky et al 2002) to yield the following possible effect categories: benign, possibly damaging, probably damaging, unknown. This server has also been used to estimate, from 100 randomly chosen mutations among the 2064 possible mSNCs for the *OTC* gene, what fraction of these mutations is expected to cause disease. Only 26% of the mutations were found to be probably damaging, 9% were categorized as possibly damaging, 40% were considered benign, and in 26% the effect was unknown (no prediction)
[d]Since the three published methods are devised for missense changes rather than for deletions, when using these methods we have equated the Glu273del mutation to a Glu273Gly mutation. Therefore, the predictions correspond to the deletion of the side-chain of Glu273

4 (patient 21) (Table 2) and also causes abnormal splicing (Garcia-Pérez et al 1995; Grompe et al 1989). Two characteristic OTC-inactivating mutations, p.Asn161Ser and p.His302Arg, are represented among our patients (patients 15 and 18, respectively). These mutations affect active-centre invariant residues (Shi et al 1998) including the ornithine binding signature (HCLP) residue His302 which belongs to the catalytic charge relay system (Jin et al 1997). The very characteristic mild mutations p.Arg40His and p.Ala208Thr are also represented in our sample, in patients 11 and 17, respectively. These mutations have been reported in paternal transmission and were found to cause overt disease only in some males (Matsuda et al 1996; Van Diggelen et al 1996). The latter is also observed here in patient 11, who has a healthy 4-year-old brother carrying the p.Arg40His mutation. The only mutation recurring within this series, p.Lys88Asn, found in two unrelated male patients (patients 13 and 14) exhibiting similar late-onset presentations, yielded a relatively high OTC activity ($\sim$15% of normal; assayed in patient 13), suggesting the presence of a kinetically abnormal enzyme, according to the proposal (Shi et al 1998) that p.Lys88Asn impairs but does not abolish ornithine binding. In agreement with a previous report (Tuchman et al 1997), the mutation p.His202Tyr caused partial deficiency in patient 16. Two mutations found here in female patients, p.Glu52Lys (patient 20) and p.Thr178Met (patient 22), when reported in males caused, respectively, complete and partial OTCD (McCullough et al 2000; Oppliger Leibundgut et al 1995). Overall, the clinical presentations in our male patients replicate those in previous reports of the same mutations, supporting the constancy of the gross presentation for a given OTC mutation.

Acknowledgements Supported by Red REDEMETH G03/054 and project PI052838 of the Instituto de Salud Carlos III of the Spanish Ministry of Health. We thank Paz Briones (Institut de Bioquímica

Clínica, Barcelona) for OTC activity assay, and the clinicians and hospitals that provided the patients and cared for them: M. T. García-Silva and E. Martín (Hospital 12 de Octubre), M. Martínez-Pardo (Hospital Ramón y Cajal), R. A. Lama (Hospital La Paz), C. Pedrón (Hospital U. Niño Jesús), M. J. García-Muñoz (Centro de Diagnóstico de Enfermedades Moleculares, CBM-UAM) and P. Jara (Hospital La Paz), all from Madrid; M. A. Vilaseca (Hospital Sant Joan de Deu), A. Ribes (Institut de Bioquímica Clínica), G. Pintos (Hospital Germans Trias i Pujol), M. Torrent (Hospital de la Santa Creu i Sant Pau) and M. del Toro (Hospital Vall d'Hebron), all from Barcelona; J. Dalmau and A. M. García (Hospital La Fe) and A. Molina (Hospital Clínico Universitario), from Valencia; C. Sierra (Hospital Carlos Haya, Málaga); M. A. Ramos (Hospital Virgen del Camino, Pamplona); A. Trujillo (Hospital Josep Trueta, Girona); L. Peña (Hospital de Las Palmas, Gran Canaria). We also thank the Unitat Científico-Tècnica de Support de l'Institut de Recerca de l'Hospital Universitari Vall d'Hebron for DNA sequencing.

References

Antonarakis SE, Krawczak M, Cooper DN (2001) The nature and mechanisms of human gene mutation. In: Scriver CR, Beaudet Al, Sly WS, Valle D, eds; Childs B, Kinzler KW, Vogelstein B, assoc, eds. *The Metabolic and Molecular Bases of Inherited Disease*, 8th edn. New York: McGraw-Hill, 343–377.

Bernstein HJ (2000) Recent changes to RasMol, recombining the variants. *Trends Biochem Sci* **25**: 453–455.

Bhardwaj U, Zhang YH, McCabe ER (2003) Neonatal hemoglobinopathy screening: molecular genetic technologies. *Mol Genet Metab* **80**: 129–137.

Brusilow SW, Horwich AL (2001) Urea cycle enzymes. In: Scriver CR, Beaudet Al, Sly WS, Valle D, eds; Childs B, Kinzler KW, Vogelstein B, assoc, eds. *The Metabolic and Molecular Bases of Inherited Disease*, 8th edn. New York: McGraw-Hill, 1909–1963.

Climent C, Rubio V (2002) Identification of seven novel missense mutations, two splice-site mutations, two microdeletions and a polymorphic amino acid substitution in the gene for ornithine transcarbamylase (OTC) in patients with OTC deficiency. *Hum Mutat* **19**: 185–186.

Doménech i Massons JM (1980) *Bioestadística. Métodos estadísticos para investigadores.* Barcelona: Herder, 61–91.

Galvin P, Clarke LA, Harvey S, Amaral MD (2004) Microarray analysis in cystic fibrosis. *J Cys. Fibros* **3**(Suppl 2): 29–33.

Garcia-Perez MA, Sanjurjo P, Rubio V (1995) Demonstration of the spf-ash mutation in Spanish patients with ornithine transcarbamylase deficiency of moderate severity. *Hum Genet* **95**: 183–186.

Genet S, Cranston T, Middleton-Price HR (2000) Mutation detection in 65 families with a possible diagnosis of ornithine carbamoyltransferase deficiency including 14 novel mutations. *J Inherit Metab Dis* **23**: 669–676.

Green NS, Pass KA (2005) Neonatal screening by DNA microarray: spots and chips. *Nat Rev Genet* **6**: 147–151.

Grompe M, Muzny DM, Caskey CT (1989) Scanning detection of mutations in human ornithine transcarbamoylase by chemical mismatch cleavage. *Proc Natl Acad Sci USA* **86**: 5888–5892.

Hata A, Tsuzuki T, Shimada K, Takiguchi M, Mori M, Matsuda I (1998) Structure of the human ornithine transcarbamylase gene. *J Biochem (Tokyo)* **103**: 302–308.

Jin L, Seaton BA, Head JF (1997) Crystal structure at 2.8 Å resolution of anabolic ornithine transcarbamylase from *Escherichia coli*. *Nature Struct Biol* **4**: 622–625.

Jones TA, Zou J, Cowan S, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the localization of errors in these models. *Acta Crystallogr Sect A* **47**: 110–119.

Kim GH, Choi JH, Lee HH, Park S, Kim SS, Yoo HW (2006) Identification of novel mutations in the human ornithine transcarbamylase (OTC) gene of Korean patients with OTC deficiency and transient expression of the mutant proteins in vitro. *Hum Mutat* **27**: 1159.

Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* **24**: 946–950.

Lindgren V, de Martinville B, Horwich AL, Rosenberg LE, Francke U (1984) Human ornithine transcarbamylase locus mapped to band Xp21.1 near the Duchenne muscular dystrophy locus. *Science* **226**: 698–700.

Matsuda I, Matsuura T, Nishiyori A, et al (1996) Phenotypic variability in male patients carrying the mutant ornithine transcarbamylase (OTC) allele, Arg40His, ranging from a child with an unfavourable prognosis to an asymptomatic older adult. *J Med Genet* **33**: 645–648.

McCullough BA, Yudkoff M, Batshaw ML, Wilson JM, Raper SE, Tuchman M (2000) Genotype spectrum of ornithine transcarbamylase deficiency: correlation with the clinical and biochemical phenotype. *Am J Med Genet* **93**: 313–319.

Merritt EA, Murphy MEP (1994) Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr Sect D* **50**: 869–873.

Morizono H, Tuchman M, Rajagopal BS, et al (1997) Expression, purification and kinetic characterization of wild-type human ornithine transcarbamylase and a recurrent mutant that produces 'late onset' hyperammonaemia. *Biochem J* **322**: 625–631.

Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31**: 3812–3814.

Nuzum CT, Snodgrass PJ (1976) Multiple assays of the five urea-cycle enzymes in human liver homogenates. In: Grisolia S, Báguena R, Mayor F, eds. *The Urea Cycle*. New York: Wiley, 325–349.

Oppliger Leibundgut EO, Liechti-Gallati S, Colombo JP, Wermuth B (1995) Ornithine transcarbamylase deficiency: new sites with increased probability of mutation. *Hum Genet* **95**: 191–196.

Oppliger Leibundgut EO, Wermuth B, Colombo JP, Liechti-Gallati S (1996) Ornithine transcarbamylase deficiency: characterization of gene mutations and polymorphisms. *Hum Mutat* **8**: 333–339.

Raijman L (1976) Enzyme and reactant concentrations and the regulation of urea synthesis. In: Grisolia S, Báguena R, Mayor F, eds. *The Urea Cycle*. New York: Wiley, 243–254.

Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNCs: server and survey. *Nucleic Acids Res* **30**: 3894–3900.

Reish O, Plante RJ, Tuchman M (1993) Four new mutations in the ornithine transcarbamylase gene. *Biochem Med Metab Biol* **50**: 169–175.

Segues B, Veber PS, Rabier D, et al (1996) A 3-base pair in-frame deletion in exon 8 (delGlu272/273) of the ornithine transcarbamylase gene in late-onset hyperammonemic coma. *Hum Mutat* **8**: 373–374.

Shapiro MB, Senapathy P (1987) RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* **15**: 7155–7174.

Shi D, Morizono H, Aoyagi M, Tuchman M, Allewell NM (1998) 1.85-Å resolution crystal structure of human ornithine transcarbamoylase complexed with *N*-phosphonacetyl-L-ornithine. *J Biol Chem* **273**: 34247–34254.

Shi D, Morizono H, Aoyagi M, Tuchman M, Allewell NM (2000) Crystal structure of human ornithine transcarbamylase complexed with carbamoyl phosphate and L-norvaline at 1.9 Å resolution. *Proteins* **39**: 271–277.

Shi D, Morizono H, Yu X, Tong L, Allewell NM, Tuchman M (2001) Human ornithine transcarbamylase: crystallographic insights into substrate recognition and conformational changes. *Biochem J* **354**: 501–509.

Tejedor D, Castillo S, Mozas P, et al (2005). Reliable low-density DNA array based on allele-specific probes for detection of 118 mutations causing familial hypercholesterolemia. *Clin Chem* **51**: 1137–1144.

Tuchman M, Plante RJ (1995) Mutations and polymorphisms in the human ornithine transcarbamylase gene: mutation update addendum. *Hum Mutat* **5**: 293–295.

Tuchman M, Plante RJ, McCann MT, Qureshi AA (1994) Seven new mutations in the human ornithine transcarbamylase gene. *Hum Mutat* **4**: 57–60.

Tuchman M, Morizono H, Rajagopal BS, Plante RJ, Alewell NM (1997) Identification of 'private' mutations in patients with ornithine transcarbamylase deficiency. *J Inherit Metab Dis* **20**: 525–527.

Tuchman M, Jaleel N, Morizono H, Sheehy L, Lynch MG (2002) Mutations and polymorphisms in the human ornithine transcarbamylase (OTC) gene. *Hum Mutat* **19**: 93–107.

Van Diggelen OP, Zaremba J, He W, et al (1996) Asymptomatic and late-onset ornithine transcarbamylase (OTC) deficiency in males of a five-generation family, caused by an A208T mutation. *Clin Genet* **50**: 310–316.

Wang Z, Moult J (2001) SNCs, protein structure, and disease. *Hum Mutat* **17**: 263–270.

Yamaguchi S, Brailey LL, Morizono H, Bale AE, Tuchman M (2006) Mutations and polymorphisms in the human ornithine transcarbamylase (OTC) gene. *Hum Mutat* **27**: 626–632.