Diez-Fernandez C., Gallego J., Haberle J., Cervera J., Rubio V. (2015). The Study of Carbamoyl Phosphate Synthetase 1 Deficiency Sheds Light on the Mechanism for Switching On/Off the Urea Cycle. J Genet Genomics 42(5): 249-60.
Author's last version

The Study of Carbamoyl Phosphate Synthetase 1 Deficiency Sheds Light on the Mechanism for Switching On/Off the Urea Cycle

**Carmen Díez-Fernández,^a José Gallego,^b Johannes Häberle,^c Javier Cervera,^{a,d}
Vicente Rubio^{a,d,*}**

^aInstitute for Biomedicine of Valencia of the Spanish Research Council, Valencia 46010, Spain

^bMedical School of the Catholic University of Valencia, Valencia 46001, Spain

^cUniversity Children's Hospital Zurich and Children's Research Center, Zurich 8032, Switzerland

^dCentre for Biomedical Networking Research on Rare Diseases of the Carlos III Health Institute, Valencia 46010, Spain

*Corresponding author.

Tel: +34 96 3391 772, fax: +34 96 3690 800; rubio@ibv.csic.es (V. Rubio)

Abstract

Carbamoyl phosphate synthetase 1 (CPS1) deficiency (CPS1D) is an inborn error of the urea cycle having autosomal (2q34) recessive inheritance that can cause hyperammonemia and neonatal death or mental retardation. We have analyzed the effects on CPS1 activity, kinetic parameters and enzyme stability of missense mutations reported in patients with CPS1 deficiency that map in the 20-kDa C-terminal domain of the enzyme. This domain turns on or off the enzyme depending on whether the essential allosteric activator of CPS1, N-acetyl-L-glutamate (NAG), is bound or is not bound to it. To carry out the present studies, we have exploited a novel system that allows the expression *in vitro* and the purification of human CPS1, thus permitting site-directed mutagenesis. These studies have clarified disease causation by individual mutations, identifying functionally important residues, and revealing that a number of mutations decrease the affinity of the enzyme for NAG. Patients with NAG affinity-decreasing mutations might benefit from NAG site saturation therapy with N-carbamyl-L-glutamate (a registered drug, analogue of NAG). Our results, together with additional present and prior site-directed mutagenesis data for other residues mapping in this domain, suggest an NAG-triggered conformational change in the $\beta 4$ - $\alpha 4$ loop of the C-terminal domain of this enzyme. This change might be an early event in the NAG activation process. Molecular dynamics simulations that were restrained according to the observed effects of the mutations are consistent with this hypothesis, providing further backing for this structurally plausible signaling mechanism by which NAG could trigger urea cycle activation *via* CPS1.

Keywords: Urea cycle diseases; Inborn errors; Hyperammonemia; Site-directed mutagenesis; Restrained molecular dynamics; Allosteric regulation; Carbamoyl phosphate synthetase 1; Enzyme.

INTRODUCTION

Carbamoyl phosphate synthetase 1 (CPS1) deficiency (CPS1D; OMIM #237300), a recessively inherited autosomal (2q34) (McReynolds et al., 1981) inborn error of the urea cycle (Freeman et al., 1964; Gelehrter and Snodgrass, 1974), has an estimated incidence of 1/50000 to 1/300000 (Uchino et al., 1998; Summar et al., 2013). CPS1 is the entry point of ammonia, the nitrogenous waste product of protein catabolism, into the urea cycle (Fig. 1A). Therefore, CPS1D causes pure hyperammonemia (Häberle and Rubio, 2014), leading to encephalopathy and even death (Brusilow and Horwich, 2001), and to depletion of downstream urea cycle intermediates, particularly of citrulline (Häberle and Rubio, 2014).

A large repertory of mutations affecting the *CPS1* gene (OMIM #608307; 201,425 nucleotides; start/end chromosome 2 coordinates, 211,342,405/211,543,830, plus strand; <http://www.genecards.org/cgi-bin/carddisp.pl?gene=CPS1>) has been compiled from patients with CPS1D (Häberle et al., 2011). Over 50% of these mutations are missense changes spreading over the entire 1462-residue mature CPS1 polypeptide (Nyunoya et al., 1985; Haraguchi et al., 1991). The *CPS1* gene encompasses 4500 coding nucleotides over 38 exons (Funghini et al., 2003; Häberle et al., 2003; Summar et al., 2003). It may be difficult to ascertain the responsibility of a given *CPS1* missense mutation in causing CPS1D, particularly for mutations mapping outside the two catalytic domains of the enzyme (the two phosphorylation domains, Fig. 1B) which bind the substrates and catalyze the three-step CPS1 reaction (Alonso et al., 1992; Alonso and Rubio, 1995).

Our present work dealt with the analysis of the effects of CPS1D-associated mutations (called here clinical mutations) that affect a non-catalytic domain of human CPS1, the C-terminal domain of 20 kDa (Häberle et al., 2011). This domain is called the allosteric domain (abbreviated ASD) (Fig. 1B) because it binds N-acetyl-L-glutamate (NAG) (Rodríguez-Aparicio et al., 1989; Pekkala et al., 2009), the essential allosteric activator of CPS1. Without NAG, CPS1 is inactive (Rubio et al., 1981, 1983), possibly reflecting the need to stop catalysis by the enzyme (Shigesada et al., 1978; Stewart and Walser, 1980) before ammonia levels are too low (Fig. 1A). Too much decrease in the ammonia level would lead to depletion of ammonia-derived amino acids such as glycine, glutamate and glutamine (Bender, 2011), and possibly to protein catabolism. NAG is a proper effector of the CPS1 switch because its levels reflect the nitrogen burden manifested in the glutamate level (Fig. 1A). This is so because NAG has a short half-life (Morita et al., 1982) and it is made from glutamate by an enzyme (NAG synthase) exhibiting a high K_m for glutamate (Sonoda and Tatibana, 1983).

To explore the effects of ASD-mapping missense clinical mutations (abbreviated ASD clinical mutations), we utilized a novel system for production and mutagenesis of recombinant CPS1 that uses baculovirus and insect cells (Díez-Fernández et al., 2013). We had already applied this system to analyze the effects of some ASD missense mutations (Pekkala et al., 2010; Díez-Fernández et al., 2013). We now extend this analysis to all the reported ASD clinical mutations, as well as to some mutations designed to test the role of the ASD (Fig. 1B, vertical lines, and Tables 1 and 2). Analysis of the effects of these mutations has helped assess disease causality, opening the way to improved genetic counselling and even to individualized therapy. Furthermore, we now shed some light on the as yet unclarified NAG activation process, and define better the NAG site. We had previously localized the NAG site (Fig. 1C)

(Pekkala et al., 2009) by photoaffinity labeling and by *in silico* docking in the deposited (but unpublished) crystal structure of the isolated human ASD free from NAG [Protein Databank (PDB; www.rcsb.org) file 2YVQ; Xie et al., 2007]. We now have found that some ASD mutations have effects that are inconsistent with the previously proposed NAG site. With the help of molecular dynamics (MD), applying restraints based on the site-directed mutagenesis results, we now propose a refined NAG site structure where a conformational change in the $\beta 4$ - $\alpha 4$ loop could be the initial signal in NAG activation.

RESULTS

Clinical ASD domain mutations.

The twelve reported (Häberle et al., 2011) CPS1D-associated missense mutations mapping in the ASD are listed in Table 1. All of them affect residues that are invariant or highly conserved in NAG-sensitive CPSs. Except for two mutations (R1371L and Y1491H), they were given unanimous predictions of being likely to have negative effects by two widely used pathogenicity prediction servers, Polyphen-2 and MutPred (Li et al., 2009; Adzhubei et al., 2010) (Table 1).

The clinical mutations highlighted in bold face in Table 1 were studied experimentally here. The study began with the expression and purification of each mutant form (Fig. 2A). For all purified mutants, we determined the activity in a standard assay, the thermal stability (Fig. 2B), and the kinetic parameters for ATP, ammonia, bicarbonate and NAG (see Materials and Methods). All the velocities (Fig. 2C and D) refer to one mg of the purified protein. All the mutants exhibited hyperbolic kinetics for the substrates and for NAG. Unless indicated, the kinetic parameters for the substrates were similar to those of wild-type CPS1 (Díez-Fernández et al., 2013). The kinetic

parameters for NAG, relative to those for the pure recombinant wild-type enzyme (Díez-Fernández et al., 2013), are shown in Fig. 2D and E.

Impact of the clinical mutations on enzyme stability.

Unlike the case for clinical mutations mapping in the integrating domain (Díez-Fernández et al., 2014), which drastically reduced CPS1 production in the present expression system, none of the five ASD clinical mutations tested here, (R1371L, T1391M, L1398V, P1439L and P1462R; in bold type in Table 1) decreased importantly CPS1 production or purification (Fig. 2A). Only for the L1398V mutation there was some reduction (about 50%) of enzyme production and, correspondingly, of the purity of the final enzyme preparation (Fig. 2A). Nevertheless, the enzyme form carrying this mutation exhibited a decrease of about 5°C in its thermal stability (Fig. 2B), relative to wild-type CPS1, suggesting that L1398V might cause deficiency by speeding CPS1 inactivation. Another two ASD mutations, A1378T and L1381S, were previously found to cause, respectively, thermal destabilization and abolition of CPS1 expression because of degradation (Díez-Fernández et al., 2013) (Table 1). Interestingly, the A1378T, L1381S and L1398V mutations map in the very crowded and highly hydrophobic patch between the central β sheet and the α_3 layer of the domain (Fig. 1D) [the ASD is folded as an $\alpha_3\beta_5\alpha_2$ sandwich (Fig. 1C) where strands and helices alternate, beginning with $\beta 1$].

Effects of clinical ASD mutations on CPS1 activity and the kinetic parameters for NAG.

The R1371L, T1391M, L1398V and P1462R mutations decreased $\geq 80\%$ the specific activity (the activity per mg of pure CPS1) of the enzyme, determined using a

substrate-rich and NAG-rich enzyme activity assay (Fig. 2C). Of these mutations, P1462R caused nearly complete (about 97%) inactivation. Previously, the clinical mutations R1453W and R1453Q were found to inactivate the otherwise apparently well folded mutant CPS1 (Table 1), whereas T1443A and Y1491H decreased by about 20-fold and about 4-fold, respectively, the specific activity of the enzyme (Pekkala et al., 2010; Díez-Fernández et al., 2013). Thus, eight out of the twelve reported ASD clinical mutations (Häberle et al., 2011) significantly decrease enzyme specific activity. Of the patients carrying these activity-decreasing mutations, those for which the clinical condition was known presented severe deficiency (Table 1).

The study of the NAG activation kinetics of the purified mutant proteins (Fig. 2D and E) revealed that the R1371L and T1391M mutations increase by two orders of magnitude the NAG concentration required for half-maximal activation of CPS1 (K_a^{NAG}). The P1439L mutation also increased K_a^{NAG} by about 15-fold. These important decreases in the affinity of the enzyme for NAG should account for the clinical deficiency, given the NAG levels prevailing in the liver (Tuchman and Holzkecht, 1990). The same should be the case for the T1443A and Y1491H clinical mutations, found earlier (Pekkala et al., 2010; Díez-Fernández et al. 2013) to increase K_a^{NAG} by about 160-fold and about 50-fold, respectively (Table 1). Overall, in five out of the twelve ASD clinical mutations the K_a^{NAG} was significantly increased.

The large decrease in the observed enzyme activity caused by the P1462R mutation was largely due to decreased velocity at saturation of NAG ($V^{[NAG] = \infty}$; Fig. 2D), whereas the K_a^{NAG} remained essentially normal (Fig. 2E). Substrate kinetics for this mutant was normal for all substrates (not shown) except for ATP. Thus, the apparent K_m^{ATP} was increased about 10-fold (Fig. 2F), which should not substantially decrease enzyme activity in our ATP-rich standard assay, but may contribute to the

deficiency *in vivo*. Given the lack of substrate-binding or catalytic machinery in the ASD, the decreased $V^{[NAG]} = \infty$ with normal affinity for NAG observed for the P1462R mutant indicates that bound NAG elicits a poorer activation in this mutant than in wild-type CPS1. Thus, the transmission of the NAG signal to the catalytic domains appears hampered. This may also be the reason for the inactivation caused by the ASD clinical mutations R1453W and R1453Q (Pekkala et al., 2009, 2010).

Effects of ASD mutations on signal transmission and NAG binding

We got insight on the mechanism of signal transmission from the ASD to the catalytic domains by introducing the non-clinical (that is, not found in patients) D1322L mutation (Table 2; this table lists non-clinical mutations, with those studied experimentally here highlighted in bold-type). D1322 is a carbamate phosphorylation domain residue that is ion-paired to R1453 (Fig. 1C), judged from the structural model of the complete CPS1 molecule (Martínez et al., 2010) [this model is based on the structure of *Escherichia coli* CPS, the only CPS molecule that has been structurally characterized in its entirety (Thoden et al., 1999); *E.coli* CPS is not activated by NAG and is active in the absence of effectors (Meister, 1989)]. If, as proposed above, R1453Q and R1453W inactivate CPS1 by abolishing transmission of the NAG signal to the catalytic domains, the D1322L mutation should also be inactivating, which is indeed consistent with the observations made (Table 2). Although the D1322L mutant was soluble and was expressed as abundantly as the wild-type enzyme and purified similarly (Fig. 3A), its activity was <1% of that of wild-type CPS1 (Fig. 3B).

We also examined the correctness of the previously proposed NAG site (Fig. 1C) (Pekkala et al., 2009) in the light of the effects of clinical mutations studied here or reported recently (Pekkala et al., 2010; Díez-Fernández et al. 2013), and by introducing

four additional non-clinical mutations designed to test specific traits of the site (W1410A, K1444A, F1445A and N1449A; Table 2 in bold-type and Fig. 3). The large K_a^{NAG} effects of the reported N1437D, N1440D, and T1394A mutations as well as of the presently studied R1371L and T1391M clinical mutations (Fig. 2E) were consistent with the previously predicted interactions of R1371, T1391, T1394, N1437 and N1440 with NAG (Pekkala et al., 2009) (Fig. 4A). The about 15-fold increase in the K_a^{NAG} triggered by the P1439L clinical mutation (Fig. 2E) can be accounted by changes in the conformation of the β 4- α 4 loop (composed of residues 1438-1445), since this loop includes P1439 and is an important part of the proposed NAG site (Fig. 1C and 4A). The localization of the NAG site was further supported by the results obtained here with the N1449A mutant (Fig. 3). This mutant exhibited about 10-fold increase in the K_a^{NAG} (Fig. 3D), compatible with a predicted hydrogen bond between the N1449 side-chain and the NAG α -carboxylate (Fig. 1C). However, surprisingly, the W1410A, K1444A and F1445A mutations (Fig. 3), which alter residues that would form a lid in the putative NAG site of our previous model (Fig. 1C), did not trigger important changes in specific activity or in NAG activation kinetics (Fig. 3B–D and Table 2). Furthermore, the T1443A clinical mutation was reported to increase K_a^{NAG} by about 160-fold (Table 2) (Díez-Fernández et al., 2013), suggesting a strong interaction between the side-chain of T1443 and the bound NAG molecule. Such interaction was absent in our previous model for NAG binding (Pekkala et al., 2009) (Fig. 1C), which is based on the experimental crystal structure of the NAG-free ASD (Xie et al., 2007). In this structure, the side chain of T1443 is far from the NAG site (Figs. 1C and 4A). Altogether, these observations suggest that in the NAG-bound form of the ASD, the final part of the β 4- α 4 loop changes its conformation relative to the NAG-free form (Xie et al., 2007), placing the side chain of T1443 near NAG and altering the location of the neighbouring

W1410, K1444 and F1445 residues, which would fail to interact with the ligand (Fig. 4B).

Modeling of the NAG site

To build a refined model of the NAG site of CPS1 that was consistent with the mutagenesis analyses, we used restrained MD simulations of the CPS-NAG complex. The restraints were based on the K_a^{NAG} variations induced by the ASD mutations (Tables 1 and 2) as well as on protein-NAG contacts deduced from previous analyses (Pekkala et al., 2009) (see Materials and Methods for the restraints applied and for details on the approach used). The MD simulations used as a starting point the experimental ASD structure (Xie et al., 2007) together with the coordinates of the NAG molecule previously obtained by unrestrained docking (Pekkala et al., 2009). In the energy-minimized model resulting from the MD simulations, the position of the NAG molecule was very similar to that proposed previously (Fig. 4A and B). The γ -COO⁻ group of NAG interacts with the side chains of R1371 and T1394 in the more exposed region of the binding site. The NAG N atom forms a hydrogen bond with T1391, and one oxygen of the α -COO⁻ group is hydrogen-bonded by the side-chains of N1440 and N1449. At the floor of the site, the acetamido methyl group of NAG is surrounded by the L1363, I1423 and I1452 hydrophobic residues (Fig. 4C), whereas the acetamido carbonyl establishes a hydrogen-bonding interaction with N1437 (Fig. 4A). The most important variation with respect to the previous model was a change in the conformation of the β 4- α 4 loop (Fig. 4A-D), brought about by a NAG-T1443 hydrogen-bonding MD restraint based on the effect of the T1443A mutation (Table 1). As a result, the side chain of T1443 is turned inwards and forms a hydrogen bond with the other α -COO⁻ oxygen of NAG, whereas the side chains of the neighbouring K1444 and F1445

residues, which blocked entry to the empty binding site in the crystallographic structure of the isolated domain, are oriented outwards (Fig. 4D). These changes agree with the limited effect on NAG activation of the K1444A and F1445A mutations (Table 2).

To evaluate the new model of the ASD-NAG complex, we removed the NAG molecule and performed unrestrained docking calculations with GOLD (Verdonk et al., 2003). These calculations yielded total convergence (100%) of the NAG docked poses (20 of 20 solutions with root mean square conformational deviations, $RSMD \leq 0.3 \text{ \AA}$) (Fig. 4C). These poses generated by docking were very close to the energy-minimized conformation obtained by restrained MD (average $RSMD 0.57 \text{ \AA}$). Furthermore, the docking scores were 1.39 times superior (63.2 ± 0.5 vs. 45.5 ± 2.4) to those obtained for our previous model with the same methodology and scoring function (Pekkala et al., 2009).

DISCUSSION

The present and earlier experimental results (Pekkala et al., 2010; Díez-Fernández et al., 2013, 2014) corroborate the value of the baculovirus/insect cell system used here for assessing the effects of CPS1D missense mutations. In the case of the ASD, only for one of the twelve reported clinical mutations (Häberle et al., 2011), P1411L, the expression studies could not ascertain the disease causality of the mutation (Table 1) (Pekkala et al., 2010). Three ASD clinical mutations destabilized CPS1 and eight hampered or abolished enzyme activity (Table 1). Therefore, decreased specific activity is a frequent consequence of ASD clinical mutations. In line with the presence of the NAG site in the ASD, a reduced affinity for NAG is also a common consequence of these mutations (five of the twelve clinical ASD mutations). The three clinical mutations that inactivated the enzyme or that decreased its activity without increasing

K_a^{NAG} may do so by blocking or hampering transmission of the NAG activating signal. NAG activation of CPS is both a V and a K allosteric process in which V_{max} is increased and the K_m for ATP is decreased (Rubio et al., 1983). This possibly explains the increase in K_m^{ATP} associated with the inadequate NAG activation proposed for the near-inactivating P1462R mutation.

In addition to being valuable for making genotype-phenotype correlations, the present expression/mutagenesis system could help guide therapy. Patients carrying "kinetic" mutations causing increases in K_a^{NAG} could benefit from N-carbamyl-L-glutamate (NCG) administration. This deacylase-resistant NAG analog (Rubio and Grisolia, 1981) and registered drug could artificially help saturate the NAG site of these mutants. Furthermore, NCG stabilizes CPS1 (Díez-Fernández et al., 2013), opening the way to testing whether NCG might be beneficial for patients with mutations causing CPS1 destabilization. These mutations can also be identified with the present system (Díez-Fernández et al., 2014), although in the case of the ASD only few clinical mutations cause substantial destabilization. Nevertheless, these mutations reveal that the ASD contributes to CPS1 stability. Such contribution has already been detected for the corresponding domain of *E. coli* CPS when the deletion of this domain was found to lower the denaturation temperature of the enzyme (Cervera et al., 1993). The clustering of clinical destabilizing mutations in the hydrophobic patch between the central β sheet and the α_3 layer of the ASD (Fig. 1D) suggests a particularly important contribution of this region to proper ASD and CPS1 folding.

Our analysis provides some interesting information on CPS1 activation by NAG. In particular, a refined model has been built that is consistent with the effects on K_a^{NAG} of all available mutations. The more salient difference between this model and the structure of the NAG-free crystal structure of the ASD is a conformational change in the

final part of the $\beta 4$ - $\alpha 4$ loop to accommodate NAG (Fig. 4A–D). As a consequence of this movement, the side-chains of F1445 and T1443 exhibit large displacements (Fig. 4C and D). In the NAG-free crystal form, these side-chains point, respectively, towards the NAG site and away from it whereas in the NAG-bound form they adopt the reverse orientations (Fig. 4D). The interaction of T1443 and the lack of interaction of F1445 with NAG would explain the respective effect and absence of effect on K_a^{NAG} of the corresponding alanine mutations. This movement could importantly alter the relations of the NAG site with the adjacent carbamate phosphorylation domain (Figs. 1C and 4E), acting as an on/off switch for this catalytic domain, thus possibly being the first event in the NAG activation process.

Contacts between the ASD and the carbamate phosphorylation domain mediated by helix $\alpha 4$ and strand $\beta 5$ are observed in *E. coli* CPS (Thoden et al., 1999) and, correspondingly, in the structural model for the entire CPS1 (Martínez et al., 2010) (Figs. 1C and 4E). The present results suggest that these contacts may be essential for stabilizing the NAG-activated conformation of the carbamate phosphorylation domain. Since P1462 is immediately upstream of $\beta 5$ (Fig. 4E), the P1462R mutation may cause near-inactivation by altering the position of strand $\beta 5$ relative to the carbamate phosphorylation domain. R1453 belongs to helix $\alpha 4$ (Fig. 4E), and the mutations at R1453, as well as the mutation of D1322, the ion-pair partner of R1453 across the interdomain divide (Fig. 1C), may cause inactivation because the abolition of the ion pair would also disturb the ASD-carbamate phosphorylation domain interaction. Actually, D1322 belongs to a loop that is sandwiched between the ASD and another loop (residues 778–787) from the bicarbonate phosphorylation domain (Fig. 1C). Therefore, the NAG activating signal might also propagate *via* this conduit to the bicarbonate phosphorylation domain. In any case, further progress towards clarification

of the path of the activating signal from the ASD to both catalytic domains may require the determination of CPS1 structures in NAG-free and NAG-bound forms.

MATERIALS AND METHODS

Human CPS1 production

Pure recombinant human mature liver CPS1 with an N-terminal His₆-tag, and the desired site-directed mutants of this enzyme were produced from the pFastBac-CPS1 vector as previously described (Díez-Fernández et al., 2013). The same purification procedure (including cell centrifugation, lysis, centrifugal clarification, Ni-affinity chromatography and centrifugal ultrafiltrative concentration) proved appropriate for wild type and mutant enzyme forms. Purity was monitored by SDS-PAGE (8% polyacrylamide gels) (Laemmli, 1978) and Coomassie staining. Protein was determined according to Bradford's method (Bradford, 1976).

CPS1 activity assays

Activity was determined at 37°C by monitoring carbamoyl phosphate production as citrulline (Nuzum and Snodgrass, 1976) in a 10-min ornithine transcarbamylase-coupled assay (Díez-Fernández et al., 2013). The standard assay mixture contained 50 mmol/L glycyl-glycine pH 7.4, 70 mmol/L KCl, 1 mmol/L dithiothreitol, 20 mmol/L MgSO₄, 5 mmol/L ATP, 5mmol/L MgCl₂, 35 mmol/L NH₄Cl, 50 mmol/L KHCO₃, 10 mmol/L NAG, 5 mmol/L L-ornithine and 4 U/mL ornithine transcarbamylase. One enzyme unit makes 10 µmol citrulline in the 10-min assay. Substrate kinetics and NAG activation kinetics were studied by varying the concentration of one substrate or of NAG while other assay components were kept fixed at the concentrations used in the standard reaction mixture. ATP was added as an equimolar mixture with MgCl₂.

Therefore, since 20 mmol/L MgSO_4 was present in the reaction mixture, Mg^{2+} was always in 20 mmol/L excess over ATP. Kinetic data were fitted to hyperbolae using the GraphPad Prism program (GraphPad Software, San Diego, CA). Values for activity and for $V^{[NAG] = \infty}$ are referred to 1 mg of protein, thus being specific activities. These values and K_a^{NAG} values are shown as means $\pm$ SE for at least three replicate estimates. To better compare the magnitude of the changes in the activity and the kinetic constants with those reported previously, the values for these constants and for their standard errors have been normalized by dividing them by the corresponding mean values for the same parameters for wild-type CPS1 (tabulated in Díez-Fernández et al., 2013).

To monitor the thermal stability of CPS1 mutants, 0.5 mg/mL of the indicated enzyme form in 50 mmol/L glycyl-glycine pH 7.4, 10% glycerol, 0.5 mol/L NaCl, 30 mmol/L imidazole and 2 mmol/L dithiothreitol were incubated for 15 min at the indicated temperature. Then, the mixtures were rapidly cooled at 0°C and enzyme activity was determined immediately in the standard assay at 37°C. Data are expressed as means $\pm$ SE for at least duplicate assays.

Restrained molecular dynamics and docking calculations

We first used MD simulations with restraints based on the K_a^{NAG} variations induced by CPS mutations (Tables 1 and 2) and on the contacts obtained in previous unrestrained docking runs (Pekkala et al., 2009). The calculations used as a starting point the PDB structure of the human CPS C-terminal domain (PDB file 2YVQ) (Xie et al., 2007) together with the coordinates of the NAG molecule previously obtained by docking and energy minimization (Pekkala et al., 2009). The $^{1415}\text{GQNPS}^{1419}$ sequence missing in the crystallographic structure was modelled with MOE (CCG Inc., Canada) using a method based on PDB searches. The potential energy of the initial system was progressively

minimized before subjecting the CPS-NAG complex to two MD runs of 200 ps and 100 ps, respectively, at 300 K. During these runs, a soft positional restraint was applied to all protein atoms except those of $\beta 1\text{-}\alpha 1$, $\beta 2\text{-}\alpha 2$, $\beta 3$ and $\beta 4\text{-}\alpha 4$ residues close to the NAG binding site. The temperature of the system was slowly reduced to 0 K during the second half of each run, and NAG soft positional restraints were applied in the first simulation but eliminated in the second run. The following five hydrogen-bonding restraints were applied during the simulations, based on mutant K_a^{NAG} increments equal or greater than 50-fold in Tables 1 and 2 and the corresponding hydrogen-bonding contacts observed in unrestrained docking analyses (Pekkala et al., 2009): $\text{O}_{\gamma\text{-COO}}^{\text{NAG}}\text{-HN}_{\eta}^{\text{R1371}}$, $\text{HN}^{\text{NAG}}\text{-O}_{\gamma}^{\text{T1391}}$, $\text{O}_{\gamma\text{-COO}}^{\text{NAG}}\text{-HO}_{\gamma}^{\text{T1394}}$, $\text{O}_{\text{Ac}}^{\text{NAG}}\text{-HN}_{\delta}^{\text{N1437}}$ and $\text{O}_{\alpha\text{-COO}}^{\text{NAG}}\text{-HN}_{\delta}^{\text{N1440}}$. An additional $\text{O}_{\alpha\text{-COO}}^{\text{NAG}}\text{-HO}_{\gamma}^{\text{T1443}}$ restraint was imposed on the remaining $\text{O}_{\alpha\text{-COO}}$ atom of NAG on the basis of the 160-fold increase in K_a^{NAG} detected for the T1443A mutant (Table 1). Throughout these calculations, we used the ff10 force field of AMBER 8.0 (Case et al., 2005), and a generalized Born model for simulating an aqueous environment. The final coordinates of the human CPS-NAG complex were generated by minimizing the potential energy of the last MD snapshot. This model was evaluated by means of unrestrained docking calculations. These were carried out using the GOLD package (version 5.2) (Verdonk et al., 2003) with the GoldScore fitness function, as in the previous calculations (Pekkala et al., 2009).

Severity assessment

The severity of the clinical mutations was inferred from the clinical presentation, being graded as severe for neonatal presentations and/or death, and mild for late presentations or prolonged survival. The PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>) (Adzhubei et al., 2010) and MutPred (<http://mutpred.mutdb.org/>) (Li et al., 2009)

servers were used to assess *in silico* the disease-causing potential of the clinical mutations. PolyPhen-2 grades the probability of a damaging effect of an amino-acid substitution, as (from high to low) Probably damaging, Possibly damaging, and Benign. MutPred gives a g score corresponding to the probability that a given amino acid substitution was deleterious/disease-associated.

Other techniques

Structural analysis, superimpositions, and representations have been performed with Pymol (DeLano Scientific; <http://www.pymol.org>) using the experimental crystal structure of the ASD (PDB file 2YVQ) (Xie et al., 2007) or the *in silico* models for the NAG-bound form of the ASD (Pekkala et al., 2009 and the present work) and for the entire CPS1 molecule (Martínez et al., 2010). Amino acid conservation was determined by ClustalW sequence alignment (Larkin et al., 2007) of either CPS1, CPSIII or other CPSs from 14, 6 and 26 species, respectively.

Acknowledgements

This work was supported by grants from the Alicia Koplowitz Foundation, the Valencian government (No. PrometeoII/2014/029 to V.R.), the Spanish government (Nos. BFU2011-30407 to V.R., BFU2012-30770 to J.G. and a FPU fellowship to C.D.-F.), and the Swiss National Science Foundation (No. 310030_127184 to J.H.). We thank Belén Barcelona (IBV-CSIC, Valencia, Spain) for the structural model of the complete enzyme and for help with Fig. 4D. The original mutation analysis of some CPS1D patients was kindly supported by Orphan Europe SARL (France).

References

- Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248-249.
- Alonso, E., Cervera, J., Garcia-Espana, A., Bendala, E., Rubio, V., 1992. Oxidative inactivation of carbamoyl phosphate synthetase (ammonia). Mechanism and sites of oxidation, degradation of the oxidized enzyme, and inactivation by glycerol, EDTA, and thiol protecting agents. *J. Biol. Chem.* 267, 4524–4532.
- Alonso, E., Rubio, V., 1995. Affinity cleavage of carbamoyl-phosphate synthetase I localizes regions of the enzyme interacting with the molecule of ATP that phosphorylates carbamate. *Eur. J. Biochem.* 229, 377–384.
- Bender, D.A., 2012. Amino acid metabolism. 3rd edition. Wiley-Blackwell, Oxford.
- Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254.
- Brusilow, S.W., Horwich, A.L., 2001. Urea cycle enzymes. In: Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., Child, B., Kinzler, K.W., Vogelstein, B. (Eds.), *The Metabolic and Molecular Bases of Inherited Disease*, 8th edition. McGraw-Hill, New York, USA, pp. 1909-1963.
- Case, D.A., Cheatham T.E. III, Darden, T., Gohlke, H., Luo, R., Merz, K.M. Jr., Onufriev, A., Simmerling, C., Wang, B., Woods, R., 2005. The Amber biomolecular simulation programs. *J. Computat. Chem.* 26, 1668-1688
- Cervera, J., Conejero-Lara, F., Ruiz-Sanz, J., Galisteo, M.L., Mateo, P.L., Lusty, C.J., Rubio, V., 1993. The influence of effectors and subunit interactions on *Escherichia*

- coli* carbamoyl-phosphate synthetase studied by differential scanning calorimetry. J. Biol. Chem. 268, 12504-12511.
- Díez-Fernández, C., Hu, L., Cervera, J., Häberle, J., Rubio, V., 2014. Understanding carbamoyl phosphate synthetase (CPS1) deficiency by using the recombinantly purified human enzyme: effects of CPS1 mutations that concentrate in a central domain of unknown function. Mol. Genet. Metab. 112, 123-132.
- Díez-Fernández, C., Martínez, A.I., Pekkala, S., Barcelona, B., Pérez-Arellano, I., Guadalajara, A.M., Summar, M., Cervera, J., Rubio, V., 2013. Molecular characterization of carbamoyl-phosphate synthetase (CPS1) deficiency using human recombinant CPS1 as a key tool. Hum. Mutat. 34, 1149-1159.
- Eeds, A.M., Hall, L.D., Yadav, M., Willis, A., Summar, S., Putnam, A., Barr, F., Summar, M.L., 2006. The frequent observation of evidence for nonsense-mediated decay in RNA from patients with carbamyl phosphate synthetase I deficiency. Mol. Genet. Metab. 89, 80-86.
- Freeman, J.M., Nicholson, J.F., Masland, W.S., Rowland, L.P., Carter, S., Curnen, E.C., 1964. Ammonia intoxication due to a defect in urea synthesis. J. Pediat. 65, 1039-1040.
- Funghini, S., Donati, M.A., Pasquini, E., Zammarchi, E., Morrone, A., 2003. Structural organization of the human carbamyl phosphate synthetase I gene (*CPS1*) and identification of two novel genetic lesions. Hum. Mutat. 22, 340-341.
- Gelehrter, T.D., Snodgrass, P.J., 1974. Lethal neonatal deficiency of carbamyl phosphate synthetase. N. Engl. J. Med. 290, 430-433.
- Häberle, J., Rubio, V., 2014. Hyperammonemia and related disorders. In: Blau, N., Duran, R., Gibson, K.M., Blaskovics, M., Dionisi-Vici, C., (Eds.), Physician's

- Guide to the Diagnosis, Treatment and Follow-up of Inherited Metabolic Diseases. Springer-Verlag, Heidelberg, Germany, pp. 67-72.
- Häberle, J., Schmidt, E., Pauli, S., Rapp, B., Christensen, E., Wermuth, B., Koch, H.G., 2003. Gene structure of human carbamylphosphate synthetase 1 and novel mutations in patients with neonatal onset. *Hum. Mutat*, 21, 444.
- Häberle, J., Shchelochkov, O.A., Wand, J., Katsonis, P., Hall, L., Reiss, S., Eeds, A., Willis, A., Yadav, M., Summar, S., Consortium TUCD, Lichtarge, O., Rubio, V., Wong, L.J., Summar, M., 2011. Molecular defects in human carbamoyl phosphate synthetase I: Mutational spectrum, diagnostic and protein structure considerations. *Hum. Mutat.* 32, 579-589.
- Haraguchi, Y., Uchino, T., Takiguchi, M., Endo, F., Mori, M., Matsuda, I., 1991. Cloning and sequence of a cDNA encoding human carbamyl phosphate synthetase I: molecular analysis of hyperammonemia. *Gene* 107, 335-340.
- Hong, J., Salo, W.L., Lusty, C.J., Anderson, P.M., 1994. Carbamyl phosphate synthetase III, an evolutionary intermediate in the transition between glutamine-dependent and ammonia-dependent carbamyl phosphate synthetases. *J. Mol. Biol.* 243, 131-140.
- Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774-797.
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948.

- Li, B., Krishnan, V.G., Mort, M.E., Xin, F., Kamati, K.K., Cooper, D.N., Mooney, S.D., Radivojac, P., 2009. Automated interference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 25, 2744-2750.
- Martínez, A.I., Pérez-Arellano, I., Pekkala, S., Barcelona, B., Cervera, J., 2010. Genetic, structural and biochemical basis of carbamoyl phosphate synthetase 1 deficiency. *Mol. Genet. Metab.* 101, 311-323.
- McReynolds, J.W., Crowley, B., Mahoney, M.J., Rosenberg, L.E., 1981. Autosomal recessive inheritance of human mitochondrial carbamyl phosphate synthetase deficiency. *Am. J. Hum. Genet.* 33, 345-353.
- Meister, A., 1989. Mechanism and regulation of the glutamine-dependent carbamyl phosphate synthetase of *Escherichia coli*. *Adv. Enzymol. Relat. Areas Mol. Biol.* 62, 315-374.
- Morita, T., Mori, M., Tatibana, M., 1982. Regulation of N-acetyl-L-glutamate degradation in mammalian liver. *J. Biochem.* 91, 563-569.
- Nuzum, C.T., Snodgrass, P.J., 1976. Multiple assays of the five urea cycle enzymes in human liver homogenates. In: Grisolia, S., Báguena, R., Mayor, F., (Eds.), *The Urea Cycle*. John Wiley and Sons, New York, USA, pp. 325-349
- Nyunoya, H., Broglie, K.E., Widgren, E.E., Lusty, C.J., 1985. Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. *J. Biol. Chem.* 260, 9346-9356.
- Pekkala, S., Martínez, A.I., Barcelona, B., Gallego, J., Bendala, E., Yefimenko, I., Rubio, V., Cervera, J., 2009. Structural insight on the control of urea synthesis: identification of the binding site for N-acetyl-L-glutamate, the essential allosteric activator of mitochondrial carbamoyl phosphate synthetase. *Biochem. J.* 424, 211-220.

- Pekkala, S., Martínez, A.I., Barcelona, B., Yefimenko, I., Finckh, U., Rubio, V., Cervera, J., 2010. Understanding carbamoyl-phosphate synthetase 1 (CPS1) deficiency by using expression studies and structure-based analysis. *Hum. Mutat.* 31, 801-808.
- Rodriguez-Aparicio, L.B., Guadalajara, A.M., Rubio, V., 1989. Physical location of the site for N-acetyl-L-glutamate, the allosteric activator of carbamoyl phosphate synthetase, in the 20-kilodalton COOH-terminal domain. *Biochemistry* 28, 3070-3074.
- Rubio, V., Britton, H.G., Grisolia, S., 1983. Mitochondrial carbamoyl phosphate synthetase activity in the absence of N-acetyl-L-glutamate. Mechanism of activation by this cofactor. *Eur. J. Biochem.* 134, 337-343.
- Rubio, V., Grisolia, S., 1981. Treating urea cycle defects. *Nature* 292, 496.
- Rubio, V., Ramponi, G., Grisolia, S., 1981. Carbamoyl phosphate synthetase I of human liver. Purification, some properties and immunological cross-reactivity with the rat liver enzyme. *Biochim. Biophys. Acta.* 659, 150-160.
- Shigesada, K., Aoyagi, K., Tatibana, M., 1978. Role of acetylglutamate in ureotelism. Variations in acetylglutamate level and its possible significance in control of urea synthesis in mammalian liver. *Eur. J. Biochem.* 85, 385-391.
- Sonoda, T., Tatibana, M., 1983. Purification of N-acetyl-L-glutamate synthetase from rat liver mitochondria and substrate and activator specificity of the enzyme. *J. Biol. Chem.* 258, 9839-9844.
- Stewart, P.M., Walser, M., 1980. Short term regulation of ureagenesis. *J. Biol. Chem.* 255, 5270-5280.
- Summar, M.L., 1998. Molecular genetic research into carbamoyl-phosphate synthase I: molecular defects and linkage markers. *J. Inherit. Metab. Dis.* 21 (Suppl 1), 30-39.

- Summar, M.L., Hall, L.D., Eeds, A.M., Hutcheson, H.B., Kuo, A.N., Willis, A.S., Rubio, V., Arvin, M.K., Schofield, J.P., Dawson, E.P., 2003. Characterization of genomic structure and polymorphisms in the human carbamyl phosphate synthetase I gene. *Gene* 311, 51-57.
- Summar, M.L., Koelker, S., Freedenberg, D., Le Mons, C., Häberle, J., Lee, H.S., Kirmse, B., The European Registry and Network for Intoxication Type Metabolic Diseases (E-IMD), The Members of the Urea Cycle Disorders Consortium (UCDC), 2013. The incidence of urea cycle disorders. *Mol. Genet. Metab.* 110, 179-180.
- Thoden, J.B., Raushel, F.M., Benning, M.M., Rayment, I., Holden, H.M., 1999. The structure of carbamoyl phosphate synthetase determined to 2.1 Å resolution. *Acta Crystallogr. D. Biol. Crystallogr.* 55, 8-24.
- Tuchman, M., Holzknecht, R.A., 1990. N-acetylglutamate content in liver and gut of normal and fasted mice, normal human livers, and livers of individuals with carbamyl phosphate synthetase or ornithine transcarbamylase deficiency. *Pediatr. Res.* 27, 408-412.
- Uchino, T., Endo, F., Matsuda, I., 1998. Neurodevelopmental outcome of long-term therapy of urea cycle disorders in Japan. *J. Inherit. Metab. Dis.* 21 (Suppl. 1), 151-159.
- Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W., Taylor, R.D., 2003. Improved protein-ligand docking using GOLD. *Proteins* 52, 609-623.
- Xie, Y., Ihsanawati, K., Kishishita, S., Murayama, K., Takemoto, C., Shirozu, M., Yokoyama, S., 2007. Crystal structure of MGS domain of carbamoyl-phosphate synthetase from homo sapiens. Protein Data Bank DOI, [10.2210/pdb2yvq/pdb](https://doi.org/10.2210/pdb2yvq/pdb)

Legends to figures

Fig. 1. The CPS1-NAG switch for urea cycle control.

A: Simplified view of the urea cycle to highlight its control at the level of CPS1. The double arrows denote bidirectional fluxes between the elements that are linked. Enzymes are boxed and abbreviated as follows: NAGS, NAG synthase; OTC, ornithine transcarbamylase; ASS, argininosuccinate synthetase; ASL, argininosuccinate lyase; ARG1, arginase 1. ORNT 1, ornithine/citrulline antiporter. For simplicity, not all products of the different reactions are indicated. The intramitochondrial part of the cycle, where the switch mechanism operates, is highlighted. **B:** Scheme of the mature CPS1 polypeptide (N-terminal mitochondrial targeting sequence removed), indicating its two moieties (top) that are homologous to the small and large subunits of *E. coli* CPS, the two halves of the large moiety (middle), and the domain composition (lower bar) with domain names above, domain boundaries given as residue numbers, and domain functions shown below (“??” means unknown function), including the domain localization of the three steps of the CPS1 reaction. The blue arrow indicates carbamate migration between both phosphorylation domains, a process that is unlikely to involve the integrating domain (Thoden et al., 1999). Vertical lines towards the C-end map CPS1D missense mutations and rationally-designed mutations analyzed here (listed in Tables 1 and 2). The longer lines indicate that two different mutations affect the same residue. **C:** Stereo view pair (two views distinguished by a 5° rotation around the vertical axis, to allow 3-D imaging with a stereoscopic viewer) of the crystallographic structure of the allosteric domain with NAG bound as previously modeled (Pekkala et al., 2009). Spheres and explicitly shown amino acid side-chains mark residues hosting missense mutations that are discussed here and identified by labeling (residues underlined are those hosting mutations studied experimentally here; the asterisks mark

residues hosting CPS1D-associated mutations). They are colored red if the mutation decreases stability, blue if it increases the K_a^{NAG} , purple if it causes inactivation or strong k_{cat} reduction, and orange for little or no effect. The R1453:D1322 ion pair is illustrated with a dotted line. Some ASD secondary structure elements (including the $\beta 4$ - $\alpha 4$ loop) are labeled. The C^α trace is shown in black and NAG in sticks and colored (C, N and O atoms, yellow, blue and red, respectively). The indicated loops belonging to both phosphorylation domains from the superimposed structural model of human CPS1 (Martínez et al., 2010) are shown colored, with the side-chain of D1322 in sticks.

D: Hydrophobic nucleus hosting the three CPS1D-associated ASD mutations that destabilized CPS1. The residues forming this nucleus are labeled and their hydrocarbon side chains are shown in sticks representation and colored grey, except for the three residues that host the destabilizing mutations, which are labeled with larger font in red, while their side-chains are colored yellow. Secondary structure elements are shown in cartoon representation and are also labeled. Figure is prepared from PDB file 2YVQ.

Fig. 2. Production, stability, activity and kinetic properties of wild-type and CPS1D-associated mutant CPS1 forms.

A: SDS-PAGE (8% polyacrylamide, Coomassie blue staining) of purified human recombinant CPS1, either wild-type (WT) or carrying the indicated mutations. St, protein markers (PageRuler prestained protein, Thermo Scientific, USA), with masses indicated in kDa. The arrowhead signals the position of the CPS1 band. **B:** Inactivation upon 15-min heating at the indicated temperatures for recombinant human CPS1, either wild-type or carrying the indicated mutations. For clarity, a single line was fitted to the R1371L, T1391M, P1439L and P1462R mutants because they were highly similar in terms of stability. The horizontal dashed line marks 50% inactivation, whereas the

vertical dashed lines mark the temperature at which 50% inactivation occurs for the corresponding enzyme form. **C–F**: Effects of CPS1D-associated ASD mutations on enzyme activity (**C**), $V^{[NAG] = \infty}$ (**D**), K_a^{NAG} (**E**), and ATP substrate kinetics (**F**). The activities and $V^{[NAG] = \infty}$ values were always referred to 1 mg of protein. Results are given as fractions of the corresponding mean values for wild-type CPS1 (Fernández-Díez et al., 2013). The hyperbolas shown in (**F**) for the wild-type and the P1462R mutant forms of CPS1 are those for K_m^{ATP} values of 0.61 mmol/L and 6.2 mmol/L, respectively, and for relative velocities at infinite concentration of ATP of 1 and 0.118, respectively.

Fig. 3. Production, activity and kinetic properties of ASD mutant forms not identified in CPS1D.

A: SDS-PAGE (8% polyacrylamide, Coomassie blue staining) of the wild type enzyme (WT) and the indicated mutants. St, protein markers (PageRuler prestained protein ladder), with masses indicated in kDa. The arrowheads signal the position of the CPS1 band. **B–D**: Effects of the ASD mutations tested on enzyme activity (**B**), $V^{[NAG] = \infty}$ (**C**), and K_a^{NAG} (**D**). Other details are as given in Fig. 2.

Fig. 4. The NAG site, proposed NAG-triggered changes, and the ASD surface involved in the interactions with the carbamate phosphorylation domain.

NAG binding to CPS1 as previously proposed (**A**) and as inferred here (**B**) on the basis of the analysis of all the existing site-directed mutagenesis data and restrained molecular dynamics (MD) simulations. Some relevant residues whose participation in the site was tested by site-directed mutagenesis are shown in sticks representation, with hydrogen bonding to NAG illustrated with broken lines. **C**: Stereo view pair (two views differed

by a 5° rotation around the vertical axis, to allow 3-D imaging with a stereoscopic viewer) of unrestrained docking of NAG to the human CPS1 protein structure obtained by restrained MD simulation. The model is in sticks representation, with N and O atoms in deep blue and red, respectively. The 20 NAG binding poses (sticks with grey-colored carbon atoms) show 100% convergence. The amino acids side chains proposed to surround the NAG molecule are shown with carbon atoms in light blue in the conformation resulting from restrained MD, superimposed on the conformation (carbon atoms in black) observed in the crystal structure of the NAG-free site (PDB file 2YVQ).

D: Stereo view pair (see **C**) of the superimposition of the NAG site in its empty form observed in the crystal structure of the isolated ASD (grey, PDB file 2YVQ) and in the NAG bound form modelled here by restrained MD (yellow). NAG and some amino acid side-chains are shown in sticks representation and are labelled. Some secondary structure elements are shown in cartoon representation and are also labeled. **E:** The ASD (experimental NAG-free form, PDB 2YVQ) as seen from the carbamate phosphorylation domain, with cyan coloring of the regions involved in the interactions with this last domain. Residues involved in the interactions were identified in the CPS1 structure model (Martínez et al., 2010) using the PISA server (Krissinel and Henrick, 2007) at http://www.ebi.ac.uk/pdbe/prot_int/pistart.html. Several secondary structure elements including helix $\alpha 4$, and the residues marked as red dots are labeled. The NAG site is identified by the bound activator.

Figure 1

Figure 1

Figure 2

Figure 2

Figure 3

Figure 3

Figure 4

Figure 4

Table 1

Table 1. CPS1 allosteric domain mutations found in CPS1 deficiency patients, and their effects^a

Amino acid change ^b	Amino acid in CPS ^{c,d}			Patient	Severity ^c		Proposed residue role (new NAG model ^e)	Major effects of recombinant CPS1	Reference
	1	III	Other		Server-predicted				
					PolyPhen-2	MutPred (g-score)			
R1371L	R	R/q	Variable	Unknown^f	Benign	0.51	Ion pair NAG γ-COO⁻ (restraint)	$\uparrow K_a^{NAG}$ ~100-fold	Häberle et al., 2011
A1378T	A	A	Apolar	Severe	Prob. damag.	0.92	α_3/β_5 hydrophobic nest	$\downarrow$ thermal stability ^g	Eeds et al., 2006
L1381S	L	L/f	L/F/v	Severe	Prob. damag.	0.92	α_3/β_5 hydrophobic nest	Degradation ^g	Summar, 1998
T1391M	T	T	Variable	Severe^h	Prob. damag.	0.80	H-bond with NAG NH (restraint)	$\uparrow K_a^{NAG}$ ~200-fold	Häberle et al., 2011
L1398V	L	L	L/Y/f/w/i	Severe^h	Prob. damag.	0.72	α_3/β_5 hydrophobic nest	$\downarrow$ thermal stability	Häberle et al., 2011
P1411L	P	P	Variable	Mild ⁱ	Prob. damag.	0.84	β_3 - α_3 loop, near NAG site	Modest V_{max} effect ^j	Eeds et al., 2006 Summar, 1998
P1439L	P	P	Variable	Severe^h	Prob. damag.	0.66	β_4-α_4 loop conformation	$\uparrow K_a^{NAG}$ ~15-fold	Häberle et al., 2011
T1443A	T	T/S	Variable	Severe	Poss. damag.	0.74	H-bond with NAG α -COO ⁻ (restraint) ^k	$\uparrow K_a^{NAG}$ ~160-fold ^g	Eeds et al., 2006
R1453W	R	R	R/l/i	Severe	Prob. damag.	0.82	NAG signal transmission to the catalytic domains	Inactivation ^j	Häberle et al., 2011 Pekkala et al., 2010
R1453Q	R	R	R/l/i	Severe	Prob. damag.	0.89	NAG signal transmission to the catalytic domains	Inactivation ^j	Pekkala et al., 2010
P1462R	P/A	P	Variable	Unknown^l	Poss. damag.	0.85	Conformation of signal transmission region	$\downarrow V^{NAG}=\infty$ ~30-fold $\uparrow K_M^{ATP}$ ~10-fold	Häberle et al., 2011
Y1491H	Y/F	Y/h	Variable	Unknown ^m	Benign	0.9	Interdomain signal transmission	$\uparrow K_a^{NAG}$ ~50-fold ^j	Summar, 1998

^aAmino acids are shown in single-letter code. Mutations studied experimentally here are highlighted in bold-type. ^bTranslation of the cDNA reference sequence NM_001875.4 (GenBank). Nucleotide 136 in this sequence is considered +1, since it is the A of the translation initiation codon. ^cFor details see Materials and Methods. ^dAmino acids found in low frequency are shown in low case. "Variable" indicates the occurrence at

a given position of >four different amino acids that do not share the same chemical characteristics (polar, apolar, charged, etc.). ^eWhen the contact was used as a restraint for model generation, it is specified. ^fNo clinical information available for the patient carrying this mutation. ^gDíez-Fernández et al., 2013. ^hUnknown when originally reported. New data gathered on the patient. ⁱThe late onset and coexistence with a null second allele (Q478*) indicates residual activity. ^jPekkala et al., 2010. ^kNot present in previous docking model (Pekkala et al., 2009). ^lThe patient is alive, but the second CPS1 missense allele (Y959C) might not be inactivating. ^mThe patient had a late onset presentation, but the second allele was not identified.

Table 2. CPS1 allosteric domain mutations designed to test functional features and their effects^a

Amino acid change	Amino acid in CPS			Proposed role of the residue (new NAG model)	Major effect of recombinant CPS1
	I	III	Other		
D1322L^b	D	D	D/y	NAG signal transmission	Inactivation
T1391V	T	T	Variable	H-bond with NAG NH (restraint)	$\uparrow K_a^{NAG} \sim 300\text{-fold}^c$
T1394A	T	T	T/d/y	H-bond with NAG $\gamma\text{-COO}^-$ (restraint)	$\uparrow K_a^{NAG} \sim 600\text{-fold}^c$
W1410A	W	W	Variable	$\beta 3\text{-}\alpha 3$ loop, near NAG site^d	No substantial effect
W1410K	W	W	Variable	$\beta 3\text{-}\alpha 3$ loop, near NAG site ^d	$\uparrow K_a^{NAG} \sim 40\text{-fold}^c$
N1437D	N	N	Variable	H-bond with NAG O _{Ac} (restraint)	$\uparrow K_a^{NAG} \sim 50\text{-fold}^c$
N1440D	N	N	Variable	H-bond with NAG $\alpha\text{-COO}^-$ (restraint)	$\uparrow K_a^{NAG} \sim 90\text{-fold}^c$
K1444A	K/r	K/R	Variable	Part of $\beta 4\text{-}\alpha 4$ loop^e	$\uparrow K_a^{NAG} \sim 3\text{-fold}$
F1445A	F/y/h	F/q/h/y	Variable	Part of $\beta 4\text{-}\alpha 4$ loop^e	No substantial effect
N1449A	N	N	Variable	H-bond with NAG $\alpha\text{-COO}^-$	$\uparrow K_a^{NAG} \sim 10\text{-fold}$

^aMutations studied experimentally here are highlighted in bold-type. For other details see Table 1. ^bD1322 belongs to the carbamate phosphorylation domain (see text). ^cPekkala et al., 2009. ^dNAG site lid residue and H-bond to NAG $\alpha\text{-COO}^-$ in previous docking model (Pekkala et al., 2009). ^eNAG site lid residue in previous docking model (Pekkala et al., 2009).