



PhD THESIS

**USING RECOMBINANT HUMAN CARBAMOYL PHOSPHATE
SYNTHETASE 1 (CPS1) FOR STUDYING THIS ENZYME'S
FUNCTION, REGULATION, PATHOLOGY AND STRUCTURE**

Author:

CARMEN DÍEZ FERNÁNDEZ

Directors:

Drs. Vicente Rubio Zamora y Javier Cervera Miralles

May 2015



Vicente Rubio Zamora, Doctor en Medicina, Profesor de Investigación del Consejo Superior de Investigaciones Científicas (CSIC) y Profesor Titular de Bioquímica (en Excedencia) de la Universidad de Valencia,

y

Javier Cervera Miralles, Doctor en Biología, Investigador jubilado del Centro de Investigación Príncipe Felipe de Valencia, de la Fundación Valenciana de Investigaciones Biomédicas de Valencia,

CERTIFICAN:

Que D^a, Carmen Díez Fernández, Ingeniera Agrónoma por la Universidad Politécnica de Valencia, con DNI 29208872E, ha realizado bajo su dirección el trabajo de Tesis Doctoral que lleva por título (en lengua inglesa): **“Using recombinant human carbamoyl phosphate synthetase 1 (CPS1) for studying this enzyme’s function, regulation, pathology and structure”**.

Que están de acuerdo con los contenidos del trabajo, realizado en la modalidad de acopio de publicaciones.

Que entienden que el trabajo reúne los requisitos para ser susceptible de defensa ante el Tribunal apropiado, por incluir dos publicaciones ya aparecidas y una sometida a evaluación, en revistas internacionales de amplia difusión y en las que la doctoranda es primera autora; así como porque el contenido de las otras secciones de dicho trabajo de tesis reflejan adecuadamente el estado del conocimiento y discuten apropiadamente el contenido y las implicaciones del trabajo realizado; porque las conclusiones recogen un avance suficiente en el campo objeto de estudio como para hacer a la candidata merecedora del doctorado; y, finalmente, porque les consta que ninguno de los coautores de dichas publicaciones, todos ellos ya doctores, han utilizado ni utilizarán las mismas para confeccionar otro trabajo de tesis doctoral.

Valencia, a 3 de Marzo de 2015.



Fdo. Vicente Rubio Zamora



Fdo. Javier Cervera Miralles

ACEPTACIÓN DE LOS COAUTORES PARA QUE LA DOCTORANDA
PRESENTE LOS SIGUIENTES ARTÍCULOS COMO TESIS; Y RENUNCIA
EXPRESA A PRESENTARLOS COMO PARTE DE OTRA TESIS
DOCTORAL

*ACCEPTANCE OF THE COAUTHORS FOR THE SUBMISSION BY THE
DOCTORAL CANDIDATE OF THE FOLLOWING RESEARCH
PUBLICATIONS OR MANUSCRIPTS AS PART OF HER DOCTORAL
THESIS, DECLINING THE USE OF THESE PUBLICATIONS AS A PART OF
ANOTHER THESIS WORK*

Los/las coautores/coautoras de los artículos científicos citados al pie, Drs. Ana Isabel Martínez, Satu Pekkala, Belén Barcelona, Isabel Pérez-Arellano, Ana María Guadalajara, Marshall Summar, Liyan Hu, José Gallego, Johannes Häberle, Javier Cervera Miralles y Vicente Rubio Zamora, aceptan que la doctoranda D^a. Carmen Díez Fernández, con DNI 29208872E, presente los artículos listados al pie como parte de la Tesis Doctoral con título “**USING RECOMBINANT HUMAN CARBAMOYL PHOSPHATE SYNTHETASE 1 (CPS1) FOR STUDYING THIS ENZYME’S FUNCTION, REGULATION, PATHOLOGY AND STRUCTURE**”, y renuncian a presentarlos como parte de otra tesis doctoral.

*The coauthors of the research papers listed below, Drs. Ana Isabel Martínez, Satu Pekkala, Belén Barcelona, Isabel Pérez-Arellano, Ana María Guadalajara, Marshall Summar, Liyan Hu, José Gallego, Johannes Häberle, Javier Cervera Miralles and Vicente Rubio Zamora, agree with the inclusion of the below listed research papers as a part of the Doctoral Thesis of Ms. Carmen Díez Fernández (Spanish DNI 29208872E), entitled “**USING RECOMBINANT HUMAN CARBAMOYL PHOSPHATE SYNTHETASE 1 (CPS1) FOR STUDYING THIS ENZYME’S FUNCTION, REGULATION, PATHOLOGY AND STRUCTURE**”. They decline to use them as a part of another doctoral thesis.*

Artículos científicos:
Scientific articles:

[1] Díez-Fernández C, Martínez AI, Pekkala S, Barcelona B, Pérez-Arellano I, Guadalajara AM, 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.

[2] 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.

[3] Díez-Fernández C, Gallego J, Häberle J, Cervera J, Rubio V. (2015) Experimental studies on the inherited metabolic error carbamoyl phosphate synthetase 1 deficiency shed light on the mechanism for switching on/off the urea cycle. Submitted to Journal of Genetics and Genomics.

Valencia, 3 de marzo 2015.

Valencia, March 3, 2015.



Dr. Javier Cervera Miralles



Dr. Vicente Rubio Zamora



Dr. Ana Isabel Martinez



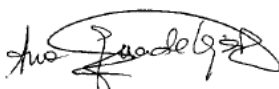
Dr. Satu Pekkala



Dr. Belén Barcelona



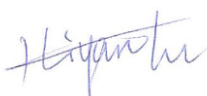
Dr. Isabel Pérez-Arellano



Dr. Ana María Guadalajara



Dr. Marshall Summar



Dr. Liyan Hu



Dr. Johannes Häberle



Dr. José Gallego

To carry out this PhD work entitled "**USING RECOMBINANT HUMAN CARBAMOYL PHOSPHATE SYNTHETASE 1 (CPS1) FOR STUDYING THIS ENZYME'S FUNCTION, REGULATION, PATHOLOGY AND STRUCTURE**", the doctoral candidate, Carmen Díez-Fernández was supported first by a Predoctoral fellowship of the Centro de Investigación Príncipe Felipe and then by an FPU fellowship (AP2009-2261) of the Ministry of Education of the Spanish Government, with additional funds from the same last source for a short-stay work visit in the Children's Hospital of Zürich.

The work was carried out within group 739 of CIBERER-Instituto de Salud Carlos III (IP, V. Rubio) under the auspices of the following research projects:

- BFU 2011-30407 of the Ministerio de Ciencia e Innovación of the Spanish Government, entitled "Luz estructural sobre señalización y regulación por nitrógeno y sobre biosíntesis de arginina/urea, sus errores congénitos y su conexión con biología del envejecimiento" (IP, V. Rubio).

- Prometeo/2009/051 of the Conselleria d'Educació de la Generalitat Valenciana, entitled: "BioMedec. Genes, proteínas y rutas de señalización en enfermedades raras" (investigators, F. Palau, P. Sanz y V. Rubio).

- Fundación Alicia Koplowitz 2012: "Encefalopatía hiperamoniémica por déficit de carbamil fosfato sintetasa 1 (CPS1). Hacia el diagnóstico, consejo genético y tratamiento a través de la estructura atómica y estudios funcionales de CPS1" (IPs, J. Cervera y V. Rubio).

- In Zurich, the work was carried out with the Swiss National Science Foundation grant 310030_127184 to Dr. Johannes Häberle.

Summary

Carbamoyl phosphate synthetase 1 (CPS1), a 1462-residue mitochondrial enzyme, catalyzes the entry of ammonia into the urea cycle, which converts ammonia, the neurotoxic waste product of protein catabolism, into barely toxic urea. The urea cycle inborn error and rare disease CPS1 deficiency (CPS1D) is inherited with mendelian autosomal recessive inheritance, being due to *CPS1* gene mutations (>200 mutations reported), and causing life-threatening hyperammonemia.

We have produced recombinantly human CPS1 (hCPS1) in a baculovirus/insect cell expression system, isolating the enzyme in active and highly purified form, in massive amounts. This has allowed enzyme crystallization for structural studies by X-ray diffraction (an off-shoot of the present studies). This hCPS1 production system allows site-directed mutagenesis and enzyme characterization as catalyst (activity, kinetics) and as protein (stability, aggregation state, domain composition). We have revealed previously unexplored traits of hCPS1 such as its domain composition, the ability of glycerol to replace the natural and essential CPS1 activator N-acetyl-L-glutamate (NAG), and the hCPS1 protection (chemical chaperoning) by NAG and by its pharmacological analog N-carbamyl-L-glutamate (NCG).

We have exploited this system to explore the effects on the activity, kinetic parameters and stability/folding of the enzyme, and to test the disease-causing nature, of mutations identified in patients with CPS1 deficiency (CPS1D). These results, supplemented with those obtained with other non-clinical mutations, have provided novel information on the functions of three non-catalytic domains of CPS1.

We have introduced three CPS1D-associated mutations and one trivial polymorphism in the glutaminase-like domain of CPS1, supporting a stabilizing and an activity-enhancing function of this non-catalytic domain. Two mutations introduced into the bicarbonate phosphorylation domain have shed light on bicarbonate binding and have directly confirmed the importance of this domain for NAG binding to the distant (in the sequence) C-terminal CPS1 domain. The introduction of 18 CPS1D-associated missense mutations mapping in a clinically highly eloquent central non-catalytic domain have proven the disease-causing nature of most of these mutations while showing that in most of the cases they trigger enzyme misfolding and/or destabilization. These results, by proving an important role of this domain in the structural integration of the multidomain CPS1 protein, have led us to call this domain the Integrating Domain.

Finally, we have examined the effects of eight CPS1D-associated mutations, of one trivial polymorphism and of five non-clinical mutations, all of them mapping in the C-terminal domain of the enzyme where NAG binds, whereas we have re-analyzed prior results with another four clinical and five non-clinical mutations affecting this domain. We have largely confirmed the pathogenic nature of the clinical mutations, predominantly because of decreased activity, in many cases due to hampered NAG binding. A few mutations had substantial negative effects on CPS1 stability/folding. Our analysis reveals that NAG activation begins with a movement of the final part of the $\beta 4$ - $\alpha 4$ loop of the NAG site. Transmission of the activating signal to the phosphorylation domains involves helix $\alpha 4$ from this domain and is possibly transmitted by the mutually homologous loops 1313-1332 and 778-787 (figures are residue numbers) belonging, respectively, to the carbamate and bicarbonate phosphorylation domains. These two homologous loops are called from here on Signal Transmission Loops.

Resumen

La carbamil fosfato sintetasa 1 (CPS1), una enzima mitocondrial de 1462 residuos, cataliza la entrada del amonio en el ciclo de la urea, que convierte esta neurotoxina derivada del catabolismo de las proteínas en urea, que es mucho menos tóxica. El déficit de CPS1 (CPS1D) es un error innato del ciclo de la urea, una enfermedad rara con herencia mendeliana autosómica recesiva, que se debe a mutaciones en el gen *CPS1* (>200 mutaciones descritas) y que cursa con hiperamonemia.

Hemos producido CPS1 humana recombinante (hCPS1) en un sistema de expresión de células de insecto y baculovirus, y la hemos aislado en forma activa y muy pura y en cantidad elevada. Ello ha permitido la cristalización de la enzima para estudios estructurales de difracción de rayos-X (un trabajo derivado de esta tesis pero no incluido en ella). Este sistema de producción de hCPS1 permite la realización de mutagénesis dirigida y la caracterización de la enzima como catalizador (actividad, cinética) y como proteína (estabilidad, estado de agregación y composición de dominios). Hemos revelado características de la hCPS1 antes no exploradas como es la composición de dominios, la capacidad que tiene el glicerol para reemplazar al activador natural y esencial de la CPS1, N-acetil-L-glutamato (NAG), y la protección de la hCPS1 por NAG y por su análogo farmacológico N-carbamil-L-glutamato (NCG) (chaperonas químicas).

Hemos utilizado este sistema para explorar los efectos en la actividad, los parámetros cinéticos y la estabilidad/plegamiento de la enzima, y para comprobar la naturaleza patogénica de mutaciones identificadas en pacientes con CPS1D. Estos resultados, junto con los obtenidos con otras mutaciones no clínicas, han aportado información novedosa sobre tres de los dominios no catalíticos de la CPS1.

Las observaciones realizadas tras introducir en el dominio de tipo glutaminasa de la enzima tres mutaciones asociadas a CPS1D y un polimorfismo trivial, apoyan la contribución de este dominio no catalítico a la estabilidad y a aumentar la actividad de la enzima. Dos mutaciones introducidas en el dominio de fosforilación de bicarbonato han arrojado algo de luz sobre el modo de unión del bicarbonato (un sustrato). Los resultados de estas mutaciones también han confirmado la contribución de este dominio para la unión de NAG, cuyo sitio de unión se encuentra en el dominio C-terminal de CPS1, bastante alejado (en la secuencia) del dominio de fosforilación de bicarbonato. Además, hemos introducido 18 mutaciones de cambio de sentido asociadas a CPS1D, las cuales están localizadas en un dominio no catalítico, central y de elevada elocuencia clínica. Estos resultados han demostrado la naturaleza patogénica de estas mutaciones, ya que en la mayoría de los casos estas mutaciones producen un mal plegamiento o/y desestabilización de la enzima. Debido a que estos resultados han puesto de manifiesto el importante papel de este dominio en la

integración estructural de la proteína multidominio CPS1, lo hemos llamado Dominio Integrador.

Finalmente, hemos examinado los efectos de ocho mutaciones asociadas a CPS1D, de un polimorfismo trivial y de cinco mutaciones no clínicas, todas ellas localizadas en el dominio C-terminal de la enzima, donde se une NAG. Además, hemos reanalizado resultados anteriores con otras cuatro mutaciones clínicas y cinco no clínicas afectando a este dominio. Hemos confirmado el carácter patogénico de las mutaciones clínicas, las cuales predominantemente causan una disminución en la actividad enzimática, en muchos casos debida a que la unión de NAG se encuentra obstaculizada. Unas pocas mutaciones mostraron efectos negativos sustanciales en la estabilidad/plegamiento de la CPS1. Nuestros análisis revelan que la activación por el NAG empieza con un movimiento de la parte final del bucle $\beta 4-\alpha 4$ del sitio de NAG. La transmisión de la señal activadora a los dominios de fosforilación implica a la hélice $\alpha 4$ de este dominio y posiblemente se transmite a través de los bucles homólogos 1313-1332 y 778-787 (numeración de los residuos) pertenecientes, respectivamente, a los dominios de fosforilación de carbamato y bicarbonato. Por ello, hemos llamado a ambos bucles homólogos Bucles de Transmisión de la Señal.

Resum

La carbamil fosfat sintetasa 1 (CPS1), un enzim mitocondrial de 1462 residus, catalitza l'entrada d'amoni en el cicle de la urea, que convertix l'amoni, producte neurotòxic del catabolisme de les proteïnes, en urea, una molècula molt poc tòxica. El dèficit de CPS1 (CPS1D) és un error innat del cicle de la urea, una malaltia rara amb herència mendeliana autosòmica recessiva, que es deu a mutacions en el gen CPS1 (>200 mutacions descrites) i que cursa amb hiperamonièmia.

Hem produït CPS1 humana recombinant (hCPS1) en un sistema d'expressió de cèl·lules d'insecte i baculovirus, i l'hem aïllada en forma activa i molt pura i en gran quantitat. Això ha permès la cristal·lització de l'enzim per a estudis estructurals amb difracció de raïos-X (treball no inclòs en esta tesi). Aquest sistema de producció de hCPS1 permet la realització de mutagènesi dirigida i la caracterització de l'enzim com a catalitzador (activitat, cinètica) i com a proteïna (estabilitat, estat d'agregació i composició de dominis). Hem revelat característiques de la hCPS1 no explorades abans com és la composició de dominis, la capacitat que té el glicerol per a reemplaçar l'activador natural i essencial de CPS1, N-acetil-L-glutamat (NAG), i la protecció de la hCPS1 per NAG i pel seu anàleg farmacològic N-carbamil-L-glutamat (NCG) (xaperones químiques).

Hem utilitzat aquest sistema per a explorar els efectes en l'activitat, els paràmetres cinètics i l'estabilitat/plegament de l'enzim, i per a comprovar la naturalesa patogènica de mutacions identificades en pacients amb CPS1D. Aquestos resultats, junt amb els obtinguts amb altres mutacions no clíniques, han aportat informació nova sobre tres dels dominis no catalítics de la CPS1.

Les observacions, després d'introduir tres mutacions associades a CPS1D i un polimorfisme trivial en el domini tipus glutaminasa de CPS1, recolzen la contribució d'aquest domini no catalític a l'estabilitat i a l'optimització de l'activitat enzimàtica. Dues mutacions introduïdes en el domini de fosforilació de bicarbonat han esclarit el mode d'unió de bicarbonat. Els resultats d'aquestes mutacions també han confirmat la contribució d'aquest domini per a la unió de NAG, el lloc d'unió de la qual es troba en el domini C-terminal de CPS1, prou allunyat (en la seqüència) del domini de fosforilació de bicarbonat. A més, hem introduït 18 mutacions de canvi de sentit associades a CPS1D, les quals estan localitzades en un domini no catalític, central i d'elevada eloqüència clínica. Aquestos resultats han demostrat la naturalesa patogènica d'aquestes mutacions, ja que, en la majoria dels casos produïxen un mal plegament o/i desestabilització de l'enzim. Pel fet que aquestos resultats han posat de manifest

l'important paper d'aquest domini en la integració estructural de la proteïna multidomini CPS1, l'hem anomenat Domini Integrador.

Finalment, hem examinat els efectes de huit mutacions associades a CPS1D, un polimorfisme trivial i cinc mutacions no clíniques, totes elles localitzades en el domini C-terminal de l'enzim, on s'unix NAG. A més, hem reanalitzat resultats anteriors amb altres quatre mutacions clíniques i cinc no clíniques que afecten aquest domini. Hem confirmat el caràcter patogènic de les mutacions clíniques, les quals predominantment causen una disminució en l'activitat enzimàtica, en molts casos pel fet que la unió de NAG es troba obstaculitzada. Unes poques mutacions van mostrar efectes negatius substancials en l'estabilitat/plegament de CPS1. Les nostres anàlisis revelen que l'activació de NAG comença amb un moviment de la part final del bucle $\beta 4-\alpha 4$ del lloc de NAG. La transmissió del senyal activadora als dominis de fosforilació involucra l'hèlix $\alpha 4$ d'aquest domini i es transmet, possiblement, a través dels bucles homòlegs 1313-1332 i 778-787 (numeració dels residus), pertanyents, respectivament, als dominis de fosforilació de carbamato i bicarbonat. Per això, hem anomenat a ambdós bucles homòlegs Bucles de Transmissió del Senyal.

Agradecimientos / Acknowledgements

A lo largo de estos años de investigación he recibido el apoyo de muchas personas. La presentación de esta tesis no habría sido igual sin la aportación de cada una de ellas.

En primer lugar quiero agradecer a mis directores de tesis, el Prof. Vicente Rubio y el Dr. Javier Cervera, por abrirme las puertas al mundo de la investigación, al haberme dado la oportunidad de desarrollar mi trabajo de doctorado en sus laboratorios; por haberme enseñado lo que es la ciencia y el valor y la importancia del trabajo bien hecho; por la paciencia y dedicación que han tenido conmigo.

Muchas gracias a mi tutor en la Universidad Politécnica de Valencia, Dr. Ramón Serrano y a su mujer Mariche, por haberme apoyado siempre desde que nos conocimos en la Escuela de Agrónomos, por haberme transmitido ilusión por la ciencia, por haberme brindado la oportunidad de impartir clases con ellos y por haberme ayudado y animado en todo momento.

I would also like to thank Dr. Johannes Häberle, for giving me the possibility of enjoying a short stay in his laboratory in Zurich Children Hospital, for showing me the importance of research and the value of our work. As well, I am grateful to all the colleagues in the lab in Zurich for their help and patience during my adaptation there.

A mis compañeros del IBV: Clara, Carles, Sergio, Juanma, Nadine, Alicia, Lorena, Ariel y Belén. Son todos magníficos compañeros y amigos. Por su ayuda, por compartir conmigo sus experiencias y consejos y por los buenos momentos que hemos pasado juntos. En particular quería agradecer a Super Nadine por su ayuda en el día a día; a Clara, Carles y Sergio, expertos cristalógrafos, por sus sabios consejos y a Belén, una muy buena amiga, que además de aportarnos el modelo estructural de CPS1, me ha apoyado y ayudado tanto científica como personalmente. A mi buena amiga Laura, por su amistad, apoyo y tantas cenas juntas.

A mis compañeros del CIPF, donde lleve a cabo los dos primeros años de esta tesis: Lorena, Rober, Belén, Pilar, Ana y Satu. A pesar de habernos tenido que separar, siempre han estado a mi lado y han compartido conmigo su ayuda, experiencia, consejos y apoyo.

A toda la gente del IBV y del CIPF, investigadores, becarios, técnicos, personal de la biblioteca, administración, informática, limpieza, mantenimiento,

seguridad, informática, por haberme acogido tan bien y por solucionar todos los problemas que surgían en el día a día.

A todas mis amigas, en especial a Consuelo y Natalia, por estar siempre a mi lado, por sus mensajitos de ánimo, por sus visitas inesperadas y por las carreras por el río a cualquier hora del día o de la noche.

A mi buena amiga María y a su madre M^a Dolores, por haberme cuidado todos estos años.

A mi familia, la más maravillosa del mundo. Quiero darles las gracias por haberme animado todos estos años a conseguir mi sueño, por su apoyo constante, por siempre ser tan cariñosos, por cuidarme en los buenos momentos y ayudarme a levantarme en los no tan buenos, por aportarme la fuerza que necesitaba (incluidas las comidas tan buenas y los pastelitos diarios) y por enseñarme a no rendirme nunca. A mi abuela, por rezar por mí cada día, por las conversaciones telefónicas y por los guisos deliciosos. A mis sobrinos Luis y Adriana, que siempre me regalan sonrisas y carcajadas.

A Jaime, por la mitocondria que me dibujó, por la paciencia que ha tenido todo este tiempo, por su creatividad para lidiar con mis cambios de humor, por su comprensión y su cariño.

Esta tesis esta dedicada a mis padres, a mi hermano y a Jaime.

Non-standard Abbreviations

3D: Three-dimension

6xHis or His₆-tag: 6 histidine tag

AcNPV: *Autographa californica* nuclear polyhedrosis virus

ALT: Alanine transaminase

AMPPNP: Adenylyl-imidodiphosphate

ARG1: Arginase type 1

ASD: C-Terminal allosteric domain of CPS1

ASL: Argininosuccinate lyase

ASS: Argininosuccinate synthetase

AST: Aspartic transaminase

ATC: Aspartate transcarbamylase

attTn7: Target site of the Tn7 transposon

BPSD: Bicarbonate phosphorylation domain of CPS1

CAD: Multifunctional polypeptide that catalyzes the initial three steps of pyrimidine biosynthesis. It is composed of CPS II, dihydroorotase and aspartate transcarbamylase (in this order from N-terminus to C-terminus)

CP: Carbamoyl-phosphate

CPS: Carbamoyl-phosphate synthetase

CPS1: CPS devoted to making urea. It uses ammonia and is NAG-activated

CPS1D: Deficiency of CPS1

CPSII: CPS used for making pyrimidines in eukaryotes

CPSIII: CPS devoted to making urea in fishes. It uses glutamine but it is activated by NAG.

CPSD: Carbamate phosphorylation domain of CPS1

Chymo: Chymotrypsin

DHO: Dihydroorotase

DTT and dithiothreitol: 1,4-dithio-D-threitol

ABBREVIATIONS

FUM: Fumarase

G6PDH: glucose-6-phosphate dehydrogenase

GDH: L-Glutamate dehydrogenase

GLNase: Glutaminase

GSD: Glutaminase-like domain of CPS1

hCPS1 or hCPS: human CPS1

HK: Hexokinase

ID: Integrating domain, a central domain of CPSs

ISD: Most N-terminal domain of CPS1

K_a : For hyperbolic activation, concentration of the activator that causes half-maximal activation

Kana: Kanamycin

LacZ α : Gene for α peptide of LacZ

LDH: Lactate dehydrogenase

MDH: Malate dehydrogenase

m.o.i.: Multiplicity of infection, number of infecting viral particles per cell to be infected

NAG and acetylglutamate: N-Acetyl-L-glutamate

NAGS: N-Acetyl-L-glutamate synthase

NAGSD: NAGS deficiency

NCG: N-Carbamyl-L-glutamate

OMIM or MIM: Online Mendelian Inheritance in Man, the online version of the McKusick book, "Mendelian Inheritance in Man", a repertoire of genetic diseases and pathology-relevant human genes and gene products. It gives a number code to each reported genetic disease. Can be accessed freely at <http://www.ncbi.nlm.nih.gov/omim>

Ori-pUC: Origin of replication of the pUC plasmid

ORN: L-Ornithine

ORNT 1: Ornithine/citrulline antiporter

OTC: Ornithine transcarbamylase

ABBREVIATIONS

PDB: Protein Data Bank, a repertoire of macromolecular structures that can be accessed at <http://www.rcsb.org/pdb/home/home.do>

Pfu Turbo: *Pyrochoccus furiosus* (an organism that lives at nearly 100°C) DNA polymerase

PK: Pyruvate kinase

P_{PH}: Polyhedrin promoter

PRPP: Phosphoribosyl pyrophosphate

rhCPS1: Recombinant human CPS1

SDS-PAGE: Polyacrylamide gel electrophoresis in the presence of the protein-binding detergent sodium dodecyl sulphate

Sf: the moth *Spodoptera frugiperda*

St: commercial protein markers

SV40 poly A: Simian virus 40 polyadenylation signal

Tetra: Tetracyclin gene resistance

TEV: Tobacco Etch Virus

Tn7L and Tn7R: left and right arms of the Tn7 transposon

UFSD: previous acronym for the integrating domain (ID) of CPSs

ABBREVIATIONS

Contents

	Page
1. INTRODUCTION.....	1
1.1 Biological context and scope of this thesis.....	3
1.2 The urea cycle and the consequences of its derangements...	5
1.3 Biological significance of CPS1 and of its allosteric regulation by NAG.....	9
1.4 A brief review of knowledge about CPS1 that is relevant for this work.....	12
1.4.1 <i>Early findings</i>	12
1.4.2 <i>Initial work with pure CPSs</i>	13
1.4.3 <i>Reactional mechanism</i>	14
1.4.4 <i>Human CPS1</i>	15
1.4.5 <i>NAG activation and domain structure</i>	15
1.4.6 <i>Cloning and sequencing of CPS1 cDNA and gross enzyme architecture</i>	16
1.4.7 <i>Crystal structure of E. coli CPS</i>	17
1.4.8 <i>Crystallographic data on the structure of the C-terminal domain of human CPS1, and NAG site delineation</i>	19
1.4.9 <i>Modelling of the CPS1 structure</i>	20
1.4.10 <i>CPS1 production</i>	21
1.5 CPS1 deficiency, the CPS1 gene and its mutations.....	22
1.6 Other roles of CPS1 in pathology.....	26
1.6.1 <i>Secondary hyperammonemia</i>	26
1.6.2 <i>A CPS1 polymorphism involved in vascular reactivity and in other traits related to pathologies</i>	27

CONTENTS

	Page
1.7 Previous in vitro expression and structure-function studies on CPSs.....	29
1.8 The baculovirus/insect cell expression system.....	31
1.9 Site-directed mutagenesis.....	38
1.10 CPS1 activity assays.....	41
1.11 Study goals, contents, involved laboratories, and role of myself.....	44
2. RESULTS	47
2.1 Results. Chapter 1.....	49
2.1.1 <i>Abstract</i>	50
2.1.2 <i>Introduction</i>	51
2.1.3 <i>Patients and Methods</i>	54
2.1.3.1 Patients and CPS1 mutations.....	54
2.1.3.2 Recombinant human CPS1 production..	54
2.1.3.3 Enzyme purification.....	57
2.1.3.4 Enzyme activity assays.....	58
2.1.3.5 Other techniques.....	59
2.1.4 <i>Results and Discussion</i>	60
2.1.4.1 Producing human CPS1 in a baculo-virus/insect cell system.....	60
2.1.4.2 Recombinant human CPS1 represents well the natural liver enzyme.....	61
2.1.4.3 Glycerol partially replaces NAG in the activation of human CPS1.....	62

	Page
2.1.4.4 Limited proteolysis reveals the hCPS1 domain organization.....	63
2.1.4.5 Testing the effects of clinical CPS1D mutations on enzyme functionality.....	64
2.1.4.6 Influence of the substrates and of NAG on the resistance of human CPS1 to proteolytic or thermal inactivation.....	68
2.1.4.7 Final comments.....	69
2.1.5 <i>Acknowledgements and Disclosure statement</i>	74
2.1.6 <i>References</i>	75
2.2 Results. Chapter 2	80
2.2.1 <i>Abstract and key words</i>	81
2.2.2 <i>Introduction</i>	82
2.2.3 <i>Materials and Methods</i>	86
2.2.3.1 Patients and CPS1 Mutations.....	86
2.2.3.2 Production of CPS1 carrying the desired mutations.....	86
2.2.3.3 Enzyme activity assays.....	88
2.2.3.4 Other assays.....	88
2.2.4 <i>Results</i>	90
2.2.4.1 CPS1D mutations that affect the UFSD	90
2.2.4.2 UFSD mutations negatively influence the yield of CPS1.....	91
2.2.4.3 Eight UFSD mutations strongly impair CPS1 activity.....	93
2.2.4.4 Activity and stability changes with another eight UFSD mutations.....	93

CONTENTS

	Page
2.2.4.5 K_m effects are the major changes associated with mutations affecting Arg850.....	96
2.2.4.6 UFSD mutations generally decrease V_{max} of the enzyme.....	98
2.2.5 <i>Discussion</i>	99
2.2.6 <i>Supplementary material: Structural rationalization of the mutations effects</i>	101
2.2.7 <i>Acknowledgements/conflict of interest statement</i>	104
2.2.8 <i>References</i>	104
2.3 Results. Chapter 3	110
2.3.1 <i>Abstract and key words</i>	111
2.3.2 <i>Introduction</i>	112
2.3.3 <i>Results</i>	115
2.3.3.1 Clinical ASD domain mutations studied	115
2.3.3.2 Impact of the clinical mutations on enzyme stability.....	115
2.3.3.3 Effects of clinical ASD mutations on CPS1 activity and on the kinetic parameters for NAG.....	116
2.3.3.4 The effects of ASD mutations shed further light on signal transmission and NAG binding.....	120
2.3.3.5 Modeling of the NAG site.....	122
2.3.4 <i>Discussion</i>	123
2.3.5 <i>Materials and Methods</i>	127
2.3.5.1 Human CPS1 production.....	127
2.3.5.2 CPS1 activity assays.....	128

CONTENTS

	Page
2.3.5.3 Restrained molecular dynamics (MD) and docking calculations.....	128
2.3.5.4 Other techniques.....	129
2.3.6 <i>Acknowledgements</i>	130
2.3.7 <i>References</i>	130
3. DISCUSSION.....	135
4. CONCLUSIONS.....	149
5. REFERENCES.....	153

CONTENTS

1 Introduction

INTRODUCTION

1.1 Biological context and scope of this thesis

Carbamoyl-phosphate synthetases (CPSs) are complex multidomain enzymes [1-3] that catalyse the synthesis of carbamoyl phosphate (CP) from two molecules of ATP and one molecule of each bicarbonate and ammonia (Fig. 1). Ammonia is used as such by the enzyme that is responsible for initiating the urea cycle and which is the subject of this study, CPS1. However, all other CPSs use ammonia derived internally from glutamine, an amino acid that is hydrolyzed by these other types of CPS.

Search for CPS sequences in the many genomes deposited in sequence databases reveals that CPSs are present in virtually all organisms. The wide species distribution of CPSs reflects the fact that carbamoyl phosphate (CP) is the initial precursor of pyrimidines [4-6] and arginine [6-9]. Arginine is a crucial amino acid. It is a key protein component, and a precursor of nitric oxide, creatine, polyamines, agmatine and proline (Fig. 1) [10]. Furthermore, in some fishes, in amphibians and mammals, arginine has a central role in urea production. In these last two types of vertebrates urea is synthesized to get rid of ammonia. By catalyzing the entry of ammonia into the urea cycle, CPS1 has paramount importance in these organisms for ammonia detoxification [11].

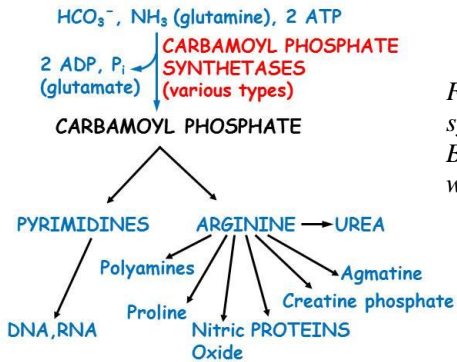


Figure 1. Roles of carbamoyl phosphate synthetases. Blue arrows, CPS reaction. Black arrows, metabolic conversions in which the indicated molecules are involved.

This thesis work deals with human CPS1 and with its deficiency (CPS1D) (OMIM #237300; <http://www.ncbi.nlm.nih.gov/omim>). CPS1D is an autosomal recessive inborn error of the urea cycle, which causes hyperammonemia, potentially leading to death or mental retardation [12]. More than 200 CPS1D-associated mutations have been reported in the *CPS1* gene (located in 2q34). The majority of these mutations are missense changes (Fig. 2). In many cases, the disease-causing role of these missense mutations is unproven [12]. This situation was recently alleviated when our group reported [13] a baculovirus/insect cell expression system that permitted *in vitro* production of rat CPS1, opening the way for accessible site-directed mutagenesis studies with this enzyme [14].

INTRODUCTION

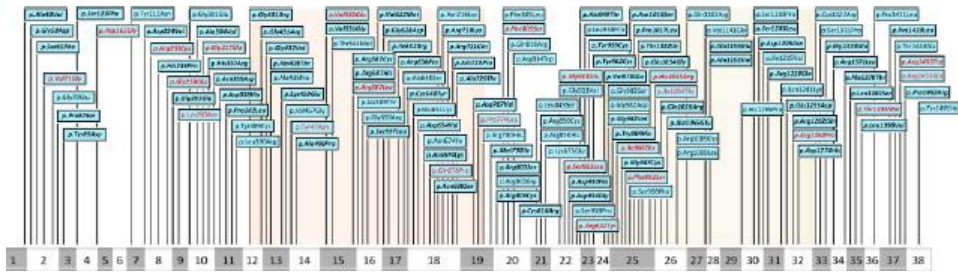


Figure 2. The *CPS1* coding sequence split into its different exons (bottom), to show the location of the different missense mutations (banners) recently compiled or reported in *CPS1* deficiency [12]. The figure aims at giving a view of the abundance and spreading of the mutations over the entire sequence, rather than to give specific information on each mutation (blurred because the figure has been taken from the original publication [12], being a part of Fig. 1 of that publication).

The present work adapts this baculovirus/insect cell expression system to endow it with the ability to produce human CPS1. Recombinant human CPS1 is produced and characterized here, showing that it has essentially the same properties as human CPS1 obtained from its natural source (the liver) [15,16]. With the recombinant enzyme we have been able to examine properties of human CPS1 that had not been studied previously, such as the domain composition of the enzyme. Furthermore, abundant expression of recombinant human CPS1 has opened the way to crystallographic studies that are culminating in the determination of the crystal structure of the human enzyme (collaborative work with other group members not reported here).

In the recent mutational repertory of *CPS1D* [12] the density of missense mutations was found not to be constant along the *CPS1* gene and protein. A particularly high frequency of missense mutations was identified in a central domain of CPS1 having no known function. We have applied here the baculovirus/insect cell expression system to investigate by site-directed mutagenesis the effects on the functionality and stability of human CPS1 of all the missense mutations mapping in this central domain that were reported in *CPS1D* patients. With this approach we aimed at testing the disease-causing role of these mutations and also to determine the functions of this central domain and to clarify the reasons for the high mutational eloquence in *CPS1D*.

We also present here a site-directed mutagenesis study of the C-terminal domain of human CPS1, a domain called the allosteric domain [1-3] because it hosts the site for the essential allosteric activator of the enzyme, N-acetyl-L-

INTRODUCTION

glutamate (NAG) [13,17]. Our study includes the expression of patient-identified and of rationally designed mutations as well as the analysis of the effects of already studied mutations mapping in this domain either as a part of the present work (Chapter 1 of the Results) or of previous work of other group members with recombinant CPS1 from rat [13,14]. This study has not only provided insight into disease-causation but also has shed important light into the process of NAG activation.

The Results section of this thesis is composed of three chapters corresponding to three manuscripts (two of them published, one under review). Since these manuscripts have their own introductions, this general introduction will aim at providing a broader context for the present work, trying to avoid repetitions and redundancies.

1.2 The urea cycle and the consequences of its derangements

In vertebrates, ammonia, the final nitrogenous waste product of protein and amino acid catabolism, is a strong neurotoxin [18,19] that has to be eliminated from the organism. Fishes can do it by direct dilution in the surrounding water. Reptiles and birds, generally having limited water access particularly during development, convert the waste nitrogen into uric acid, a highly insoluble purine derivative that precipitates out of solution. In the case of mammals the viviparous foetus is mainly ammonotelic (meaning excretion of ammonia) during development, with the mother being in charge of ammonia disposal, which in this case occurs by conversion to urea (ureotelism) [20,21].

Urea [22] is the diamide of carbonic acid, a simple molecule consisting of one atom of each carbon and oxygen, of two atoms of nitrogen and four atoms of hydrogen. Since the dissociation of amidic hydrogens has very high pK values, urea is not an acid at body pH or even at urinary pH. Given the high polarity of the O and N atoms and the presence of four hydrogen atoms linked to nitrogen atoms, it can be dissolved in water at supramolar concentrations. Consequently, urea can be concentrated much in the urine of mammals, which have generally enough water availability to prevent its crystallization there. The kidneys have developed urea-impermeable tubules. In this way, while most of the water filtered in the glomerulus is resorbed later on, the urea is concentrated in the urine and becomes the main nitrogenous waste product [23]. Furthermore, perhaps because of the non-dissociability of its hydrogens, urea is little toxic, and thus it is tolerable to the body at much higher concentration than ammonia [24].

INTRODUCTION

Krebs, together with Henseleit, proposed eighty three years ago how ammonia was converted to urea in mammals, creating the concept of metabolic cycle (Fig. 3) [25,26]. The original (as redrawn by Krebs in 1976) and a more recent view [27] of this cycle are shown in Fig. 3.

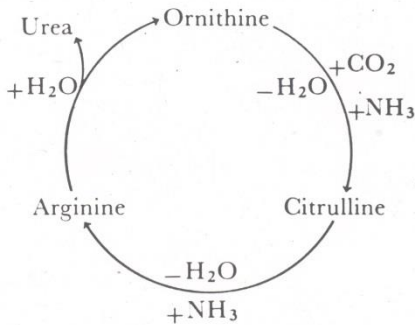
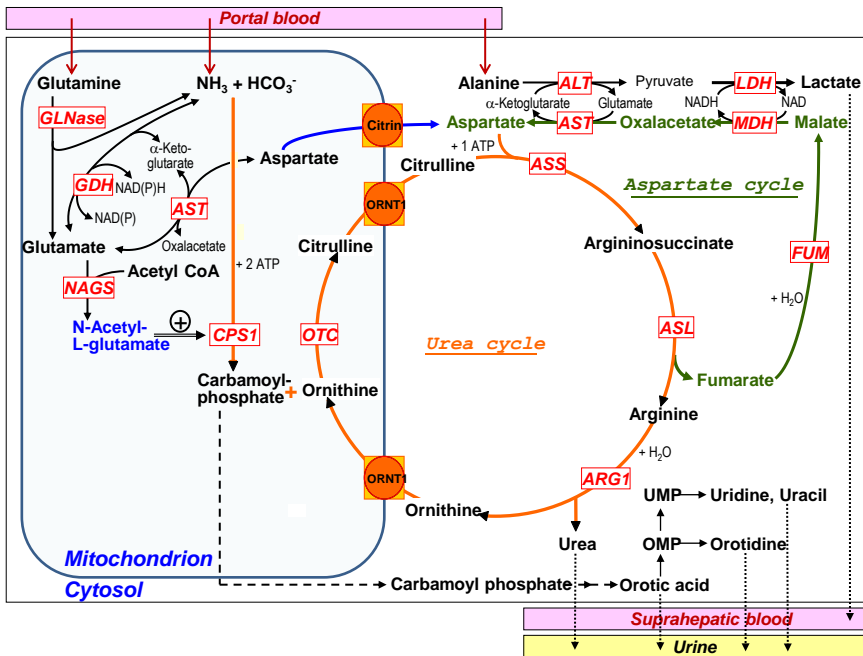


Figure 3. Early (left) and present (down) views of the urea cycle. The view on the left is as originally formulated by Krebs, in a drawing of this author in 1976 [26]. The view down (courtesy of Drs. Häberle and Rubio) of this cycle (orange arrows) includes ancillary reactions (other arrows). The relation with pyrimidine synthesis is shown also. Enzymes and transporters are shown in red and boxed.



NAGS, NAG synthase; OTC, ornithine transcarbamylase; ASS, argininosuccinate synthetase; ASL, argininosuccinate lyase; ARG1, arginase type 1; GLNase, glutaminase; GDH, glutamate dehydrogenase; AST, aspartic transaminase; ALT, alanine transaminase; MDH, malate dehydrogenase; LDH, lactate dehydrogenase; FUM, fumarase.

INTRODUCTION

In mammals, the urea cycle is found in complete form exclusively in hepatocytes, excepting the perivenous hepatocytes, which convert the ammonia to glutamine instead of to urea [28]. The cycle is split into cytosolic and intramitochondrial compartments, with the two initial steps, including the step of carbamoyl phosphate synthesis, being localized in the mitochondrial matrix (Fig. 3) [7,11].

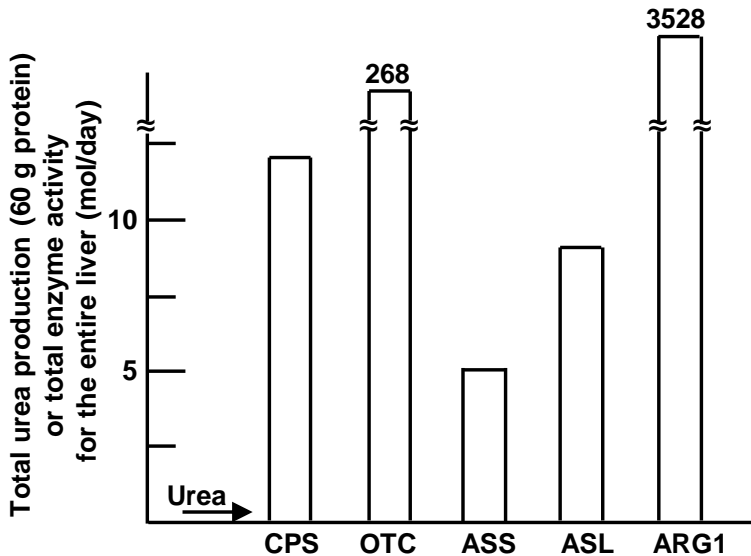


Figure 4. Total liver activities of the different urea cycle enzymes (abbreviated as in Fig. 3) given as product synthesis per day for a 1.5 kg liver. Derived from enzyme activity data in homogenates of human liver biopsies [29] assayed at optimal pH and substrate concentrations. Note the very high activities of OTC and ARG1 (268 and 3528 mole/day). The arrow indicates urea production per day for an intake of 60 g protein per day. Modified from ref. [11].

Fig. 4 shows the relative activities of the different urea cycle enzymes [29] in an entire human liver, given as ability to make urea per day, and compared with the need to make urea in an adult consuming a normal diet (containing 60g protein day). It can be seen that, in principle, there is considerable functional reserve, and that CPS1 is among the three enzymes with the lowest activity among all the urea cycle enzymes. Complete absence of any urea cycle enzyme, particularly of those catalysing the two initial steps (CPS1 and ornithine transcarbamylase), causes hyperammonemia, which, given the neurotoxicity of ammonia, soon leads to coma and death [27, 30]. Because of the considerable functional reserve (Fig. 4), heterozygous individuals for any urea cycle catalyst (except for the X-chromosome-encoded ornithine transcarbamylase) should not

INTRODUCTION

present pathology. Furthermore, since the heterozygous mother can take care of disposal of the ammonia made by the foetus, the hyperammonemic phenotype of individuals with complete deficiency is only manifested after birth, when the connection with the mother is lost and the newborn has to cope with the challenge to dispose of ammonia to urea using his/her own urea cycle enzymes.

Complete enzyme deficiency, at least of the initial two steps of the urea cycle, is therefore constantly associated with neonatal hyperammonemia (called "early onset" or "neonatal presentation") and death by approximately 72 hours after birth [27,30]. However, given the important functional reserve, patients with partial deficiencies caused by low but not zero enzyme levels or by "kinetic" enzyme variants having high K_m values for the substrates (and also, in the case of CPS1, for the essential activator NAG) have in many cases enough enzyme activity to cope with the normal burden of ammonia detoxification, presenting with hyperammonemia only when the system is overloaded with ammonia, as in catabolic situations caused by fever and infections, trauma, surgery, delivery, or in peaks of protein intake [30]. In these cases the neonate may not present early onset hyperammonemia after birth, having a delayed course that can have either a late onset presentation early in life or even much later, when growth rate (and thus, protein generation for building the body) decreases, or even in adult life. Nevertheless, even in late onset cases the outcome may be equally dramatic as in neonatal cases, since hyperammonemic crises can be deadly or can cause mental retardation irrespective of the level of residual enzyme activity [30].

Although outside the scope of this work, it is important to mention that there are effective therapeutic measures [30] that can be applied to patients with urea cycle disorders to change the natural course of the disease and even to cure it, rendering it essential to apply therapy promptly and aggressively. Our present work can be related to the diagnosis and treatment of urea cycle deficiency because of the following:

- 1) Specific disease diagnosis is now based on the detection of genetic lesions [12], which are in many cases missense mutations. Characterization of the effects of these mutations may not only help decide whether a given missense mutation is or is not disease-causing, but can also provide insight on its severity and the need to be prepared for treatment in cases in which the presence of a given mutation in the foetus is known before delivery.
- 2) Specific prognostic judgements associated with the approach used in our present work can guide therapeutic decisions. For example, the identification of highly severe mutations would be an indication for particularly close

INTRODUCTION

monitoring of the patient and might help tilt the balance in the direction of, for example, enrolling in liver cell therapy (presently an experimental treatment) and/or in liver transplantation as soon as possible.

3) Identification of specific traits that could suggest susceptibility to given therapies. In the case of CPS1D this is related only, for the moment, to N-acetyl-L-glutamate (NAG) activation (discussed below), since there is a pharmacologic NAG analog and registered orphan drug, N-carbamylglutamate [31], that could be tested for efficiency in those patients having CPS1 mutations in which the affinity for NAG is decreased, such as those that are the subject of Chapter 3 of the Results.

4) Development of CPS1-stabilizing therapies. Pharmacological chaperones for treating inborn errors have gained momentum after the discovery that the cofactor used by phenylalanine hydroxylase, tetrahydrobiopterin, can be used in the therapy of phenylketonuria, having been found to increase the level of the enzyme "*in vivo*" by increasing the stability of the enzyme [32]. As it will be shown in Chapter 1 of the Results, we provide here promising findings with carbamylglutamate that may herald this approach in the treatment of CPS1D.

In the next two sections we will concentrate on highlighting the biological significance of CPS1 as well as to review our knowledge on this enzyme.

1.3 Biological significance of CPS1 and of its allosteric regulation by NAG

Human CPS1 is inactive in the absence of NAG [16]. Since rodent studies [33,34] indicate that the level of NAG depends on the metabolic situation, CPS1 can be considered a true switch for turning the urea cycle on or off depending on the availability of NAG.

Why should the urea cycle have such a switch? The most reasonable explanation stems from the fact that ammonia is not only a toxin but also an essential life component. Although ammonia levels cannot be allowed to be too high, because of brain toxicity, they should not be too low either. Ammonia is in equilibrium with several non-essential amino acids (such as serine, glutamine, asparagine, glutamate and via glutamate with other amino acids) [35] and the levels of these amino acids would decrease with the ammonia level. Therefore, the activity of the urea cycle cannot be controlled exclusively by the availability of ammonia, because this would lead to ammonia depletion, and thus to drainage of amino acids and of proteins. NAG is synthesized in the same cells and cell compartment (the mitochondrial matrix) where CPS1 is

INTRODUCTION

present. It is made by a dedicated enzyme that uses acetyl-CoA and L-glutamate [36]. Since NAG has a short $t_{1/2}$ [37] and NAGS has a high K_m value for glutamate [36], the NAG levels is a true sensor of the nitrogen load reflected in the level of glutamate (Fig. 5).

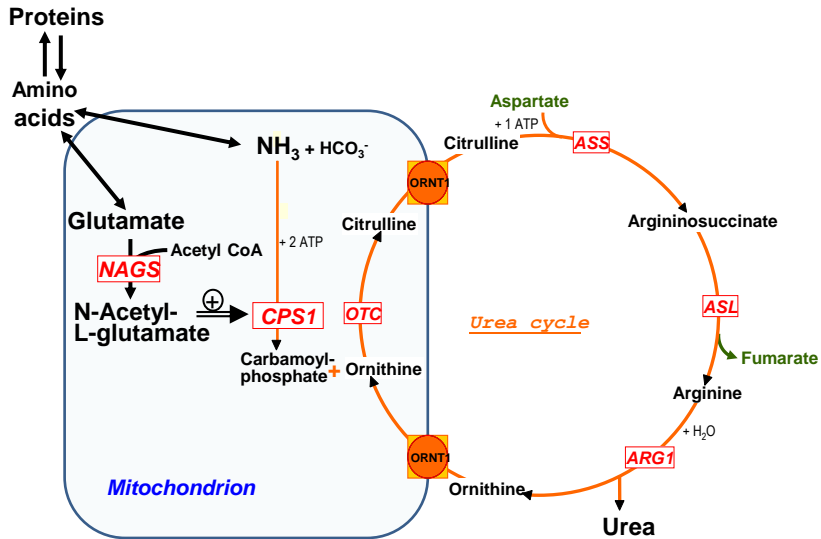


Figure 5. Role of NAG as a safety switch for preventing amino acid and protein depletion. For further details see the text.

Why was acetylglutamate chosen by natural selection for such role instead of glutamate? The answer to this question is not clear. The choosing of NAG as the CPS1 effector was a relatively recent occurrence, since NAG is a CPS activator only in animals. The CPSs from more ancient taxonomic groups (including bacteria, archaea and plants) are insensitive to NAG [1-9]. Actually, in these organisms NAG is the first committed precursor of ornithine and thus of arginine [8,9] (Fig. 6, left panel). Interestingly, as discussed below, the NAG site of CPS1 is equivalent to the site of *Escherichia coli* CPS for IMP [13], a purine nucleotide that activates this bacterial CPS [6]. Therefore, at some moment in the evolution towards animal arginine biosynthesis all the genes for the enzymes that catalyze the conversion of NAG into ornithine were lost, with exclusive preservation of NAG synthase, which became exclusively devoted to making NAG for the purpose of CPS1 activation (Fig. 6, right panel). Simultaneously CPS must have shifted its allosteric activator specificity from IMP to NAG losing the ability to be activated by ornithine (another characteristic of many bacterial CPSs) [6]. Further evolution also involved an

INTRODUCTION

increase in the on/off ratio of activities of the enzyme depending on whether is present or is not present, since bacterial CPSs are active in the absence of any effectors [6], fish CPSIII (another type of arginine-making CPS considered a close predecessor of CPS1) is partially dependent on the presence of NAG [38], whereas CPS1 has a complete dependency on this activator [16,39], representing an extreme case of allosteric activation [40].

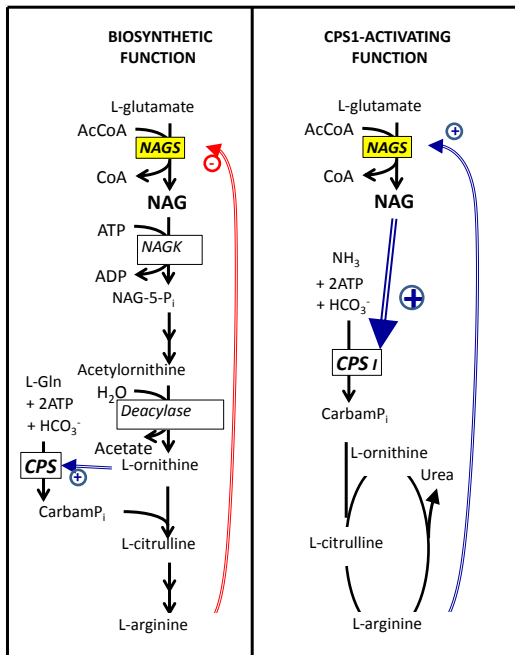


Figure 6. Shift in the function of NAG from bacteria and plants (left panel), where it is the first committed intermediate in arginine synthesis, to mammals (right panel), where it activates CPS1. Blue arrows denote activation and red arrows denote feed-back inhibition.

Fig. 7 schematizes the structural traits and Table 1 summarizes the differential properties of various types of CPS. In addition to the ability to be activated by NAG, CPS1 differs from other types of CPS in its inability to use glutamine as substrate, and in its high affinity for ammonia [16,39], approximately 100-fold higher than that of the bacterial enzyme [6]. The second of these properties, unique to CPS1, is understandable as an adaptive measure because of the major role of CPS1 as a detoxifier of ammonia at the concentrations of this metabolite that prevail in portal blood. The inability of CPS1 to utilize glutamine is also understandable, since in the liver there are two systems that coexist to get rid of ammonia [28]: the high capacity but relatively low-affinity system represented by the urea cycle, which is present in periportal hepatocytes, and the low capacity/high affinity system represented by glutamine synthetase, which is found only in perivenous hepatocytes. It makes metabolic sense to get rid of ammonia in the two metabolic processes devoted to lowering ammonia levels in the blood entering the liver, the urea cycle and the synthesis of glutamine,

INTRODUCTION

particularly since CPS1 has developed a relatively high affinity for ammonia when compared with other CPSs.

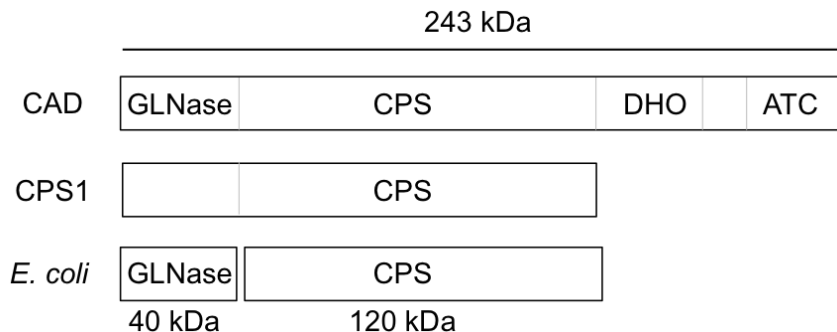


Figure 7. The three basic structural types of CPS. CAD stands for the multifunctional protein that catalyses the three initial steps of pyrimidine synthesis in animals. The abbreviations (in addition to CPS) stand for, GLNase, glutaminase; DHO, dihydroorotase; ATC, aspartate transcarbamylase. Based on a drawing in [11].

Table 1. Differential traits of four classical types of CPS

ENZYME	FUNCTION	SUBSTRATE	ACTIVE WITHOUT ACTIVATOR?	ACTIVATOR	INHIBITOR
<i>E. coli</i>	PYRIMIDINE synthesis ARGININE synthesis	Glutamine (ammonia)	YES	Ornithine IMP	UMP
CPS I	Arginine/UREA synthesis (ureotelic)	Ammonia	NO	N-Acetyl-L-glutamate	---
CPS II	PYRIMIDINE synthesis	Glutamine (ammonia)	YES	PRPP	UTP
CPS III	Arginine/UREA synthesis (ureosmotic)	Glutamine (ammonia)	YES	N-acetyl-L-glutamate	---

1.4 A brief review of knowledge about CPS1 that is relevant for this work.

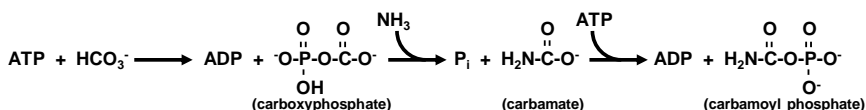
1.4.1 Early findings. The discovery of carbamoyl phosphate synthesis was a consequence of the efforts to place the urea cycle on sound enzymological bases. Krebs and Henseleit reported the cycle on the bases of experiments with liver slices in which the influence of several amino acids on the production of urea was tested [25]. Only after nearly 20 years Grisolia and Cohen [41]

INTRODUCTION

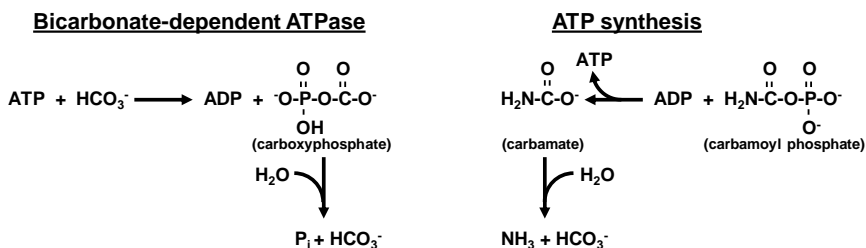
identified the actual enzymatic steps that convert ornithine to citrulline, of which the first was shown to depend on ATP and to produce an unstable compound that was initially called compound X; this compound, when mixed with ornithine and a mitochondrial extract yielded citrulline in an ATP-independent way. Jones and Lipmann showed that compound X is carbamoyl phosphate (CP) [42], and Cohen's group identified NAG as the natural activator of the enzyme catalysing the synthesis of CP [43]. These authors also were the first to purify CPS1, from frog liver [44].

1.4.2 Initial work with pure CPSs. Using the pure frog liver enzyme, the CPS1 reaction was proposed to involve two phosphorylation steps that were reflected in two slow partial reactions, one being a bicarbonate-dependent ATPase and the other being the synthesis of ATP from CP and ADP [45] (Scheme 1).

3-STEP GLOBAL REACTION



PARTIAL REACTIONS



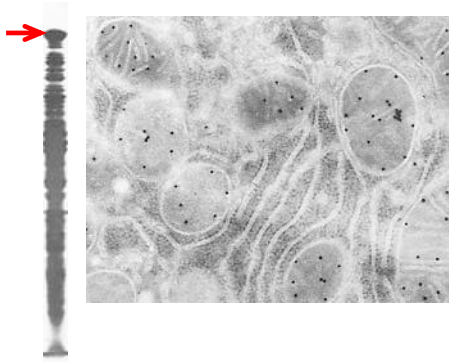
Scheme 1. Reactional steps of the CPS1 reaction (top) and partial CPS reactions (bottom). The partial reactions are aligned vertically with the reactional steps they are believed to reflect. Taken from [46].

Little later it was discovered that bacteria also use CPS for arginine and pyrimidine synthesis [47], which was the starting point for extensive studies with the bacterial enzyme, mainly led by the group of Alton Meister. For brevity, these early data on the bacterial enzyme will not be revised here unless they are directly relevant to the present work, sending the reader to a published review on bacterial CPS [6]. A procedure was soon reported for purification of rat liver CPS1 and it was proposed that NAG activation was allosteric [39],

INTRODUCTION

whereas the kinetic mechanism of the enzyme from beef liver was clarified [48,49] after the appropriate equations were developed for such a complex reaction [50]. In the middle seventies the very high abundance of CPS1 in liver mitochondria (10-25% of the mitochondrial matrix protein) was discovered [51] (Fig. 8).

Figure 8. Abundance, mass and mitochondrial localization of CPS1. Left, Coomassie staining of SDS-PAGE of mouse liver mitochondria. The arrow signals the highly abundant 160-kDa CPS1 band. Right, rat liver electron micrograph, with CPS1 immunogold labeling (black spots), revealing the mitochondrial localization of the enzyme. SDS-PAGE courtesy of A. García-España and V. Rubio. Electron micrograph courtesy of Dr. Erwin Knecht (Instituto de Investigaciones Citológicas, now Centro de Investigación Príncipe Felipe, Valencia).



1.4.3 Reactional mechanism. The groups where I have carried out my PhD work have contributed much of the significant information about CPS1 mechanism. Studies by one of my supervisors showed that the frog liver enzyme makes an unstable intermediate, called "active CO₂" [52], soon shown demonstrated also with mammalian CPS1 [53,54], believed to be carboxyphosphate [55], while Meister's group showed simultaneously the formation of carboxyphosphate with the *E. coli* enzyme [56]. These two groups proved (in the case of Dr. Rubio's group with CPS1) the existence of two separate sites for ATP, one for bicarbonate phosphorylation and the other for carbamate phosphorylation [52-54,57]. Our laboratory showed that, in the absence of ammonia, "active CO₂" (carboxyphosphate) was formed intraenzymatically and that all the substrates of the reaction were bound to the enzyme, with no product release prior to the addition of ammonia [52-55]. The properties of the two sites for ATP were characterized [53-55] (Fig. 9). The existence of an intraenzymatic reversible reaction between ATP and bicarbonate that was much faster than the ATPase partial reaction was proven by positional isotopic exchange techniques with the rat liver [55] and the bacterial enzymes [59,60].

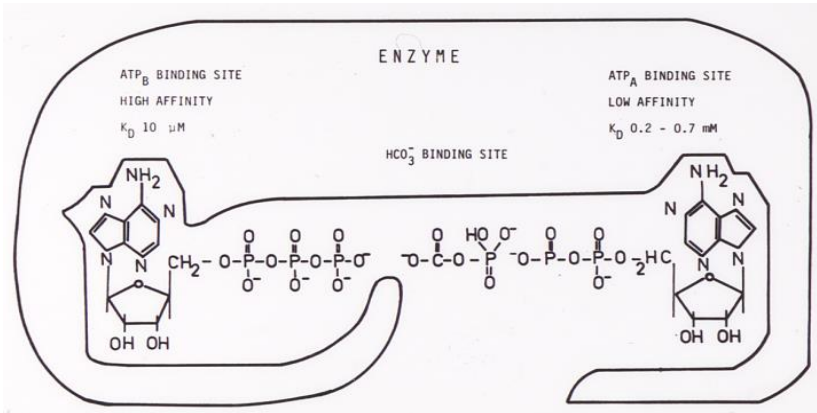


Figure 9. Scheme of CPS1 charged with all its substrates except ammonia, summarizing the data derived from Dr. Rubio's laboratory in the early 1980s. No data existed at the time on any other physical property of the enzyme except that it was a single polypeptide of 160 kDa. Figure reproduced from [58].

1.4.4 Human CPS1. Studies of an American group and of my own group [15,16] resulted in the purification and characterization of human CPS1, providing the only existing detailed enzymological data on the human enzyme until the present work was carried out. These data are summarized in Chapter 1 of the Results, and, therefore, we will not deal anymore with them here.

1.4.5 NAG activation and domain structure. Studies from my laboratories are also responsible for most of the existing knowledge about NAG activation of CPS1. Very early on they clarified NAG binding, showing the presence of a single NAG site per CPS1 molecule and proving the existence of cross-talk between the ATP site that is used to phosphorylate bicarbonate and the NAG site [61]. They proved that NAG can be partially replaced as an activator by a number of chemically unrelated cryoprotectant agents such as glycerol or sucrose [62,63]. They also showed that CPS1 has trace activity in the absence of NAG, which is of very low magnitude because of extremely large K_m values for ATP and for the two ionic activators of the enzyme, K^+ and Mg^{2+} [40]. On the basis of these data the proposal [39] that NAG is an allosteric activator of CPS1 was confirmed [40]. Evidence for NAG-triggered conformational changes was obtained by limited proteolysis studies [46]. A kinetic mechanism for NAG activation was proposed [63].

Studies with the isolated small and large subunits of bacterial CPS had shown that the small subunit binds and cleaves the glutamine used in its reaction and that the large subunit carries out the entire reaction from ammonia and binds the

INTRODUCTION

effectors [6]. By then, limited proteolysis results of my group [64] and of another three groups [65-67] revealed a domain structure of CPS1 that is reflected in Fig. 1 of Chapter 1 of the Results. The first fine localization of a function in any CPS was provided by our laboratory when it showed by photoaffinity labeling of rat CPS1, using N-chloroacetyl-L-[¹⁴C]glutamate and subsequent limited proteolysis cleavage, that NAG binds in the C-terminal domain of CPS1 (a domain of ~20 kDa) [17]. Our group also proved that this domain of *E. coli* CPS is used for binding of all its allosteric effectors [68-70], an information that found correspondence in similar results for the pyrimidine-specific CPS from mammals (CPSII) [71]. Since then the C-terminal domain of ~20 kDa of any CPS is known as the allosteric domain.

1.4.6 Cloning and sequencing of CPS1 cDNA and gross enzyme architecture. Lusty's group cloning and sequencing in the 1980's of the cDNA for *E. coli* CPS and for rat CPS1 [72-74] opened a new era in CPS studies. The sequence revealed that the large subunit of bacterial CPS, of ~120 kDa, exhibits internal homology indicating that its origin involved gene duplication and tandem fusion events [72]. This allowed mapping of some functions in different domains. Using these data, the limited proteolysis data mentioned above, differential scanning calorimetry results [69] and other data including oxidative cleavage [75] and in particular cleavage with FeATP [76], our group proposed [68] a model for the large subunit of *E. coli* CPS and for the equivalent region of any other CPS (including CPS1) in which (Fig. 10) this region was folded as a pseudohomodimer of two homologous phosphorylation domains of ~40 kDa, the more N-terminal involved in HCO₃⁻ phosphorylation and the more C-terminal in the phosphorylation of carbamate [1-3]. Given the reactional mechanism, it was inferred that the carboxyphosphate and the carbamate

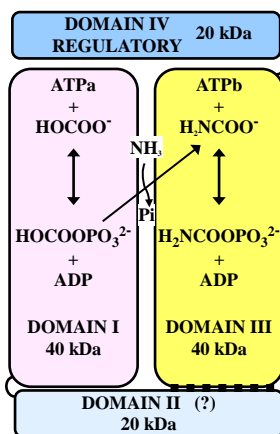


Figure 10. Structural model prevailing in the 1990s for the 120-kDa moiety of any CPS, folded as a pseudohomodimer of its two phosphorylation domains (N-terminal, pinkish, bicarbonate phosphorylation; more C-terminal, yellow, carbamate phosphorylation), with passage of unstable intermediates (carboxyphosphate or carbamate) from one to the other, in a water-shielded way. Darker blue, C-terminal allosteric or regulatory domain, which hosts the sites for the allosteric effectors and which is the subject of Chapter 3 of the Results. Light blue, intervening domain of previously unknown function which is the subject of Chapter 2 of the Results and that we now call Integrating domain.

INTRODUCTION

intermediates had to be shielded from water, and, therefore, that the two phosphorylation centres had to be connected intraenzymatically, so that either the carboxyphosphate or the carbamate would be able to migrate from one to the other phosphorylation centre (Fig. 10).

1.4.7 Crystal structure of E. coli CPS. The next large breakthrough in the CPS field was the determination, 18 years ago, of the crystal structure of *E. coli* CPS [77,78] (Fig. 11). This is the only CPS for which the structure has been determined. Our group got crystals of frog liver CPS [79] but these have never provided good-quality diffraction data. Up to now, ten *E. coli* CPS crystal structures of wild-type or mutant forms, in complexes with substrates and allosteric effectors, have been reported [Protein data bank (PDB, <http://www.rcsb.org/pdb/home/home.do>) accession numbers: 1T36, 1M6V, 1KEE, 1C30, 1CS0, 1CE8, 1BXR, 1A9X and 1JDB). All these structures contained the activator ornithine bound to them, and therefore they

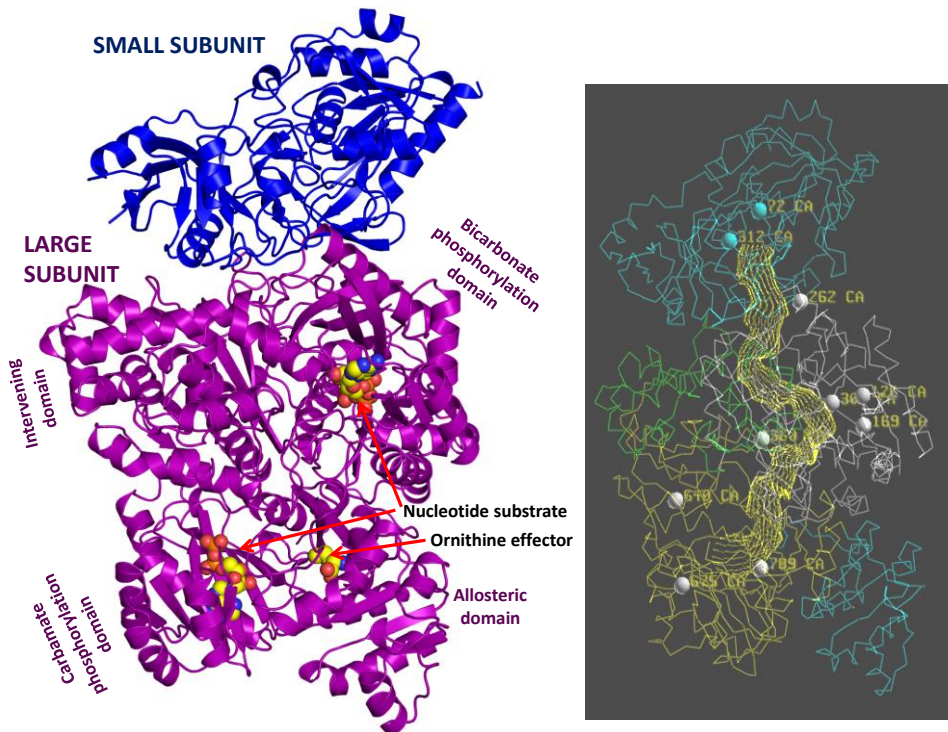


Figure 11. *E. coli* carbamoyl phosphate synthetase structure. Left, detailed structure based on PDB 1JDB. Right, tunnel spanning the whole enzyme as drawn in [85], modified to locate some clinical mutations studied in [87] (represented as spheres).

INTRODUCTION

most likely correspond exclusively to active CPS forms, providing no insight on how a CPS is activated by its effectors. Despite the evolutionary distance between *E.coli* CPS and CPS1, these two types of CPS share ~40% sequence identity over their entire lengths [74], although in CPS1 the 40-kDa and the 120 kDa moieties that correspond to the small and large subunits of *E. coli* CPS are fused into a single polypeptide. Therefore, much information from the *E. coli* CPS is of application to CPS1, and thus the main traits of this structure will be briefly summarized here.

The structure (Fig. 11, left panel) [77,78] confirmed the domain composition that had been proposed for CPS1 as well as the view that the 120-kDa moiety is folded as a pseudohomodimer [1]. Each phosphorylation domain presented the ATP grasp fold that had been identified in glutathione synthetase [80], D-Ala-D-Ala ligase [81] and biotin carboxylase [82], a fold that is composed of three subdomains. Indeed, the connection between CPS and the biotin carboxylase component of biotin-dependent carboxylases such as acetyl-CoA carboxylase had been established already by our group, which had shown that biotin carboxylase makes the same carboxyphosphate intermediate that is made by CPS [83,84]. Also as predicted by our group, the transfer of intermediates between both phosphorylation centres in the two phosphorylation domains of ~40 kDa appeared to occur via a water-shielded intramolecular tunnel [85] (Fig. 11, right panel). Actually, since in the case of bacterial CPS glutamine is the source of ammonia, the tunnel extended nearly 100 Å, from the glutaminase domain of the small subunit to the bicarbonate phosphorylation centre and from there to the carbamate phosphorylation domain at the opposite end of the large subunit (Fig. 11, right panel). It was inferred that the ammonia generated from the glutamine travels to the bicarbonate phosphorylation site via the first half-tunnel, to react with the carboxyphosphate that is waiting for it, making the carbamate, that then travels to the second active site, where it is phosphorylated by a second MgATP molecule, giving carbamoyl phosphate [86].

Since our group had found that ADP and phosphate are only released when ammonia has reacted [52-54], the two phosphorylation sites must have some type of interconnection that permits release of their product cargo only when carbamoyl phosphate is released. Such coordination was illustrated best by a clinical mutant of the carbamate phosphorylation domain that was found in a CPS1 deficiency patient and that when introduced at the equivalent conserved residue in *E.coli* CPS abolished reaction at both phosphorylation sites [87]. Such mutant appeared to hamper [87] the opening of the B subdomain [88] of the carbamate phosphorylation domain, but, interestingly, it also appeared to prevent the opening of the B subdomain of the bicarbonate phosphorylation domain, indicating that this opening is concerted by as yet unknown

INTRODUCTION

mechanisms. Our present work reported in Chapter 2 of the Results suggests that this concerted opening is mediated by the intervening domain that links both phosphorylation domains (called in this work the Integrating domain).

1.4.8 Crystallographic data on the structure of the C-terminal domain of human CPS1, and NAG site delineation. Comparison of the sequences of *E. coli* CPS with those of CPS1 rendered dubious that the allosteric domain of both CPSs were structurally similar, because they exhibited in this region quite low sequence identity [89,90]. Nevertheless, in agreement with the studies of our group [68-70,89,90], the allosteric effectors of *E. coli* CPS, IMP and UMP, were found in the crystal structures of *E. coli* CPS to bind to the allosteric domain [91,92], and ornithine was also found to bind there, although at the boundary with the carbamate phosphorylation domain [77,78].

In 2007 the crystal structure of the allosteric domain of human CPS1 lacking its 22 more C-terminal residues, produced *in vitro* in the absence of the remainder of the protein, was deposited in the PDB as entry 2YVQ, without comments or associated publications, by a structural genomics consortium [93]. This structure (Fig. 12) proved that the allosteric domain of CPS1 has nearly an identical fold to that of the corresponding domain of *E. coli* CPS. Our group analyzed this structure and identified within it the NAG site [13].

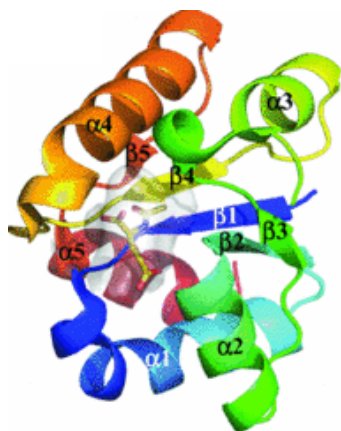


Figure 12. Structure of the C-terminal domain of human CPS1 (PDB entry 2YVQ), coloured to enhance visibility of the structure. Secondary elements are labelled. The pocket identified in [13] is shown in semi-transparent surface representation. A modeled NAG molecule bound to this pocket [13] is shown in sticks representation. Taken from [13].

This domain (Fig. 12) is built up by a central sheet of five parallel β -strands surrounded by a bundle of three α -helices on one side and a bundle of two α -helices on the other side, forming the Rossmann fold seen in proteins of the methylglyoxal synthetase family [94]. The site identified in the flexible docking studies of my group within this crystal structure fitted the earlier results of the group concerning the binding of NAG analogues and the specificity of the enzyme for NAG [95], and they also fitted the results of the group's

INTRODUCTION

photoaffinity labelling studies concerning the site of covalent anchoring of the photoactivated NAG analogue [13]. Furthermore, when these studies were carried out, the group had set up already the baculovirus/insect cell system for rat CPS1 expression, and, therefore, five putative NAG site residues were replaced by other residues, confirming the location of this site [13].

The NAG site modeled in this crystal structure (which had no bound ligands) is a pocket located between the central β -sheet and two α -helices, and appeared to be covered by a three-residue closing lid. The location of the NAG-binding site is equivalent to that of the *E.coli* CPS activator IMP [13] (Fig. 13), indicating a common origin for these sites. Many of the residues which build up the proposed NAG-binding pocket, are conserved in all NAG-regulated enzymes [13].

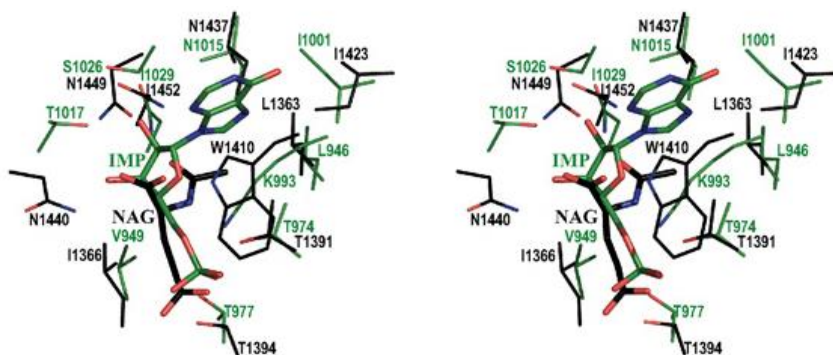


Figure 13. Stereo view of the superimposed side chains (thinner sticks) of the residues making the IMP site (green carbons) and the proposed NAG site (black carbons) in *E. coli* CPS (PDB 1CE8) and human CPS1 (PDB 2YVQ), respectively, with the bound IMP and NAG molecules shown with green and black carbons respectively. Taken from [13].

Chapter 3 of the Results represents a further development of the study of NAG binding to this domain, providing strong evidence that the NAG site shifts conformation as NAG binds to it, this likely being the initial signal for CPS1 activation. Elements of the path for cross-talk to the catalytic domains are also identified in this third part of our study.

1.4.9 Modelling of the CPS1 structure. The finding [13] that the structure of the allosteric domain of CPS1 closely resembles the structure of the allosteric domain of *E. coli* CPS despite the low sequence identity between these two domains supported the possibility of modelling with high fidelity the structure of the whole CPS1 molecule on the basis of the structure of *E. coli* CPS,

INTRODUCTION

particularly since all other parts of the enzyme exhibit much higher sequence identity than the allosteric domain [74,96]. The group of one of my PhD advisers obtained such a model [97], which is virtually identical to the model produced with a somewhat different approach in collaborative work of my other adviser [12]. Both 3D models were obtained from the *E.coli* homologue bound to the non-hydrolysable ATP analogue AMPPNP (PDB file: 1BXR). In the present work I will use our internal group model (Fig. 14), using it as the context frame on which to judge and rationalize all structural inferences made here except those concerning the allosteric domain. Nevertheless, it should be remembered that, because of the presence of ornithine bound to the template structure used from *E. coli*, this model should correspond to active CPS1, since it is based on the active *E. coli* structure.

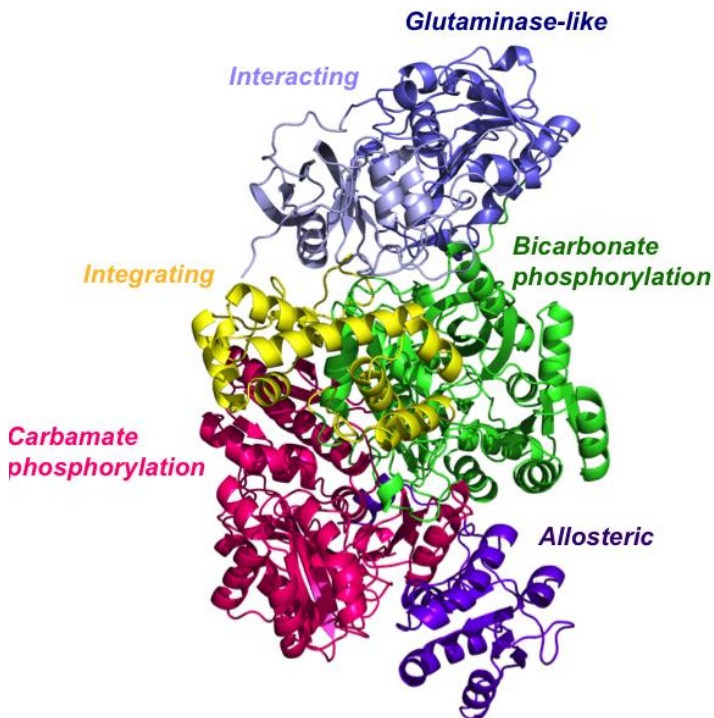


Figure 14. Structural model [97] of CPS1 used throughout the present PhD work.

1.4.10 CPS1 production. Like most mitochondrial proteins that are encoded by nuclear genes [98], CPS1 is synthesized in the cytoplasm as a larger precursor. The pre-CPS1, of 1500 amino acids, contains an N-terminal 39-residue signal peptide that directs the protein for delivery to the mitochondrial matrix, and

INTRODUCTION

which is cleaved off by the matrix-processing peptidase upon translocation [98-100]. Thus far no mutations have been identified in the region encoding this signal peptide [12].

1.5 CPS1 deficiency, the *CPS1* gene and its mutations.

CPS1D was first reported in 1964 [101], in a mild (late onset) case. The report of the first case of complete deficiency (neonatal presentation) took place 10 years later [102]. By 1976 six cases of CPS1D had been identified already [103]. Diagnosis was based on the finding of hyperammonemia, with low citrulline levels and lack of orotate excretion (orotate increases when CP accumulates, Fig. 3, bottom panel) [103]. Confirmatory diagnosis depended on enzyme activity determination in the liver, obtained by biopsy or during necropsy [29,103]. Intestinal biopsy was introduced later [104] as a less invasive means of confirmatory diagnosis. These traits, together with increased blood glutamine levels, have remained hallmarks of this deficiency for many years [105]. CPS1D was soon recognized as a genetic disease of Mendelian autosomal inheritance [106]. Indeed, by 1981, excepting liver transplantation, the main components of present-day treatment of CPS1D had been established already [107], consisting in protein restriction with essential amino acid supplementation, prevention of protein catabolism, arginine or citrulline supplementation, and the administration of benzoate or phenylacetate (nowadays replaced, when administered orally, by phenylbutyrate or by its glycerol triester [30,108]) as alternative pathway medications to bypass the urea cycle.

With the description in 1981 of NAG synthase deficiency (NAGSD) [109], it became clear that the same clinical and biochemical picture was shared by CPS1D and NAGSD, with the only distinguishing element being the results of the assays of enzyme activity in the liver (or in intestinal mucosa in the case of CPS1). The fact that NAGSD can be treated with complete recovery by administering orally N-carbamylglutamate as substitutive therapy for the missing NAG [30,31] has led to the advocacy for investigating a positive response to the administration of carbamylglutamate as a way to differentiate clinically between CPS1D and NAG synthase deficiency [110]. Nevertheless, this may not be a perfect test for distinction between both deficiencies. Some kinetic or low-stability variants of CPS1 might be improved by carbamylglutamate because of pharmacologically-assisted saturation of the NAG site of CPS1 or by pharmacological chaperone effects of this drug (see Chapter 1 of the Results).

INTRODUCTION

The sequence of the cDNA for the human *CPS1* gene became available more than 20 years ago [111]. This opened the way to identification of mutations in the *CPS1* gene, which simplified confirmatory diagnosis of CPS1D and enabled prenatal diagnosis (until then could only possible in foetal liver biopsies[112]). Thus, mutation identification is now considered the gold standard for diagnosis of CPS1D [27]. Nevertheless, the path to widespread genetic diagnosis of CPS1D has been slow, mainly because of the large size of the coding sequence (4503 nucleotides, including the stop codon) and the restricted distribution of abundant gene expression (only liver and intestinal mucosa, with some expression also in the pancreas [51,113]). mRNA-based diagnosis was progressively simplified by the use of illegitimate transcription, first in immortalised patient's leukocytes [114] or in patient's fibroblasts [115], and, more recently on phytohaemagglutinin-stimulated peripheral blood leukocytes [116].

The report of the *CPS1* gene structure [117-119] opened the way to the use of genomic DNA for genetic diagnosis. The gene (OMIM #608307) is localized in the long arm of chromosome 2, at 2q34, spanning 201425 nucleotides, with start and end coordinates within chromosome 2 at 211342405 and 211543830 of the plus strand (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=CPS1>). This gene is composed of 38 exons and 37 intervening introns [117-119]. Therefore, diagnosis using mRNA is considered less laborious than the use of genomic DNA for mutation identification, although this last approach might be alleviated in the near future with the generalization of next generation sequencing techniques.

Now, after 50 years from the original description of CPS1D, it is clear that the disease is quite rare (1:50000-1:300000) [120-122]. The first summary of mutations identified in *CPS1* patients in 24 years was published in 2011 [12]. Together with the mutations reported in two additional publications appearing after that paper [116,123], a total number of 243 CPS1D-associated genetic changes have been identified, generally with little recurrence of the mutations. As shown in Fig. 15, where the percentages of the changes corresponding to each type of mutation are represented, more than half of these mutations are missense changes. Since nonsense changes as well as most small insertions, deletions and splice site aberrations result in enzyme truncation, at least 1/3 of all the mutations should result in the lack of *CPS1* protein in the liver. In fact, this number is even higher because a number of the missense changes also cause enzyme destabilization (see the Results section).

INTRODUCTION

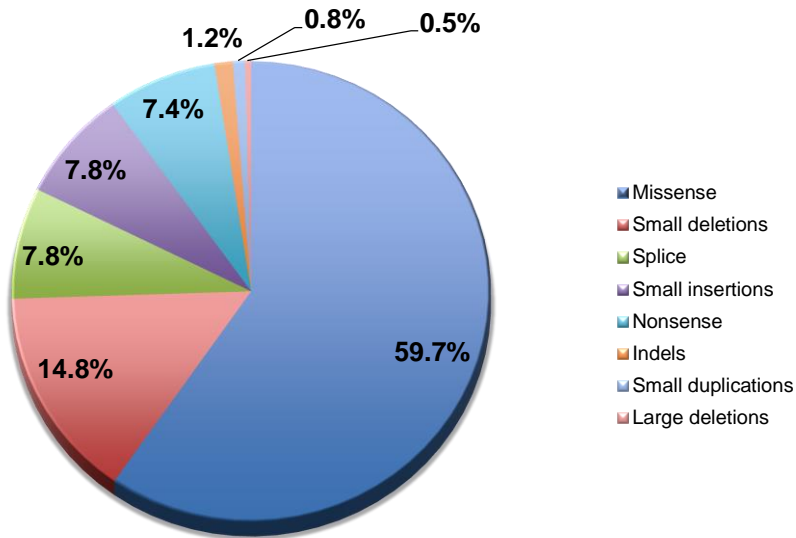


Figure. 15. Summary of all CPS1 mutations reported to date [12,116,123]

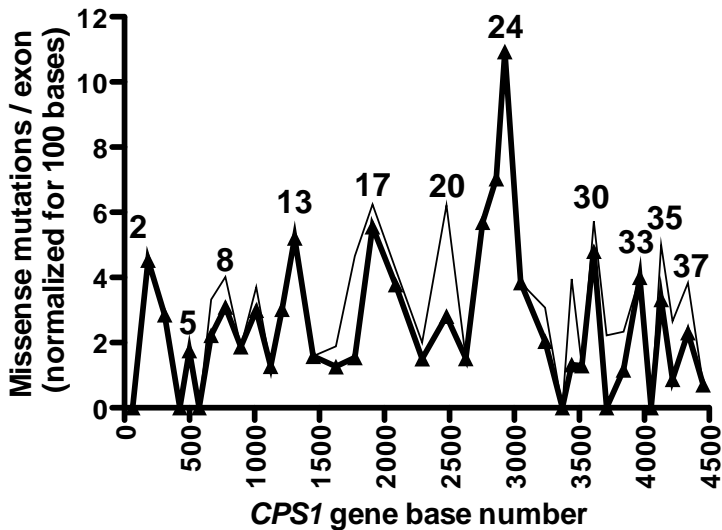


Figure. 16. Density of missense mutations per CPS1 exon. The number of missense mutations per exon, normalized per 100 nucleotides, is plotted along the coding sequence, with the symbols placed at the middle point of each exon (some exon numbers are given). Thin line, all missense mutations. Thick line and large filled triangles, missense mutations that do not fall on CpG dinucleotides. Figure modified from [12] to include the mutations described in [116,123].

INTRODUCTION

The incorporation of the novel missense mutations reported in [116,123] to the mutational database summarized in [12] does not alter the conclusion formulated in [12] that the density of missense mutations along the gene varies with the region. Fig. 16 actualizes Fig. 2B of reference [12], highlighting the fact that the distribution of the mutations is highly skewed, particularly when the contribution of the CpG mutational hotspots is subtracted, supporting the existence of highly eloquent and less eloquent regions in the enzyme.

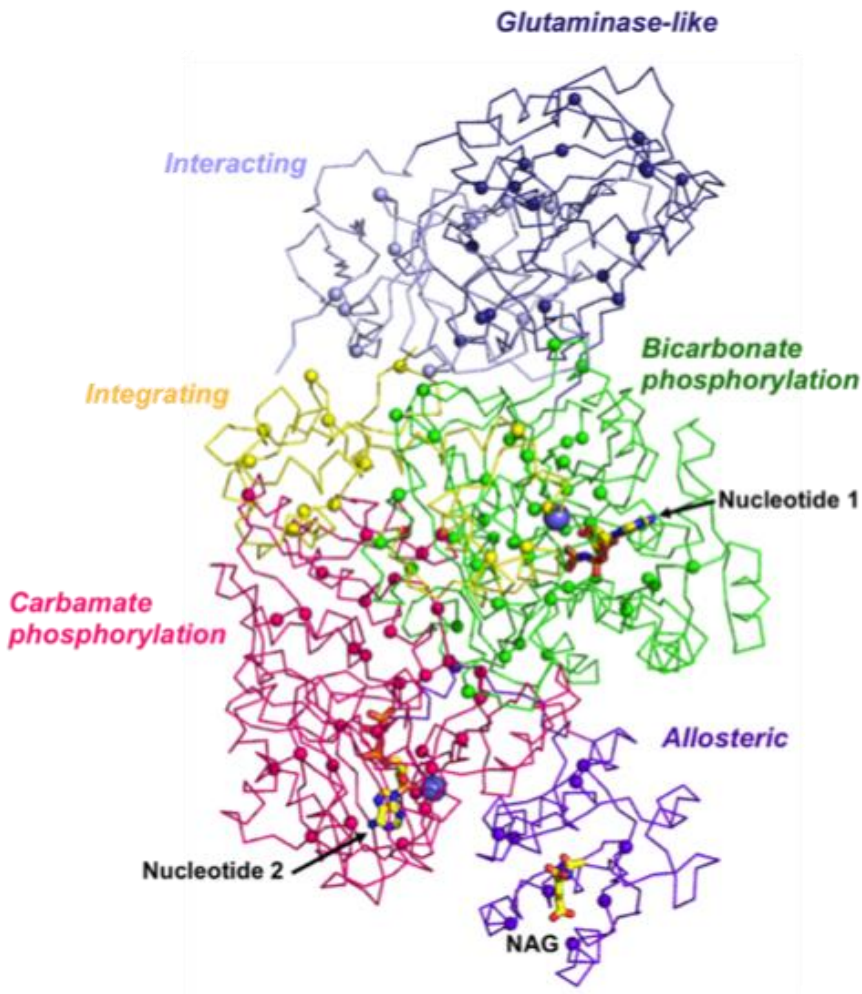


Figure 17. Residues hosting missense mutations localized in the structural model of human CPS1 that has been used here [97]. For clarity, residues have not been identified. They are colored as the domain where they map.

INTRODUCTION

Fig. 17 illustrates the distribution of the missense mutations in the structural model of the protein, giving a 3D-view of the mutations' density in different domains of the enzyme. Although the size of the protein and the large number of mutations makes impossible a very clear view, it is evident, for example, that mutations are rarer towards both ends of the enzyme (in the regions corresponding to the small subunit of *E. coli* CPS and to the allosteric domain) than in the regions containing the catalytic machinery or than in the intervening domain that links both phosphorylation domains. Actually, this skewed distribution of mutations is one of the basis of the present work: to clarify the underlying reasons for the variable eloquence of the various regions and in particular to clarify the reasons for the high eloquence of this intervening domain that, after the present work, is called the Integrating domain (Chapter 2 of the Results).

1.6 Other roles of CPS1 in pathology.

Although not the subject of the present work, it is important to briefly highlight that mutation-triggered CPS1 deficiency is not the only mechanism linking CPS1 to human pathology and that, therefore, knowledge on CPS1 may also have impact on other disease processes.

1.6.1 Secondary hyperammonemia. One of these disease processes are the secondary hyperammonemias (the hyperammonemias in which the operation of the urea cycle is decreased secondarily to other defects external to that cycle [124]) since in many of them impaired CPS1 operation is the reason for the hyperammonemia (Fig. 18). This appears to be the case in the hyperammonemia of carnitine cycle defects or of fatty acid oxidation defects, in which the impaired cellular energy production has been proposed to reduce the availability of ATP for CPS1 [124]; in mitochondrial carbonic anhydrase deficiency [125], in which bicarbonate may not be produced fast enough to feed the cycle at the CPS1 level; and in the hyperammonemias of organic acidemias, or due to valproate or, also, to fatty acid oxidation defects, via a mechanism in which poor NAG production by NAGS may decrease CPS1 activity [124,126]. An important therapeutic consequence of this CPS1-mediated mechanism of the secondary hyperammonemia is that carbamylglutamate may be effective in treating these hyperammonemias [125,127]. It is also to be noted that carnitine administration restored CPS1 expression in an animal model of a carnitine uptake defect (juvenile visceral steatosis) that associates with hyperammonemia because of decreased expression of urea cycle enzymes [128], suggesting an additional mechanism for causation and for treatment of some types of secondary hyperammonemia.

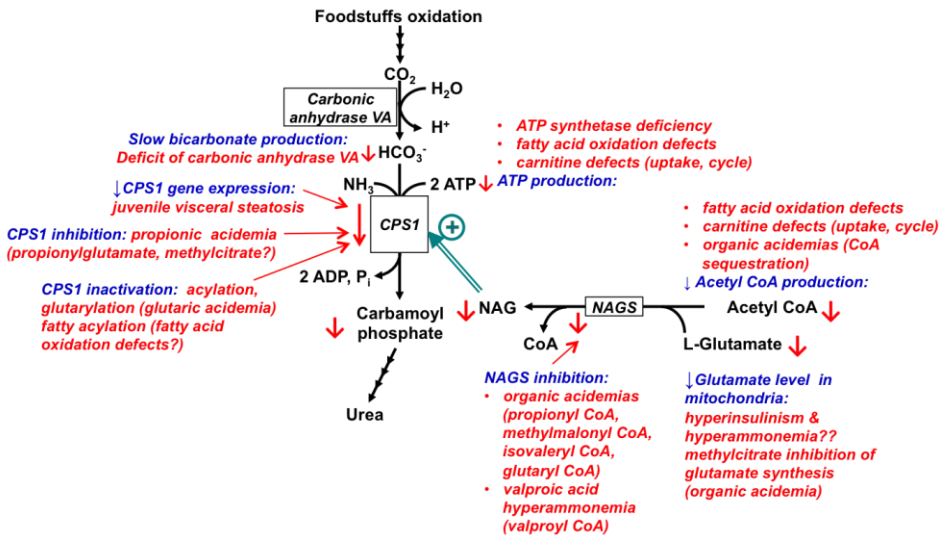


Figure 18. Possible mechanisms of secondary hyperammonemia due to decreased CPS1 activity

A novel aspect recently introduced in the understanding of CPS1 function and pathology is the realization that CPS1 is acylated at several lysines by as yet unclarified mechanisms that possibly depend on the acyl coenzyme A levels, and that it is deacylated by sirtuin 5, with claims and indirect evidences for the control of CPS1 (and the urea cycle) activity by the state of acylation [129,130]. Indeed, fatty acylation of the CPS1 active center had been reported to greatly affect CPS1 activity [131]. Since sirtuins are anti-aging proteins, CPS1 acylation/deacylation was considered a link between nitrogen metabolism and aging [132], whereas in recent work CPS1 glutarylation was considered a key component of the secondary hyperammonemia of glutaric acidemia [133]. Clearly this is a new field that requires further exploration, and that may be addressed experimentally by exploiting our present expression system reported in Chapter 1 of the Results.

1.6.2 A CPS1 polymorphism involved in vascular reactivity and in other traits related to pathologies. Apart from being involved in primary and secondary hyperammonemia, CPS1 has been involved more recently in other pathological processes less clearly connected with deficient activity of the enzyme. In particular, the p.Thr1046 form of the CPS1 non-synonymous polymorphism (polymorphisms are sequence changes which do not cause evident enzyme deficiency) p.Thr1406Asn (designated earlier on as p.Thr1045Asn) was

INTRODUCTION

reported [134] to be associated with lower blood concentrations of citrulline and higher rates of hepatic veno-occlusive disease and death than the Asn1406 variant, in adults given high-dose chemotherapy for bone marrow transplantation. This polymorphism was associated with pulmonary hypertension and decreased plasma arginine and nitric oxide metabolites in neonates with respiratory distress [135], and in those developing pulmonary hypertension after congenital heart surgery [136], leading to trials for using citrulline to prevent or treat these presentations of pulmonary hypertension [137,138].

These vascular reactivity phenotypes were rationalized on the basis of differences in nitric oxide production that were tentatively related to differences in the enzyme activity between the two polymorphs [139], given the fact that arginine is a precursor of nitric oxide. However, the effects on activity of the p.Thr1406Asn substitution appeared not to be drastic, with reports in crude extracts [139] of 20-30% higher activity and with pure recombinant CPS1 of up to 1.7-fold lower activity [140] for the Asp form. An activity effect might underlie the report of increased frequency of the Asp polymorphism in valproate-associated hyperammonemia [141,142], which would be in line with a decreased activity of this form of the enzyme [140]. However, a recent Japanese study [143] did not find an association between the p.Thr1406Asn polymorphism and valproate-induced hyperammonemia, although this last study used a pediatric cohort and defined hyperammonemia as ammonia levels $>200 \mu\text{M}$ [143], whereas in the largest earlier study finding an association the threshold ammonia level was $65 \mu\text{M}$ [142].

Homozygosity for the CPS1 Thr1406 variant was also associated with an increased risk of necrotizing enterocolitis in preterm infants [144]. This disease is also believed to be related to nitric oxide. However, this form of the polymorphism was reported not to be associated with a decreased level of arginine in preterm infants [145], despite the fact that the level of arginine is known to be decreased in preterm infants with necrotizing enterocolitis [144].

With the advent of genome-wide association studies of phenotypic characters with genetic polymorphisms, the p.Thr1406Asp polymorphism has emerged as one of the determinants of the levels of homocysteine [146,147] and of homoarginine [148] in normal subjects, of fibrinogen in women [149], of creatinine [150] and of the response to albuterol in African-American children with severe asthma [151]. The link between these phenotypic traits and CPS1 function is unknown or at best speculative for now.

1.7 Previous *in vitro* expression and structure-function studies on CPSs.

As already indicated, Chapter 1 of the Results is devoted to the "*in vitro*" production and purification of human CPS1. We will discuss here previous attempts of CPS expression. Actually, once having its gene cloned, *E. coli* CPS was easily expressed in *E. coli* cells, having been the subject of extensive site-directed mutagenesis studies (see for example [90,152-154]). Given the existence of reasonable sequence identity between the different CPSs [96], some of these bacterial results have been exploited to interpret the effects of a number of CPS1D-associated missense mutations (see for example [155]). In fact, work of my group [87] utilized the bacterial enzyme as a surrogate model for the human enzyme to explore systematically the impact of clinical mutations found in CPS1D on the activity of the enzyme.

An eukaryotic CPS, the arginine-specific yeast enzyme, was expressed recombinantly in yeast and was utilized for site-directed mutagenesis studies focusing on basic aspects of the enzyme, without making clinical inferences [156,157]. Similarly, CPSII, the CPS component of the multienzymatic protein CAD was expressed long ago in *E. coli* in complete form, as its individual domains, or as hybrid proteins with *E. coli* CPS [71,158-160], and it was used to study the function of domains or regions of this enzyme or even of individual residues [159-160]. These enzymes have not been used as models for understanding CPS1D.

Although Chinese and Korean groups claimed long ago recombinant production of CPS1 in cultured animal cells [161,162], the first well documented example of *in vitro* recombinant CPS1 expression and purification was provided in 2003 by Powers-Lee group [163], for the enzyme for the american bullfrog (*Rana catesbeina*), using *Schizosaccharomyces pombe* as the expression host. Apparently large amounts (15 mg/L of yeast culture) of pure enzyme were obtained. This enzyme was used for site directed mutagenesis studies focusing on the reasons for the ammonia or glutamine specificity of different CPSs, but not to explore CPS1 pathology despite the 74% identity between frog and human CPS1s [163]. This same system was exploited by the same group to generate reasonable amounts (2-3 mg/L of *S. pombe* culture) of pure active human CPS1 [140,165], as reported months before our laboratory published the recombinant production of rat CPS1 in baculovirus/insect cells [13]. The *S. pombe* production system was used to study the impact on enzyme functionality of the p.Thr1406Asn polymorphism discussed in the previous section, as well as to try to clarify the role of the N-terminal region of the enzyme, which was deleted [140]; and to clarify the role of two proximate cysteines classically

INTRODUCTION

considered essential for CPS1 activity [164]. Intriguingly, no further use has been reported of this system to characterize the impact in pathology of the mutations found in CPS1D patients.

As already mentioned, in the context of characterizing the site for NAG by site-directed mutagenesis (an objective that is completed in Chapter 3 of the Results of the present PhD dissertation), our group reported in 2009 [13] the production of the rat liver enzyme in baculovirus/insect cells. This production system was highly effective (10 mg of pure CPS1/L of culture), allowing to prove, by exploiting the high identity of rodent and human CPS1, the pathogenetic potential of clinical mutations found in CPS1D [14]. Since this system is thoroughly discussed in the next section of this introduction and in Chapter 1 of the Results, no further discussion of it will be made here. Indeed, Chapter 1 of the Results reports the production of the genuine recombinant human enzyme in its mature form, and, together with the two chapters following it, describes the use of this system to characterize the effects of clinical and experimental mutations of human CPS1.

A weakness of all these systems mentioned thus far, including the baculovirus/insect cell system, is that they are not able to assess the effects of mutations in the non-coding regions of the CPS1 gene, or of mutations that may affect the splicing of the gene. Summar's group [165] has developed a tool based on the generation of a highly sophisticated bacterial artificial chromosome (BAC) that is devised to express CPS1 in SV40-immortalized lung fibroblasts. This expression system relies heavily on normal BAC replication in bacteria, episomal eukaryotic replication, antibiotic selection in eukaryotic cells, and the expression of green fluorescent protein. Because of the large size of genomic inserts allowed by BACs this system was reported to provide an efficient way to test genetic variants affecting both coding and noncoding sequences, being particularly useful for assaying mutations that affect RNA processing, thus being suitable to examine mutations' effects on the nature or the levels of the mRNA produced.

As a proof of concept, this system was applied [165] to examine the mechanism of four putative splicing mutations (two sitting in exonic sequences and the other two in intronic sequences) found in CPS1D patients, locating experimentally two cryptic splicing sites for the two intronic mutations and revealing the role of nonsense-mediated decay by knocking down this process. The latter agrees with the previous finding of these authors in a cohort of 26 CPS1D patients with 52 mutant alleles, that there was evidence in many cases of nonsense-mediated decay in those patients carrying nonsense or frameshift changes or splicing defects [166].

In summary, we are now endowed with a number of powerful techniques that permit investigating the disease-causing impact and the mechanisms of damage for any mutation affecting CPS1. The present work deals with one such instrument, the use of baculovirus-insect cell expression, to analyze the characteristics of the recombinant human enzyme and to investigate the kinetic and stability derangements introduced by CPS1D-associated mutations. We will deal now with these two experimental approaches utilized here, to provide some extra insight on them, since the treatment of these two techniques in the manuscripts that constitute the three chapters of the Results is brief.

1.8 The baculovirus/insect cell expression system

Baculoviruses are large rod-shaped circular DNA viruses that infect many different species of insects [167]. They can infect mammalian cells but they do not propagate therein [168]. They are used in protein expression studies because they make massive amounts of a single protein (polyhedrin in some types of baculoviruses, granulins in other types of viruses of this family), which is used as a matrix to include therein mature virions [167,169,170]. This is the so-called occlusion-derived virus, which favors virus conservation in the external environment for extended periods, facilitating viral transmission. The larval form of the insect gets infected by eating material such as leaves that are contaminated with the occluded virus within the proteinaceous matrix. The matrix is dissolved in the gut and the virus is released for infection. The infected intestinal cells produce more virions that bud out from the cell, this time without being included in a polyhedrin body, infecting other cells.

Baculovirus-based expression systems rely on the introduction of the gene to be expressed in the baculoviral genome under the control of a strong late viral promoter such as the polyhedrin promoter [171]. Therefore, a recombinant virion has to be built where the gene of interest is incorporated appropriately into the viral genome. Classically, this was attained by lengthy recombination procedures. The recombinant virus was first isolated by a series of plate selection steps, and was then used to infect cells or even susceptible larvae, that were then homogenized and tested for expression of the introduced gene [172].

Luckow *et al.* [173] developed a faster system for generating recombinant viruses. This system, sold as the Bac-to-Bac commercial system [174], has been used here. It takes advantage of site-specific transposition by the Tn7 transposon [175], an unusual mobile DNA segment that inserts itself at high-frequency into a single specific site in the bacterial chromosome. In the absence

INTRODUCTION

of this specific site, called *attTn7* in *E. coli*, Tn7 transposes at low-frequency into many different sites. The presence of the *attTn7* site endows the transposition reaction with specificity and high frequency of transposition.

The Bac-to-Bac system uses the *Autographa californica* (a moth) nuclear polyhedrosis virus (AcNPV) [176]. The heterologous gene is cloned into the pFastBac™ (Invitrogen, Carlsbad, CA) plasmid [177], which, in its pFastBacHT version (Fig. 19, top right), has at the cloning site a polyhedrin promoter (P_{PH}) and introduces a N-terminal His₆ tag connected to the cloned sequence by a TEV protease cleavage site. This expression cassette ends with a SV40 polyadenylation signal to terminate transcription and for adding a polyA tail. The left and right arms of the Tn7 transposon flank the cassette. pFastBacHT also includes ampicillin and gentamicin resistance genes.

An essential second element in this expression system is the DH10Bac™ *E. coli* strain (Invitrogen Carlsbad, CA) (Fig. 20), which hosts a baculovirus shuttle vector (called bacmid bMON14272) with an *attTn7* target site where Tn7 will be inserted from pFastBac. The bacmid [173,177] hosts a mini-F replicon, which is a fragment of the naturally occurring F (fertility or sexual) factor of *E. coli*, that carries all the genes and sites required for replicon maintenance and control. This enables the complete baculovirus genome to replicate in *E. coli* as a large and stable low-copy number plasmid. This bacmid includes a kanamycin resistance marker for antibiotic selection. Also for selection purposes, the *attTn7* site is located within the coding sequence for the LacZ α peptide, which, when expressed, provides the missing part of LacZ encoded in the bacterial chromosome under control of the lac promoter, resulting in blue colonies upon IPTG induction if the chromogenic LacZ substrate X-gal is present. However, transfer of Tn7 to the *attTn7* site disrupts the coding sequence for the LacZ α peptide, and the corresponding colonies, although kanamycin resistant, are white on agar containing IPTG and X-gal.

Recombinant bacmids are generated by transposing the Tn7 element from a pFastBac™ donor plasmid to the *attTn7* attachment site on the bacmid. The enzyme transposase, required to effect this transfer, is provided in this system by another helper plasmid, pMON7124 (Fig. 20). This helper plasmid is carried also in the DH10Bac™ strain, providing in trans the Tn7 transposition function [178], as well as conferring tetracyclin resistance, thus endowing the cells carrying this plasmid with an extra antibiotic selection trait.

The process used for preparation of recombinant CPS1-carrying pFastBac plasmid is schematized in Fig. 19 and is summarized in the next paragraphs.

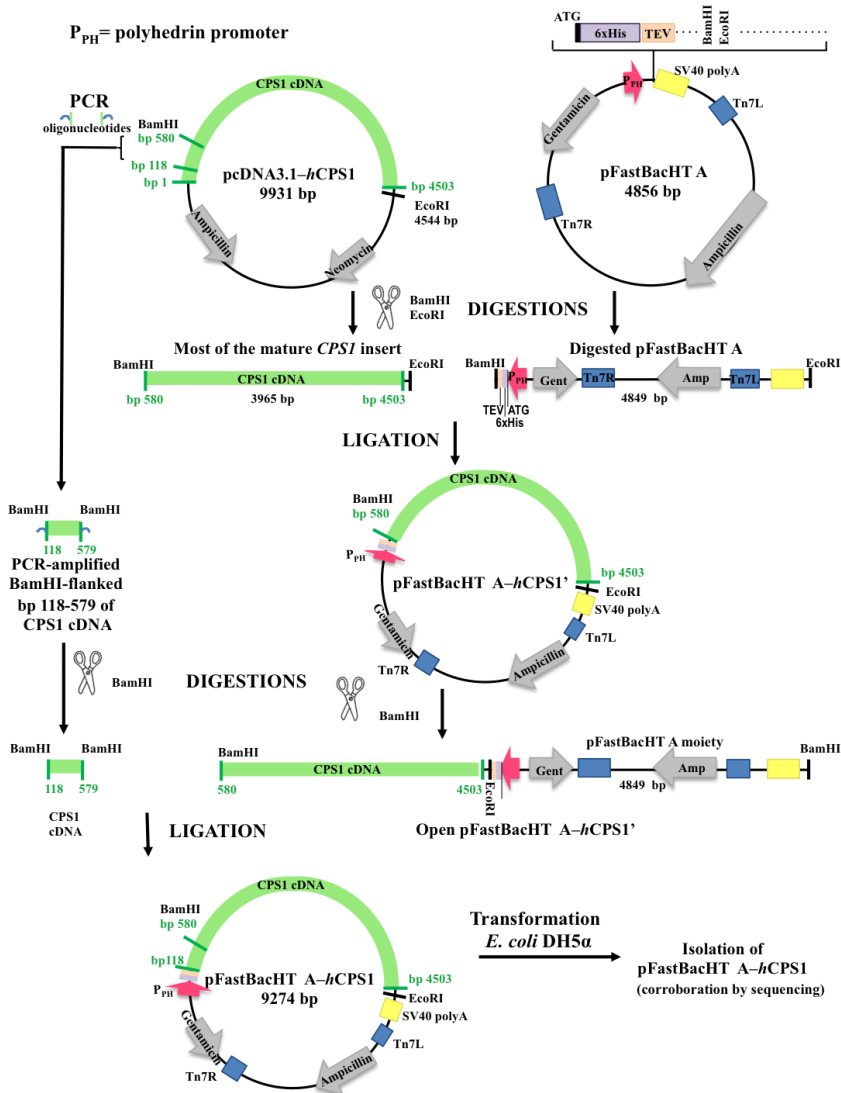


Figure 19. *CPS1* cloning in pFastBacHT A. The *CPS1* insert is in green and the bp numbering of it use as 1 the A of the initial ATG of the pre-*CPS1* coding sequence. Only the restriction sites used here are indicated. The same symbols, acronyms and color code are used throughout the figure: P_{PH}, polyhedrin promoter (pink arrow); 6xHis, 6 histidine tag (purplish box); TEV, Tobacco Etch Virus protease cleavage site (light orange box); SV40 polyA, Simian virus 40 polyadenylation signal (yellow box); Tn7L and Tn7R, left and right arms of the Tn7 transposon (blue boxes). Digestions are represented by scissors. Antibiotic resistance genes are colored grey.

INTRODUCTION

The cDNA for mature human CPS1, carried in a plasmid provided by Dr. Marshall Summar (plasmid pcDNA3.1-*hCPS1*; a derivative of the pcDNA3.1 plasmid from Invitrogen; see Chapter 1 of the Results), was introduced in a two-step process, from base 118 to the end of the ORF, in the cloning region of a variant of pFastBacHT, called pFastBacHT A [174], between the *Bam*HI and *Eco*RI sites of this cloning region. The A variant of the plasmid was used because it was the one that allowed *CPS1* gene insertion in frame (there are A, B and C versions, with B and C having one and two extra G bases, respectively, at the cloning site). This *CPS1* cDNA-bearing vector was amplified and isolated from suitable *E. coli* cells (DH5 α) by exploiting the ampicillin resistance provided by the plasmid. The presence of the insert and the correctness of the construction were checked by sequencing.

Similarly, the parental bacmid was propagated in *E. coli* DHBac10 cells as a large plasmid that confers resistance to kanamycin and can complement the *lacZ* deletion present on the chromosome to form colonies that are blue (*Lac*⁺) in the presence of a chromogenic substrate such as X-gal and the inducer IPTG.

Then (Fig. 20), the isolated plasmid carrying the cDNA for mature human *CPS1*, pFastBacHT A-*hCPS1*, was used for heat-shock transformation of *E. coli* DHBac10 cells. Transformed cells in which the Tn7 cassette had been transferred to the *att*Tn7 site were selected by the white color of the colonies on LB agar having IPTG and X-gal. Selection also included gentamicin, kanamycin and tetracyclin, to ensure the co-presence of pFastBacHT, the bacmid and the helper plasmid. Individual white colonies were inoculated in liquid LB medium with these three antibiotics. After overnight growth, the bacmid was isolated from the pelleted cells by a modified [174] alkaline-SDS miniprep procedure [179] that avoids shearing forces to prevent cleavage of the large circular DNA. The presence of the insert in the isolated bacmid was checked by PCR, corroborating the expected size of the PCR product by agarose electrophoresis.

Although production of the bacmid is an important step in the process towards protein expression, it is just a prerequisite that is still far from the final goal, the production of the protein product of interest. The path from bacmid to protein requires the conversion of the bacmid into a real virus provided with its capsid and being capable of infecting cells [173,174]. This is accomplished (Fig. 21) by introducing the bacmid via transfection into susceptible cells derived from an insect that is sensitive to the virus. Then the cells will produce virus particles hosting the gene for the desired protein under the control of the potent promoter for the late-expressed (in the viral life cycle) polyhedrin protein. In this way, the protein of interest will be produced.

INTRODUCTION

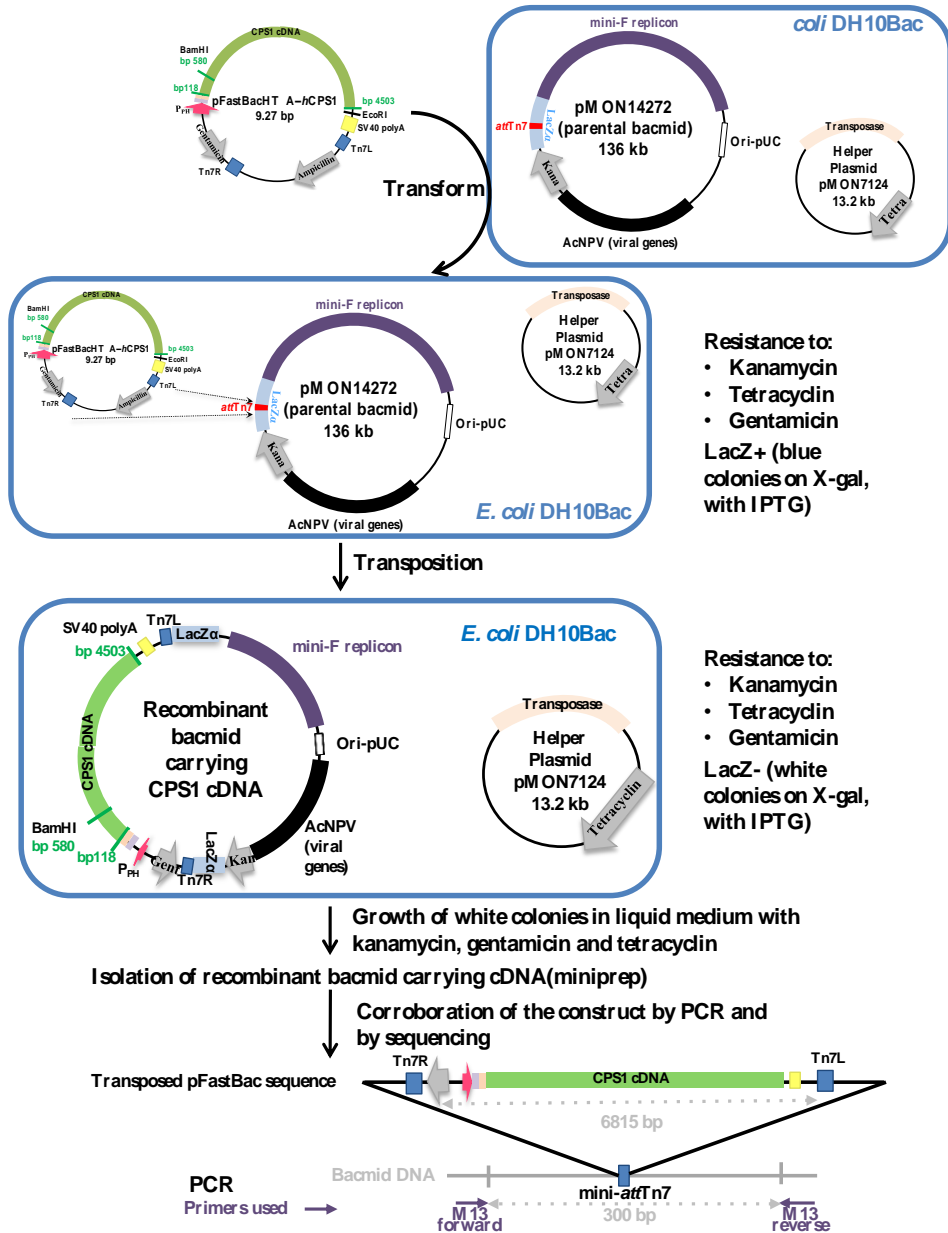


Figure 20. Recombinant bacmid generation. *bMON14272* elements: *attTn7*, target site of the *Tn7* transposon; *mini-F* replicon, fragment of the *E. coli* *F* factor that enables *E. coli* low-copy number replication; *AcNPV*, *Autographa californica* nuclear polyhedrosis virus genes; *LacZ α*, gene for α peptide of *LacZ*; *Ori-pUC*, origin of replication of the *pUC* plasmid; *Kana*, Kanamycin. In *pMON7124*, *Tetra*, tetracycline resistance gene; *Transposase*, genes for the enzymatic machinery for *Tn7* transposition.

INTRODUCTION

In practice, the process from bacmid to protein expression (Fig. 21) requires transfection optimization, the use of cells that grow fast and allow rapid viral production, and the production of large viral numbers for generating important amounts of the desired protein in relatively large volumes of cell cultures with high cell concentrations. The use of continuous insect cell lines that can be grown in medium without serum to improve transfection, the use of several culture rounds for escalation in the total number of virus particles, the titration of the virus produced (in our case using a fluorescent assay [180] rather than the classical plate assay [181]), and the optimization of the multiplicity of viral infection (number of virus particles used per cell to be infected) and of time of harvesting, are considered key factors for successful protein expression [174].

We have used in all the steps in which animal cells are utilized (Fig. 21) a commercial cell line from the moth *Spodoptera frugiperda* [182] called Sf9 cells, which are round cells that grow well in plate and in suspension. For further methodological detail I will refer here to the instruction booklet [174] of the Bac-to-Bac system. Nevertheless, I would like to underline the need to use at the transfection step a medium devised for insect cells in the absence of proteins (Grace Medium unsupplemented, protein or protein hydrolysate-free) [183] and the utilization as transfecting agent the cationic liposome/lipocomplex-forming (with DNA) commercial Cellfectin II (Life Technologies) [184]. In all other culture steps a protein-containing or serum-supplemented Grace medium can be utilized, although we have used a proprietary (from Gibco-Invitrogen) protein-free and serum-free medium (SF-900 II SFM [185]), to lower the protein contamination in the final CPS1 purification step. The use of protein-free medium made essential the addition of an anti-shearing polymeric compound, Pluronic F68 (a copolymer of polyethylene glycol and polypropylene glycol with an average mass of ~8 kDa; obtained from Sigma) to prevent cell breakage because of the orbital shaking. Nevertheless, shaking was moderate (120 rpm).

Although secreted proteins expressed in the baculovirus/insect cell expression system can be recovered from the medium, this was not the case with CPS1, and therefore in the final step (Fig. 21) the cells were separated from the medium by centrifugation, and the protein was extracted from the cell pellet by gentle breaking of the cells using freezing and thawing in an appropriate buffer.

A property of the presently used baculovirus system is that the initial pFastBac plasmid containing the insert encoding the protein can be subjected to standard site-directed mutagenesis procedures. The use of site-directed mutagenesis has been crucial in the studies reported in this PhD dissertation, and thus we will deal with it in the next section of this Introduction.

INTRODUCTION

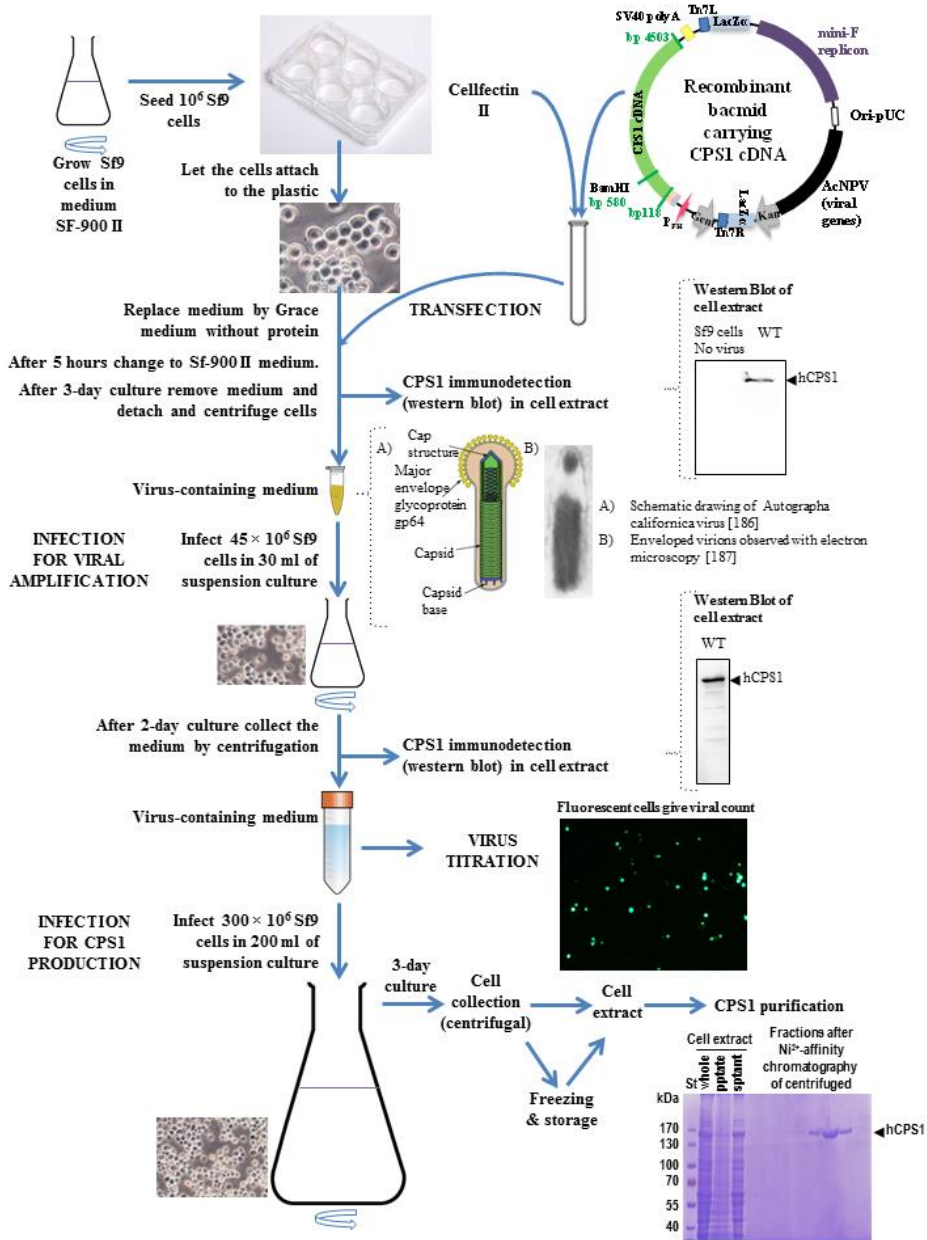


Figure 21. CPS1 production. Western blots of cell extracts (~20 μ g protein per well) were revealed with an anti-rat CPS1 rabbit antiserum (provided by our laboratory) followed by luminiscent detection with the anti-rabbit ECL system (GE-Helathcare). Coomassie-stained SDS-PAGE was 8% polyacrylamide. St, protein markers.

INTRODUCTION

1.9 Site-directed mutagenesis

Site-directed mutagenesis allows the replacement at will of a given residue of a protein by any other residue. In this way the role of particular residues and the effects of amino acid changes can be assessed [188]. This approach, heralded in the late seventies with phage DNA [189], was popularized in the eighties through the utilization of several alternative principles [190], including phosphorothioate-based approaches [191], as well as "cassette mutagenesis", in which a DNA fragment with the desired mutated sequence was introduced by restriction enzyme-mediated digestion and ligation of the mutated cassette [192]. The advent of the polymerase chain reaction (PCR) [193] led to the introduction of much easier PCR-mediated mutagenesis techniques (too numerous for discussion here) which use synthetic oligonucleotide primers in which the gene sequence has been modified to introduce the desired mutation [194].

A potential source of problems of the PCR techniques is the incomplete fidelity of the polymerase used [195], leading to the possibility of introducing undesired mutations. Although this is not generally a problem in applications of PCR in which the bulk product is sequenced, such as the analysis of mutations in patient DNA, it is a potential problem in those applications in which a single DNA molecule is cloned, such as in PCR-mediated site-directed mutagenesis.

When using a low-fidelity thermostable enzyme, such as the widely utilized Taq polymerase (which is fast, highly processive although not optimally thermostable, $t_{1/2}$ at 95°C, 40 min) [196] generally several transformed bacterial colonies are selected, sequencing the isolated plasmids, which generally gives at least one cloned mutant DNA without unwanted mutations. Alternatively, a thermostable polymerase having high fidelity (due to the exhibition of proofreading exonuclease activity) can be used [195,196] as in the Quickchange Site-directed Mutagenesis Kit (Stratagene, La Jolla, CA) used here [197]. This kit was chosen because it is easy to use, it is not labor-intensive and it can use any double-stranded plasmid. Furthermore, the rather simple mutagenesis protocol used with this kit has a reputation of high efficiency.

This kit uses a highly elegant approach to address the problem of selecting for the mutant form and against the wild-type gene, by exploiting the selectivity of the frequent-cutter (four nucleotide target site) restriction enzyme *DpnI* for a target DNA sequence with a methylated adenine at the cutting site [198]. Since PCR does not introduce methylations in the synthesized DNA, this allows discrimination between the cell-generated template wild-type DNA and the PCR-generated mutated DNA, by incorporating a *DpnI* digestion step.

INTRODUCTION

The rationale behind mutagenesis with this kit is as follows (Fig. 22). Two mutagenic forward and reverse oligonucleotides, which are complementary between themselves and to the two strands of the gene in the region to be mutated (Fig. 22, top) are synthesized and used. They are mutagenic because they host the required DNA changes (generally not exceeding three bases for simple amino acid replacements). After a high temperature denaturation of the plasmid double helix, these primers are hybridized to the plasmid strands by lowering the temperature, and then the two strands of the whole circular plasmid including the cDNA that encodes the protein to be mutated are copied by a high fidelity DNA-directed DNA polymerase that can copy processively and highly faithfully relatively large templates. This last goal is attained in this kit by using an enzyme mixture called Pfu turbo, which includes *Pyrochoccus furiosus* (an organism that lives at nearly 100°C) DNA polymerase [199] and a thermostable dUTPase added to prevent poisoning of the polymerase reaction by dUTP, thus improving PCR performance [200]. The time allocated for elongation should be sufficient to complete the copy of the entire plasmid. The entire cycle is repeated again 12-18 cycles (a small enough number to minimize the probability of introducing unwanted mutations), the mixture is left to anneal, *DpnI* is added to digest the methylated parental DNA, and, after completing the digestion, the DNA is used to transform any suitable cell type such as DH5 α cells, which will seal the staggered nicks in the mutated plasmid.

For application to the mutagenesis of CPS1 expressed in insect cells, we used as template for mutagenesis the pFastBac-CPS1 vector (see above and Fig. 22), consisting of 9242 bp (including the 4386 bp of the CPS1 insert), and the mutagenic primers indicated in the individual chapters of the Results section. Instead of using the complete commercial kit, we only utilized the proprietary *Pfu* turbo DNA polymerase/dUTPase mix, utilizing other reagents acquired separately. We did not follow strictly the kit recommendation of generating primers with a melting temperature $>75^{\circ}\text{C}$, utilizing in many cases primers with lower melting temperatures. This did not noticeably decrease the efficiency of plasmid copying in the elongation step of the cycle used (1-min 95°C -melting, 1-min annealing at a temperature appropriate for the plasmids used; 9-min elongation at 72° ; this process repeated generally only 16 times), despite the utilization of an elongation temperature that in some cases exceeded that of melting of the plasmid-DNA complex. Possibly during the low-temperature annealing step there was enough elongation to increase the melting temperature above that used in the proper elongation step. Only in few cases we had to add 2-4% dimethylsulfoxide (an additive used for long-PCR [201]) for successful results. In our hands, the efficiency of this system is close to the manufacturer claim of 80% success. We have always sequenced the entire *CPS1* gene of the mutant plasmid, and we only found very rarely unwanted mutations.

INTRODUCTION

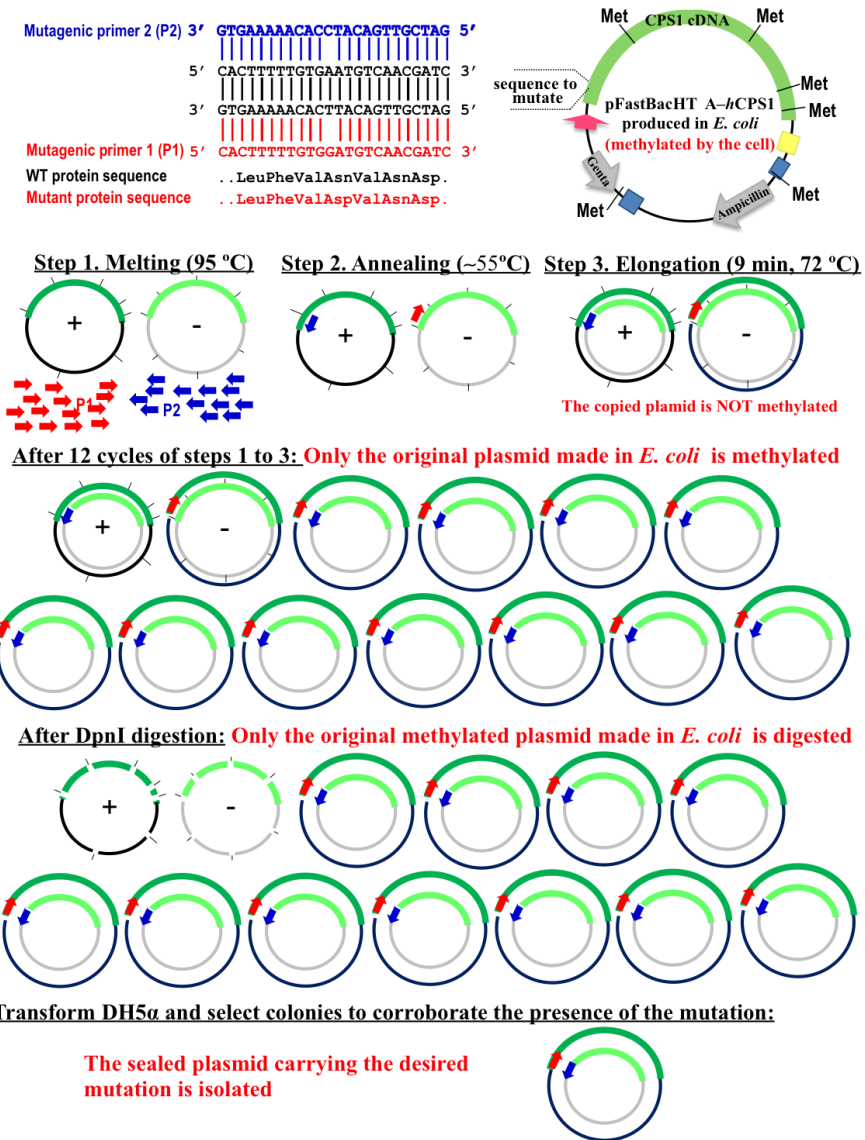
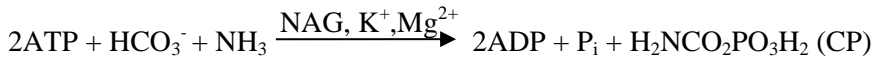


Figure 22. Site-directed mutagenesis represented for a starting single plasmid molecule. The mutagenic primers used as an example (the sequences at the top and the blue and red arrows in the remainder of the figure) are the ones used to introduce the p.Asn355Asp mutation (Chapter 1 of the Results section). The vector is shown in full at the top, but is schematized as circles in the remainder, coloring green the CPS1 cDNA insert and using different color hues for plus and minus DNA strands. Adenine methylations (represented by black lines) are non-exact examples used for illustration.

1.10 CPS1 activity assays

CPS1D refers to the existence of too little CPS1 activity to cope with the physiological needs of ammonia detoxification. This can be the consequence of too little active protein present in the tissue or of normal levels of an abnormal form of the enzyme exhibiting low specific activity (or of a combination of the two). Therefore, the assay of enzyme activities of the purified enzyme has been a key element in the present studies. I will comment here mainly the rationale of the enzyme activity determinations used in the present work.

As already indicated, CPS1 catalyzes the following reaction:



We have chosen here the utilization of optical assays that determine product concentration because they are simpler and less labor-intensive than procedures based on product isolation, such as chromatography. The choice of measuring optically product increase rather than to follow substrate decrease (as it could be done, for example, for ATP) is based on the fact that in substrate determination-based approaches the readings of the analyte (the substrate) are initially very high, decreasing the sensitivity of the assay. It is more sensitive to measure a product for which the concentration is essentially zero at the beginning of the reaction (initial velocity conditions).

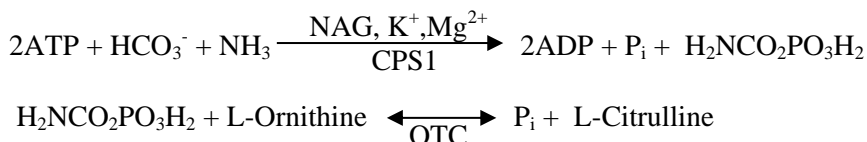
A point to remember with CPS assays is that carbamoyl phosphate is unstable, with a $t_{1/2}$ of ~40 min at 37°C and neutral pH [202], but with much faster decomposition at acidic and basic pH values [203]. This led us to choose an assay based on the continuous conversion of CP to citrulline, using the enzyme ornithine transcarbamylase (OTC) in the presence of ornithine. This is a time-honored approach [41] that was amenable to our laboratory since OTC is commercially available (from Sigma) and, furthermore, since we produce from *E. coli* pure recombinant *Enterococcus faecalis* OTC [204].

Although OTC catalyzes a reversible reaction, the equilibrium is highly displaced in the direction of citrulline synthesis [205] and, if the reaction is carried out in the absence of phosphate and in the presence of a high concentration of ornithine (the case in our assays), conversion of CP to citrulline is highly favored. The presence of ornithine in the reactional assay is not a problem with CPS1, since, in contrast with what is the case with bacterial CPS [6], ornithine is not an activator of CPS1. Other advantages of including OTC and ornithine in the CPS1 assay are the lack of any requirements for the OTC reaction other than the presence of its two substrates [205], and the fact

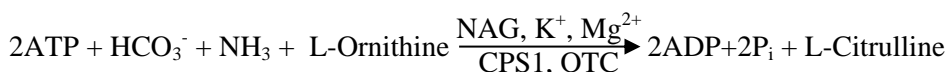
INTRODUCTION

that the immediate conversion of CP to citrulline minimizes the risk of decomposition of the CP. Furthermore, with this system the concentration of CP is expected to be always be very low, thus approaching the zero product concentration requirement of initial velocity assays.

It would be possible to monitor the CPS1 reaction with OTC coupling by measuring phosphate, or by citrulline assay, with stoichiometries of two moles of P_i released per mole of citrulline made by the coupled reactions of CPS1 and OTC:



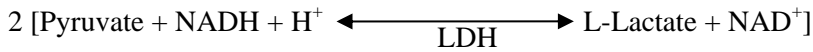
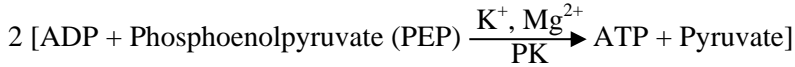
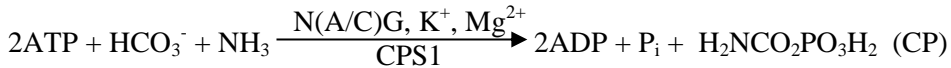
Total reaction:



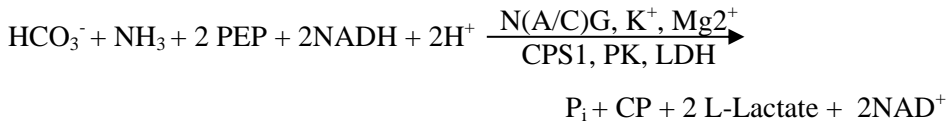
We have chosen to determine citrulline rather than P_i , at the end of a given period of assay time. This period can be lengthened in the OTC-coupled assay as much as needed, since CP is not decomposed due to its immediate conversion to citrulline. By measuring citrulline rather than P_i the total reaction is monitored, rather than risking to determine only the partial reaction of ATP hydrolysis. It has to be remembered that CP is the final product of the reaction, and thus, that it is generated only if all the reactional steps are completed. In contrast, P_i would be generated even if no citrulline were made, provided that the first step of the CPS1 reaction were operative (Scheme 1, pg. 13). Another reason for choosing citrulline is the existence of a highly sensitive color reaction for citrulline [29], which is based on the classical Archibald reaction for ureido compounds [206]. The sensitivity of this assay is good enough to allow monitoring substrate kinetics even at rather low concentrations of substrates without deviating too much from the initial velocity condition (permissivity of our assays, 25% consumption of the substrate being varied).

The N-carbamoyl-L-glutamate (NCG) NAG analog used in some of the present studies is an ureido compound, and, therefore, it produces color in the colorimetric assay for citrulline [206]. Therefore, when we have used NCG, particularly for assaying enzyme kinetics for variable NCG concentrations, we have utilized an alternative assay based on the continuous determination of ADP with pyruvate kinase (PK) and lactate dehydrogenase (LDH) [39]:

INTRODUCTION



Total reaction:



This assay is based on the fact that the reduced form of NADH, but not the oxidized form, NAD^+ , strongly absorbs light in the near ultraviolet [207] (maximum at $\lambda = 340 \text{ nm}$, with a molar extinction coefficient of $6.2 \times 10^3 \text{ Mol}^{-1} \text{ cm}^{-1}$). Provided that there is enough phosphoenolpyruvate, this assay regenerates the ATP used, and, therefore, it is also highly appropriate for measuring substrate kinetics for ATP. It is true that this assay is not useful with highly crude enzyme preparations or when the sample is contaminated with other ATP-hydrolyzing activities. Nevertheless, this problem can be alleviated in the case of CPS1 activity by determining the NAG-dependent (or NCG-dependent) ADP production, estimated as the increase in ADP production that is triggered by the addition of NAG or NCG.

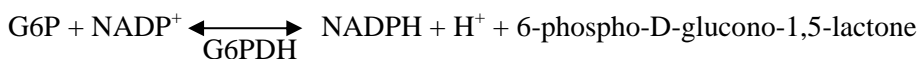
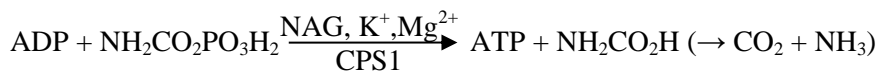
In any case, we have used the PK/LDH-coupled assay exclusively with the pure enzyme, and, therefore, the presence of spurious ATPases has not been a problem, even in the few cases in which we have determined the slow ($\sim 10\%$ of the rate of the complete activity) bicarbonate-dependent ATPase activity [16]. This partial activity (see Scheme 1, pg. 13) is measured in the absence of ammonia, and reflects the slow hydrolysis of the carboxyphosphate formed in the first step of the reaction.

In some cases, where many assays had to be carried out in parallel, we adapted the 340 nm wavelength assay to a multiwell plate. This enabled us to carry out the assay in 200 μl , using less enzyme and performing multiple assays in parallel. To standardize the plate assay, we measured the absorbance at 340 nm of increasing concentrations of NADH in both a regular spectrophotometer using a 1-cm pathlength cuvette and in the multiwell plate, and the results

INTRODUCTION

obtained indicated that the measurements done in the cuvette gave 2-fold higher absorbance than those done in 200 μ l. We used this correlation for calculations.

Another continuous assay, in this case based on the reduction of NADP (and therefore on the increase in the absorbance at 340 nm) has been utilized when needed for determining the partial reaction of ATP synthesis from ADP and carbamoyl phosphate which reflects the second phosphorylation step of the CPS1 reaction (Scheme 1, pg. 13). The coupling enzymes used are hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6PDH) [208], as follows:



In summary, we have made extensive use of enzyme activity assays. These have been generally citrulline assays carried out in test tubes in which the citrulline produced was determined after 10-30 min at 37°C (generally 10 min). However, in some cases we have determined ADP or ATP production in continuous spectrophotometric assays at 340 nm in 1-ml cuvettes or in multiwell plates.

1.11 Study goals, contents, involved laboratories, and my role in this work.

The central goals of this work are:

- a) To produce recombinant human CPS1 as an experimental platform for testing the effects of mutations and of chemical chaperoning on CPS1 functionality and stability.
- b) To investigate the role on the enzyme functionality and stability of the intervening domain that links both phosphorylation domains in the CPS1 sequence, and to clarify the reasons for the high eloquence of this domain in the database of missense mutations found in CPS1D patients.
- c) To investigate in detail the process of CPS1 activation by NAG by examining and analysing the effects of all the clinical mutations reported in the C-terminal allosteric domain as well as by analysing the effects of a number of experimental mutations introduced on rational bases.

d) To assess the disease-causing role of all the mutations reported in patients of CPS1D that map in these two domains, and to determine the mechanism of the ill-effects of these mutations.

The Results section is composed of the following three publications:

- Chapter 1:

Díez-Fernández C, Martínez AI, Pekkala S, Barcelona B, Pérez-Arellano I, Guadalajara A, 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.* 8:1149-1159.

- Chapter 2:

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.

- Chapter 3:

Díez-Fernández C, Gallego J, Häberle J, Cervera J, Rubio V. Experimental studies on the inherited metabolic error carbamoyl phosphate synthetase 1 deficiency shed light on the mechanism for switching on/off the urea cycle. Submitted to *Journal of Genetics and Genomics*.

I will summarize briefly the contents of these three chapters:

- Chapter 1 largely deals with the expression and production of the human enzyme and with the study of its properties. It not only confirms that the recombinant enzyme is a genuine representation of the normal liver enzyme, but it also studies previously unexplored properties of it, and even investigates the potentiality of chemical chaperoning it with NCG. Pilot site directed mutagenesis studies are carried out to show the value of this system for studying CPS1 deficiency.

- Chapter 2 studies by site-directed mutagenesis the impact of all the CPS1D missense mutations described so far affecting a small CPS1 domain of unknown function that had been called the UFSD, and that links both phosphorylation domains in the enzyme sequence. This domain had been found to be clinically highly eloquent despite the fact that it does not host substrate-binding or catalytic machinery. Our findings support the disease-causing role of the mutations reported to affect this domain, highlighting the value of the present expression system for ascertaining the disease-causing potential of

INTRODUCTION

CPS1D mutations. They reveal this domain's key role for proper enzyme folding and for the regulatory cross-talk between NAG and phosphorylation sites. Our work has uncovered a paramount integrating role of this domain for building and for organizing the highly complex CPS1 architecture, leading us to call this domain the “Integrating domain” of CPS1.

- Chapter 3 focuses on understanding CPS1 activation by studying or analyzing the impact of all the missense mutations lying on the C-terminal allosteric domain that have been found to date in CPS1D patients, as well as the effects of other experimental missense mutations affecting residues proposed to be involved in NAG binding or NAG activation. The analysis has included, in addition to the mutations introduced now, seven clinical and five experimental mutations that had been studied already but for which the full significance had not been inferred. Our present results and analysis clarify disease causation when appropriate, identify functionally important residues, and reveal that many of these mutations decrease the affinity of the enzyme for NAG, opening the way for NAG saturation therapy with the NAG analog and registered drug N-carbamyl-L-glutamate. Furthermore, we identify and delineate using molecular dynamics a NAG-triggered conformational change in the $\beta 4$ - α loop of the CPS1 C-terminal domain that is most likely the first event in the NAG activation process. Thus, we provide the first hint on the mechanism by which NAG can control the operation of the urea cycle.

This work has been carried out in three different laboratories: from 01.09.2009 to 31.12.2011 in the laboratory of Dr. Javier Cervera, in the Centro de Investigación Príncipe Felipe until the laboratory was closed; then the work continued from 01.01.2012 till the 31.05.2014 in the laboratory of Dr. Vicente Rubio, in the Instituto de Biomedicina de Valencia of the CSIC; with a short stay (16.11.2012-26.4.2013) in the laboratory of Dr. Johannes Häberle in the University Children’s Hospital in Zurich, Switzerland in the collaborative work that was a part of the work included in sections 2 and 3 of the Results.

As the results of this work are conveyed in three multiauthor papers, it is important to clarify which part of the work is due to myself. My personal work includes my central involvement in the writing of the three papers, in the analyses conveyed therein, whereas I have been responsible for the entire experimental work of Chapters 2 and 3, excepting the molecular dynamics and docking studies of Chapter 2 (which were carried out by Dr. José Gallego). In the case of Chapter 1 I was responsible for most of the work on human enzyme expression, purification and characterization, but I was not responsible for the studies on the rat enzyme or on mutant enzyme forms except in the case of the p.Asn355Asp mutation, which I studied myself.

2 Results

- **CHAPTER 1: Molecular characterization of carbamoyl-phosphate synthetase (CPS1) deficiency using human recombinant CPS1 as a key tool.**
- **CHAPTER 2: 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.**
- **CHAPTER 3: Experimental studies on the inherited metabolic error carbamoyl phosphate synthetase 1 deficiency shed light on the mechanism for switching on/off the urea cycle**

RESULTS

RESULTS

2.1 CHAPTER 1

MOLECULAR CHARACTERIZATION OF CARBAMOYL- PHOSPHATE SYNTHETASE (CPS1) DEFICIENCY USING HUMAN RECOMBINANT CPS1 AS A KEY TOOL

by

Carmen Díez-Fernández,^{1,2†} Ana I. Martínez,^{2†} Satu Pekkala,^{2‡} Belén
Barcelona,¹⁻³ Isabel Pérez-Arellano,^{2,3} Ana María Guadalajara,² Marshall
Summar,⁴ Javier Cervera,^{1-3*†} and Vicente Rubio^{1,4*†}

¹*Instituto de Biomedicina de Valencia (IBV-CSIC), Valencia, Spain;* ²*Centro de Investigación Príncipe Felipe, Valencia, Spain;* ³*Group 739, CIBERER, ISCIII, Spain;* ⁴*Childrens National Medical Center, Washington, District of Columbia*

Published in: Human Mutation (2013) 34:1149-1159

†CD-F and AIM have contributed equally to this work. JC and VR have contributed equally to this work.

‡Current address: Department of Health Sciences, University of Jyväskylä, Finland;

*Correspondence to: Vicente Rubio or Javier Cervera
Instituto de Biomedicina de Valencia
C/ Jaime Roig 11, Valencia-46010, Spain.
E-mail: rubio@ibv.csic.es;
E-mail: jcermir@gmail.com

Contract grant sponsors: Fundación Alicia Koplowitz 2011; Valencian Government (Prometeo 2009/051); Science Department of the Spanish Government (BFU2011-30407 and SAF2010-17933). CD-F was a FPU fellow of the Ministry of Euducation of Spain; AIM was a CIPF-Bancaja Fellow

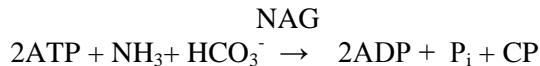
2.1.1 Abstract: The urea cycle disease carbamoyl-phosphate synthetase deficiency (CPS1D) has been associated with many mutations in the *CPS1* gene [Häberle et al. Hum Mutat 2011; 32:579-589]. The disease-causing potential of most of these mutations is unclear. To test the mutations effects, we have developed a system for recombinant expression, mutagenesis, and purification of human carbamoyl-phosphate synthetase 1 (CPS1), a very large, complex and fastidious enzyme. The kinetic and molecular properties of recombinant CPS1 are essentially the same as for natural human CPS1. Glycerol partially replaces the essential activator N-acetyl-L-glutamate (NAG), opening possibilities for treating CPS1D due to NAG site defects. The value of our expression system for elucidating the effects of mutations is demonstrated with eight clinical *CPS1* mutations. Five of these mutations decreased enzyme stability, two mutations drastically hampered catalysis, and one vastly impaired NAG activation. In contrast, the polymorphisms p.Thr344Ala and p.Gly1376Ser had no detectable effects. Site-limited proteolysis proved the correctness of the working model for the human CPS1 domain architecture generally used for rationalizing the mutations effects. NAG and its analogue and orphan drug N-carbamoyl-L-glutamate, protected human CPS1 against proteolytic and thermal inactivation in the presence of MgATP, raising hopes of treating CPS1D by chemical chaperoning with N-carbamoyl-L-glutamate.

Key words: urea cycle; CPS1 deficiency; hyperammonemia; carbamylglutamate

2.1.2 Introduction

Carbamoyl-phosphate synthetase 1 (CPS1) deficiency (CPS1D; MIM# 237300) is a rare autosomal recessive inborn error of the urea cycle [Häberle et al., 2011], the cycle that detoxifies the neurotoxin ammonia produced in body protein catabolism. Unless promptly treated, the hyperammonemia caused by CPS1D can lead to encephalopathy, coma and death or mental retardation [Brusilow and Horwich, 2001; Häberle et al., 2012]. The time of onset and severity of the presentation appear related to the amount of residual activity of the enzyme in the liver [Shih, 1976].

Human CPS1 (hCPS1), a 1462-amino acid, 160-kDa multidomain mitochondrial liver and intestinal enzymatic protein, catalyzes the complex 3-step reaction that is the first of the urea cycle [Pierson and Brien, 1980; Rubio, 1993; Rubio et al., 1981, Pekkala et al., 2010]:



(NAG = N-acetyl-L-glutamate; essential activator of CPS1; CP = carbamoyl-phosphate)

The *CPS1* gene (MIM# 608307) spans ~120 kb, it maps to 2q35 [Summar et al., 1995], comprising 38 exons and 37 introns [Summar et al., 2003]. More than 230 genetic lesions have been reported in CPS1D, with little recurrence, since most mutations are "private" to individual families [Häberle et al., 2011], with about 140 of these mutations being missense changes for which the disease-causing role has not been proven in most cases.

In an earlier study [Yefimenko et al., 2005] we attempted to infer the disease-causing potential of missense mutations found in patients with CPS1D by introducing them in recombinantly expressed *Escherichia coli* CPS, studying experimentally the consequences of such introduction on the activity or the stability of the purified enzyme. Although useful, this approach had obvious drawbacks due to the limited (~40%) sequence identity [Nyunoya et al., 1985] and the large functional differences between the bacterial and human CPSs. These differences include the use and the lack of use by bacterial CPS of, respectively, glutamine and NAG [Meister, 1989], while CPS1 cannot use glutamine as ammonia donor, utilizing ammonia with high affinity, and it needs NAG as an essential allosteric activator without which it is inactive [Rubio et al., 1981, 1983a]. In fact, the role of the CPS1 N-terminal 40-kDa region (Fig. 1A), corresponding to the glutamine-splitting, small subunit of bacterial

RESULTS. Chapter 1

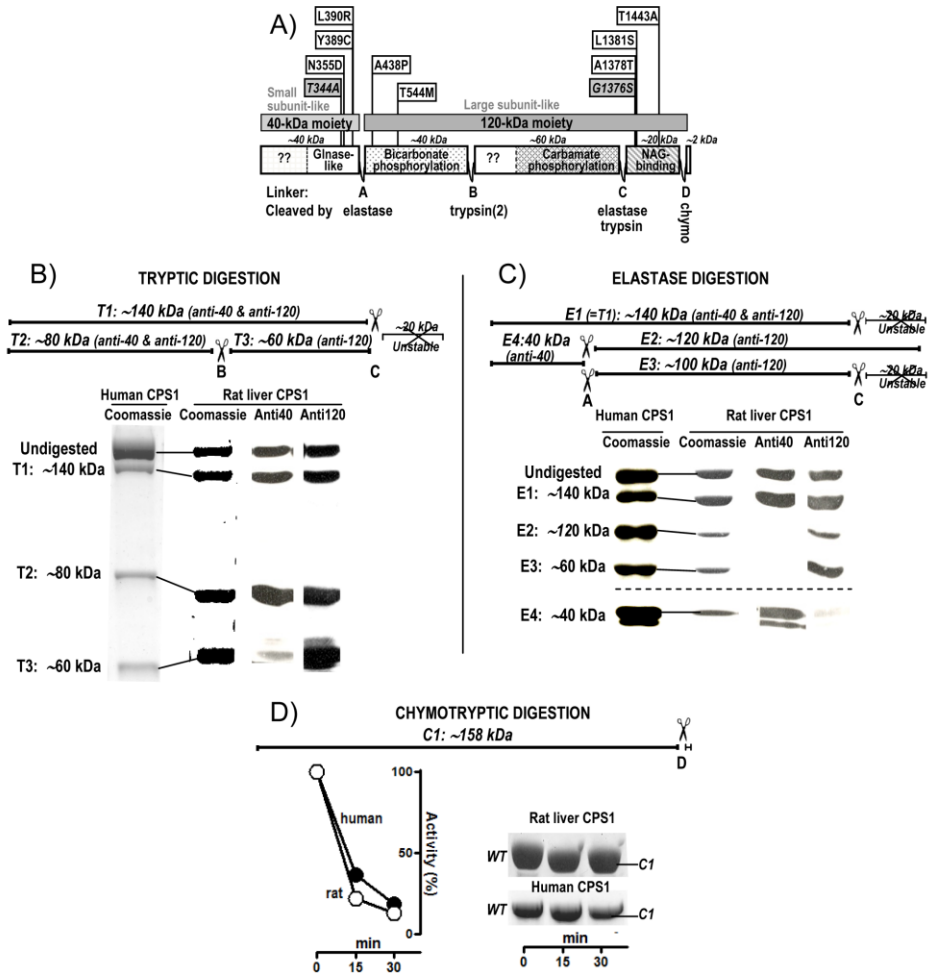


Figure 1. Domain composition of human CPS1. **A:** Linear scheme of mature CPS1, highlighting four interdomain linkers that are cleaved by the indicated proteases as identified in rodent CPS1, with the corresponding fragments masses (in kDa) above them [Marshall and Fahien, 1988]. Chymo, chymotrypsin. Trypsin(2) denotes cleavage after scission at the other tryptic site. The bars shaded grey schematize the 40-kDa N-terminal and the 120-kDa C-terminal CPS1 moieties that correspond to the small and large subunits of *E. coli* CPS, respectively. Polyclonal rabbit antibodies raised against these isolated moieties are called Anti-40 and Anti-120. Functional domains are shown in background texture and are identified. Glnase-like corresponds to the Glnase domain of bacterial CPS, but has no known function in CPS1. ??, unknown function. A dashed line separates two domains composing a proteolytic domain. (continued in the next page)

CPS [Meister, 1989; Nyunoya et al., 1985] is uncertain in CPS1, and thus, the impact of hCPS1 mutations affecting this region cannot be inferred from *E. coli* CPS studies. Furthermore, the unique characteristics among CPSs of the NAG activation of CPS1 makes difficult to infer from bacterial CPS studies the impact of hCPS1 mutations mapping in the C-terminal ~20-kDa of this enzyme, since this region (the allosteric or regulatory domain, Fig. 1A) hosts the site for NAG [Pekkala et al., 2009, Rodríguez-Aparicio et al., 1989] and must generate the allosteric signal that shift CPS1 from inactive to active.

We now exploit our recent success in producing recombinant rodent CPS1 in a baculovirus/insect cell system [Pekkala et al., 2009], to develop a similar system for recombinant production of pure human CPS1 (rhCPS1). Although harboring an N-terminal His-tag to help purification, rhCPS1 is proven here to have the same properties and characteristics as CPS1 purified from human liver [Pierson and Brien, 1980; Rubio et al. 1981] and the same domain composition as the well-studied rodent enzyme [Evans and Balon, 1988; Marshall and Fahien, 1988; Powers-Lee and Corina, 1986]. We demonstrate the value of this expression system for testing the functional impact of missense mutations found in CPS1D, thus helping infer the disease-causing role of these mutations. Furthermore, we show that NAG and its analogue and orphan drug N-carbamyl-L-glutamate (NCG) importantly decrease hCPS1 susceptibility to proteolytic

Fig. 1, cont'd: *The two polymorphisms (in italic and grey background) and the eight clinical mutations studied here are mapped in the CPS1 polypeptide with banners. B, C and D: Top part, fragments generated with trypsin, elastase or chymotrypsin [Marshall and Fahien, 1988]. The cleavage points are identified with the ABCD notation used in panel A. Each tryptic fragment is called T1-T3, elastase fragments E1-E4 and the chymotryptic one C1. Their approximate masses and their reactivity with Anti-40 and Anti-120 antibodies are given. The 20 kDa C-terminal fragment is rapidly degraded [Marshall and Fahien, 1988] and is shown crossed. Lower parts, SDS-PAGE of digested recombinant human CPS1 and, for comparison, of rat liver CPS1, stained with Coomassie or by immunoperoxidase with Anti-40 or Anti-120 after western blotting (only done with the rat enzyme). Digestions of CPS1 (1.3-2 mg/ml) were at 37°, for 15-30 min with the indicated protease (4-16 µg/ml; pancreatic, from Boheringer Mannheim or Sigma) in 35 mM Tris-HCl pH 7.4, 9% glycerol, 1.5 mM DTT, 20 mM KCl and 10 mM NAG. The enzyme was preincubated at least 15 min at 37°C prior to the addition of the protease. This addition was considered time zero. Fragments are identified in the gels as T1-T3, E1-E4 or C1. Note in (D) that while chymotrypsin inactivates rat and human CPS1 (see the plot), there is little decrease (~2 kDa) in polypeptide mass (top panels), corresponding to the loss of approximately 12 residues from the enzyme C-terminus documented earlier for rat CPS1 [Marshall and Fahien, 1988].*

attack and thermal inactivation, raising hopes that NCG might be used as a

RESULTS. Chapter 1

chemical chaperone for treating CPS1D due to misfolding-causing mutations. The present results have also produced novel information on the significance of the N-terminal and C-terminal domains of CPS1.

2.1.3 Patients and methods

2.1.3.1 Patients and *CPS1* mutations

The eight missense mutations chosen (Table 1) were reported [Eeds et al., 2006; Finckh et al., 1998; Kurokawa et al., 2007; Summar, 1998] in seven CPS1D patients with neonatal presentations, implying high disease severity. Patients 4 and 1 (Table 1) were, respectively, homozygous or compound heterozygous for one or two missense mutations, whereas in patients 2 and 6 the mutation was detected in mRNA studies that failed to detect a second mutant allele. The other three patients (Table 1) carried in one allele a missense mutation and in the other a truncation-causing change (nonsense changes in patients 3 and 7; a frameshift in patient 5) that, because of the large protein region deleted, should cause enzyme inactivation. The PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2>) [Adzhubei et al., 2010] and MutPred (<http://mutpred.mutdb.org>) [Li et al., 2009] servers assigned these mutations with a high probability of being pathogenic, whereas they made predictions of benignity for two polymorphisms causing non-synonymous amino acid substitutions [Finckh et al., 1998; Summar et al., 2003] that are studied also here as negative controls (Table 1).

All the mutations dealt with here were already included in the locus-specific database for *CPS1* (<http://www.lovd.nl/CPS1>). Amino acid conservation (Table 1) was determined by ClustalW sequence alignment of either CPS1, CPSIII or other CPSs from 15, 6 and 270 species, respectively.

2.1.3.2 Recombinant human CPS1 production

Human *CPS1* cDNA [Haraguchi et al., 1991] (GenBank entry NM_001875.4), was generated from human liver mRNA [Summar et al., 2003] as two complementary fragments by two RT-PCR reactions with appropriate primers. After sequential incorporation of these fragments into pcDNA3.1 (from Invitrogen), the complete *CPS1* cDNA was reconstructed within this plasmid by exploiting a unique *Hind*III *CPS1* site, yielding pcDNA3.1-*hCPS1*. Then (Supp. Fig. S1) a 3985 bp fragment comprising the *CPS1* open reading frame (ORF) from base 580 onwards (base 1 is the A of the translation initiation codon) was excised from this plasmid by *Bam*HI and *Eco*RI, and was ligated

Table 1. CPS1 Missense Mutations Found in Patients with CPS1 Deficiency, and Nonsynonymous Polymorphisms Studied Experimentally Here

Patient #	Allele	Nucleotide change ^a	Amino-acid change	Amino-acid in CPS			Domain	PolyPhen-2 prediction	MutPred prediction	
				I	III	Other			g score	Hypothesis
<i>Clinical mutations</i>										
1 ^b		c.1063A>G	p.Asn355Asp	N	N	Variable	Glnase-like	Probably damag.	0.875	
2		c.1166A>G	p.Tyr389Cys	Y/F	Y/F	Variable	Glnase-like	Possibly damag.	0.835	
2 ^b	1	c.1169T>G	p.Leu390Arg	L	L	L/I/F/V	Glnase-like	Probably damag.	0.929	C: Loss stability
	2		Unknown							
3 ^c	1	c.1312G>C	p.Ala438Pro	A	A	A	HCO ₃ ⁻ phosphorylation	Probably damag.	0.969	VC: Loss catalysis
	2	c.130C>T	p.Gln44*							
4 ^d	1	c.1631C>T	p.Thr544Met	T	T	A/G/S/T	HCO ₃ ⁻ phosphorylation	Possibly damag.	0.869	
	2	c.1631C>T	p.Thr544Met							
5 ^b	1	c.4132G>A	p.Ala1378Thr	A	A	A	Apolar	Probably damag.	0.919	
	2	c.3185delA	p.Asn1062Thrfss*38							
6 ^e	1	c.4142T>C	p.Leu1381Ser	L	L	L/F/Y/A	NAG-binding	Probably damag.	0.923	C: Loss stability
	2		Unknown							
7 ^b	1	c.4327A>G	p.Thr1443Ala	T	T/S	Variable	NAG-binding	Possibly damag.	0.738	
	2	c.3784C>T	p.Arg1262*							
<i>Polymorphisms</i>										
	-	c.1030A>G	p.Thr344Ala	T/S	S	Variable	Glnase-like	Benign	0.185	
	f	c.4126G>A	p.Gly1376Ser	G	A	Variable	NAG-binding	Benign	0.129	

All patients were reported to exhibit neonatal presentations. Amino acids are shown in three-letter code for the human substitutions and in one-letter code for occurrence in the various CPSs. *Variable* denotes the occurrence at a given position in the indicated group of CPSs, of >4 types of amino acids with no constant chemical characteristics (polar, apolar, charged, etc.). When the second allele is not a missense change, no data are given on residue conservation and pathogenicity potential. For localization of the CPS1 domains and the mapping of the mutations in the CPS1 polypeptide, see Figure 1A. PolyPhen-2 grades the probability of a damaging effect of an amino-acid substitution, as *probably damaging*, *possibly damaging* and *benign*. MutPred gave a g score corresponding to the probability that a given amino-acid substitution was deleterious/disease-associated. When indicated, this server was very confident (VC) or confident (C) that the indicated changes caused loss of stability or loss of a catalytic residue.

^a cDNA reference sequence NM_001875.4 (GenBank). +1 corresponds to the A of the translation initiation codon in the reference sequence (according to journal guidelines under www.hgvs.org/mutnomen).

^b Eeds et al. (2006).

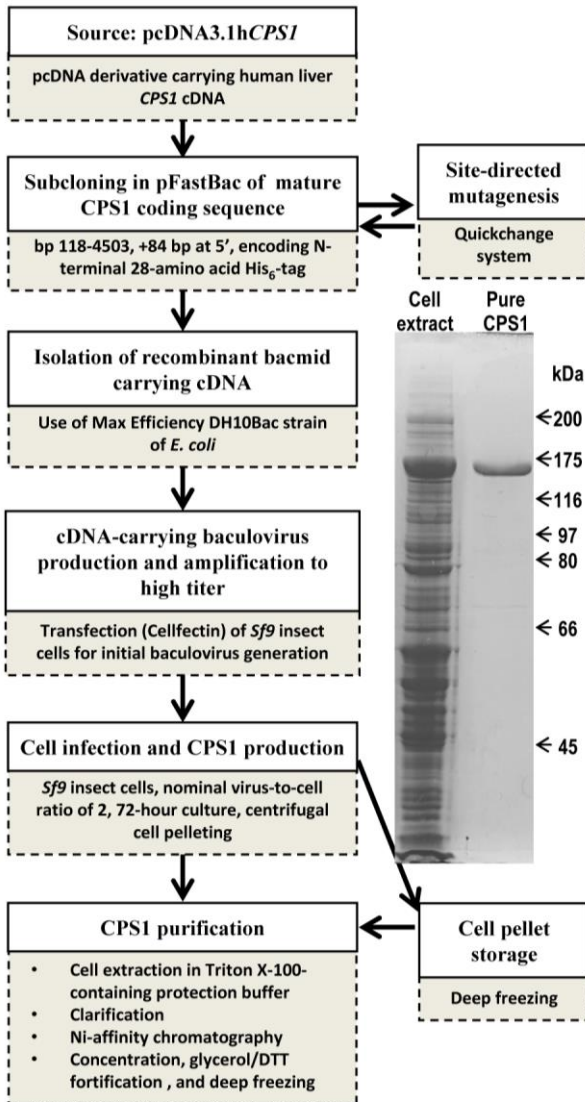
^c Kurokawa et al. (2007).

^d Finckh et al. (1998).

^e Summar (1998).

^f Summar et al. (2003).

RESULTS. Chapter 1



Supplementary Figure S1. Diagram schematizing the steps of the production of recombinant human CPS1. The gel (SDS-PAGE, Coomassie staining, the arrows give the positions of protein standards of the indicated polypeptide masses) illustrates the presence of abundant soluble CPS1 protein in the postcentrifugal supernatant of the cell extract (left track) and the essential homogeneity of the purified protein (right track).

into pFastBac. The ORF encoding mature CPS1 (lacking the N-terminal mitochondrial targeting sequence, bp 1-117) was completed by in-frame ligation of a PCR-generated fragment comprising bp 118-579 of the *CPS1* cDNA (primers: Cloning_F and Cloning_R, Supp. Table S1; they incorporate a *Bam*HI site for cloning). This yielded pFastBac-CPS1, which encodes mature hCPS1 (amino acids 40-1500) preceded N-terminally by the His₆-tag MSYYHHHHHHHDYDIPTTENLYFQGAMDP. Site-directed mutagenesis of pFastBac-CPS1 was performed by the overlapping extension method

(Quickchange kit from Stratagene) using the forward and reverse primers given in Supp. Table S1. The correctness of the constructs, the presence of the desired mutation, and the absence of unwanted mutations, were corroborated by sequencing.

For producing rhCPS1 (Supp. Fig. S1), we used the commercial Bac-to-Bac[®] Baculovirus Expression System (Invitrogen), following the manufacturer's directions. In short, *E. coli* Max Efficiency DH10Bac cells (Invitrogen), transformed with pFastBac-CPS1, were grown on LB-agar containing 50/7/10/40/100 µg/ml of, respectively, kanamycin/gentamycin/tetracyclin/IPTG/Bluo-Gal. Individual white colonies were inoculated into 5-ml LB medium with the same antibiotics, cultured overnight, and the bacmid was isolated. The baculovirus was produced by transfecting Sf9 insect cells with CPS1 cDNA-carrying (proven by PCR) bacmid, using Cellfectin/Grace medium (5 hours, 27°C), followed by 3-day culture (27°C, six-well plate) in Sf900 medium (Invitrogen) containing 0.1% Pluronic F-68, 50 U/ml penicillin and 50 µg/ml streptomycin. The culture was centrifuged, and the supernatant was used for baculovirus enrichment by infecting with it (1:60 dilution) a suspension of 1.5×10^6 Sf9 cells/ml. After 48-hour culturing (27°C, orbital shaking at 125 rpm) and centrifugation, the supernatant was used to inoculate (1:50 dilution) a fresh cell suspension for CPS1 production, collecting the cells by centrifugation after 3 days of culture as above.

2.1.3.3 Enzyme purification

Unless indicated, all steps (Supp. Fig. S1) were at 4°C. To purify rhCPS1 (wild-type or mutant forms), the insect cell pellet from a 50-ml culture was suspended in 3 ml of a lysis solution [50 mM glycyl-glycine, pH 7.4, 1 mM dithiothreitol (DTT), 10% glycerol, 20 mM KCl, 0.1% Triton X-100, 5 µM E-64 protease inhibitor and 1% of the protease inhibitor cocktail for His-tagged proteins (Sigma product P8849)] and was thawed (melting ice) and frozen (liquid nitrogen or dry CO₂-acetone mixture) three times. After 10-min centrifugation (16,000xg) and supernatant filtration through a 0.22 µm membrane, the supernatant was applied to a HisTrap HP 1-ml column fitted in an ÄKTA FPLC system (GE Healthcare) that was equilibrated with 50 mM glycyl-glycine pH 7.4, 1 mM DTT, 10% glycerol, 0.5 M NaCl, and 20 mM imidazole. After a 10-ml column wash, a 15-ml linear gradient of 20-500 mM imidazole in the same solution was applied. The CPS1-containing fractions (monitored by SDS-PAGE) were pooled, concentrated to 3-5 mg protein/ml by centrifugal ultrafiltration (100-kDa cutoff membrane, Amicon Ultra, Millipore), enriched with 10% extra glycerol and 1 mM extra DTT, and frozen at -80°C.

RESULTS. Chapter 1

Rat liver CPS1, and *E. coli*-expressed recombinant *Enterococcus faecalis* ornithine transcarbamylase (OTC) [Barcelona-Andrés et al., 2002] were purified as reported [Alonso and Rubio, 1983; Marshall and Cohen, 1972].

2.1.3.4 Enzyme activity assays

In the standard CPS1 assay, CP was converted to citrulline, which was measured colorimetrically [Pekkala et al., 2009]. The enzyme was incubated 10 min at 37°C in an assay mixture containing 50 mM glycyl-glycine pH 7.4, 70 mM KCl, 1 mM DTT, 20 mM MgSO₄, 5 mM ATP, 35 mM NH₄Cl, 50 mM KHCO₃, 10 mM NAG, 5 mM L-ornithine and 4 U/ml OTC. When the concentration of a substrate was varied, other substrates were kept at the concentrations given above (unless indicated), with MgSO₄ being in 15 mM excess over ATP.

Since NCG yields color in the citrulline assay, when testing the NCG concentration-dependence of CPS1 activity, we used a continuous pyruvate kinase/lactate dehydrogenase coupled assay in which ADP production was monitored as NADH oxidation at 340 nm [Guthorlein and Knappe, 1968]. The assay (37°C) used the same solution as the standard assay except for the lack of OTC and ornithine and the inclusion of 2.5 mM phosphoenolpyruvate, 0.25 mM NADH, 40 µg/ml pyruvate kinase and 25 µg/ml lactate dehydrogenase.

The kinetic parameters for NAG were identical, within experimental error, in this assay and in the standard assay. This NADH oxidation-coupled assay, without NH₄Cl, was used for measuring the HCO₃⁻-dependent ATPase partial reaction of CPS1. For measurement of the partial reaction of ATP synthesis from CP and ADP, an NADP reduction-coupled assay [Yefimenko et al., 2005] was used, monitoring the absorbance at 340 nm in a mixture at 37°C containing 50 mM glycylglycine pH 7.4, 0.1 M KCl, 15 mM MgSO₄, 15 mM glucose, 0.5mM ADP, 5 mM CP, 10 mM NAG, 1 mM NADP, 1 mM DTT, 0.1 mg/ml hexokinase and 25 µg/ml glucose-6-phosphate dehydrogenase. One CPS1 unit produces per minute 1 µmol citrulline or 2 µmol ADP. The program GraphPad Prism (GraphPad Software, San Diego, California) was used for curve fitting.

2.1.3.5 Other techniques

SDS-PAGE [Laemmli, 1970] was performed in 8% polyacrylamide gels, with Coomassie staining or, for cell extracts, by western blotting/immunostaining (ECL system, GE Healthcare), utilizing an anti-rat liver CPS1 first antibody

[Alonso et al., 1989]. Western blots of rat liver CPS1 protease digests (Fig. 1) were stained with immunoperoxidase [Alonso et al., 1989] using as first antibodies rabbit antisera against the electrophoretically separated [Amero et al., 1994] N-terminal 40-kDa or C-terminal 120-kDa moieties of rat liver CPS1 (produced by limited elastase digestion [Marshall and Fahien, 1988]). Protein was determined according to [Bradford, 1976] using bovine serum albumin as standard.

Supplementary. Table S1. Synthetic oligonucleotides used in cloning and site-directed mutagenesis

Name/ mutation	Direction	Sequence 5' - 3'
Cloning_F	Forward	CGCAT <u>GGATCC</u> GTCTGTCAAGGCACAGACA ^a
Cloning_R	Reverse	TTTATTT <u>GGATCC</u> CACAAAATCCACAGG ^a
p.T344A	Forward	CTATGCCTTGGACAACGCTCTCCCTGCTGGC ^b
p.T344A	Reverse	GCCAGCAGGGAGAG CG TTGTCCAAGGCATAG ^b
p.N355D	Forward	CACTTTTTGTGGATGTCAACGATC ^b
p.N355D	Reverse	GATCGTTGACATCCACAAAAAGTG ^b
p.Y389C	Forward	CAATAGACACTGAGT GC CTGTTTGATTCC ^b
p.Y389C	Reverse	GAATCAAACAGG CA CTCAGTGTCTATTGG ^b
p.L390R	Forward	CAATAGACACTGAGTAC CGG TTTGATTCC ^b
p.L390R	Reverse	GAATCAAAC CGG TACTCAGTGTCTATTGG ^b
p.A438P	Forward	CCATTGGTCAGCCTGGAGAATTTG ^b
p.A438P	Reverse	GTAATCAAATTCTCCAG G CTGACCAATGGAC ^b
p.T544M	Forward	CATTATGGCTAT GGA AGACAGGCAGCTG ^b
p.T544M	Reverse	CTGCCTGTCTTCCATAGCCATAATGGAC ^b
p.G1376S	Forward	CCAAGATTCCTT AGT GTGGCTGAACAATTAC ^b
p.G1376S	Reverse	GTAATTGTT CAG CCACACTAAGGAATCTTGGCC ^b
p.A1378T	Forward	CCTTGGTGT GACT GAACAATTAC ^b
p.A1378T	Reverse	GTAATTGTT CAGT CACACCAAGGAATC ^b
p.L1381S	Forward	CTTGGTGTGGCTGAACAAT CAC ACAATGAAGGTTTCAAG ^b
p.L1381S	Reverse	GAAACCTT CATTGTGTG ATTGTT CAG CCACACCAAGG ^b
p.T1443A	Forward	CTTCCCAACAACAAC GCT AAATTTGTC ^b
p.T1443A	Reverse	GGACAAATTT AGCG TTGTTGTTGGGAAG ^b

^aUnderlining marks the BamH1 restriction sites

^bBold type indicate base substitutions to introduce the desired mutation.

2.1.4 Results and Discussion

2.1.4.1 Producing human CPS1 in a baculovirus/insect cell system

In the liver, CPS1 is produced as a precursor that is matured by cleavage of its N-terminal 38-39 amino acids upon entry to the mitochondria [Ryall et al., 1985]. We cloned for recombinant expression the cDNA for human liver CPS1 without the N-terminal 39 codons, which were replaced by a 28-codon N-terminal His₆-tag. This allowed testing the effects of the mutations on the mature form of the enzyme (the functional one *in vivo*). The tag simplified and speeded CPS1 purification, what is important, given the instability of mammalian CPS1 [Raijman and Jones, 1976] and its sensitivity to proteolytic attack [Guadalajara et al., 1983]. The CPS1 expressed here carries the more frequent Thr form of the p.Thr1406Asn polymorphism (rs1047891, Ensemble database; allelic frequencies of T/N forms, ~0.7/0.3) that has been associated with increased frequency of some vascular pathologies possibly related to decreased citrulline levels and nitric oxide production (see for example [Pearson et al., 2001]).

The procedure used (Supp. Fig. S1), optimized for highest rhCPS1 production, had as important elements the use of insect cells cultured for two weeks after unfreezing, and the infection of the cells with a nominal virus-to-cell ratio of 2 in the final CPS1-production step, leaving 72 hours the infected cells in culture (27°C, orbital shaking, 125 rev/min), in either 50 or 200 ml of medium, before cell harvesting. Purification was possible weeks after harvesting by freezing cells pellets at -80°C.

Prior studies with liver-purified hCPS1 [Pierson and Brien, 1980; Rubio et al., 1981] and even more extensive studies with rat liver CPS1 [Alonso et al., 1992; Guadalajara et al., 1983; Guthöhrlein and Knappe, 1968; Marshall and Fahien, 1985, 1988; Raijman and Jones, 1976] showed that mammalian CPS1 is highly instable, requiring precautions to avoid oxidation, proteolytic cleavage and inactivation of unknown cause but preventable by glycerol. Taking into account these factors, we used for cell extract preparation and in subsequent steps a neutral medium at 4°C containing 10% glycerol, 1 mM DTT and a very extensive protease inhibitor cocktail, using a fast 3-step purification protocol consisting of cell disruption by freeze-thawing in 0.1% Triton X-100-containing CPS1-protecting solution, centrifugal clarification and 0.22 µm-pore membrane filtration, and a final step of fast Ni-affinity column chromatography with imidazole gradient elution.

rhCPS1, which represented ~10% of the soluble cell protein (Supp. Fig. S1, left track of the gel), was largely soluble (not illustrated) and yielded after the column chromatography step ~15 mg per L of cell culture, of highly active (Table 2), homogeneous and pure enzyme as shown by the finding in SDS-PAGE of a single band migrating as expected for its sequence-deduced mass (163.9 kDa). (Supp. Fig. S1, right track of the gel).

Table 2. Comparison of the activity, the apparent K_m values for the substrates and the K_a values for N-acetyl-L-glutamate (NAG) or for its analog N-carbamyl-L-glutamate (NCG), for recombinant or natural human CPS1

Source	Activity U/mg	Apparent K_m or K_a value (mM)				
		ATP	HCO_3^-	NH_4^+	NAG	NCG
Recombinant ^a	1.1	0.5	4.0	1.0	0.14	2.0
Liver ^b	1.5	1.1	6.7	0.8	0.10	-
Liver ^c	1.5	0.3	2.2	1.3	0.15	2.0

^a*Present work.*

^b*Pierson and Brien, 1980.*

^c*Rubio et al., 1981.*

2.1.4.2 Recombinant human CPS1 represents well the natural liver enzyme

It was important to ascertain that rhCPS1 closely represents in all its properties natural hCPS, particularly since, to minimize enzyme inactivation, the N-terminal His₆-tag was not removed. Similarly to liver CPS1, rhCPS1 has an essential requirement for NAG (Fig. 2A). Its specific activity (Table 2) is similar to that reported for liver-isolated human CPS1 [Pierson and Brien, 1980; Rubio et al., 1981]. Apparent K_m values for ATP, HCO_3^- and NH_4^+ , and the K_a for NAG and for the drug analogue of NAG, NCG, are also within published value ranges for natural hCPS1 (Table 2 and Fig. 2) [Pierson and Brien, 1980; Rubio et al., 1981].

rhCPS1 also presents the same oligomeric state as liver-derived hCPS [Rubio et al., 1981]. Gel exclusion chromatography experiments (Fig. 3A) reveal highly similar peaks and estimated masses for the natural human enzyme and for rhCPS1: in both cases the apparent mass exceeds by <30% that of the monomer, indicating that the enzyme consists of monomers in rapid equilibrium with dimers, with strong predominance of the monomers.

RESULTS. Chapter 1

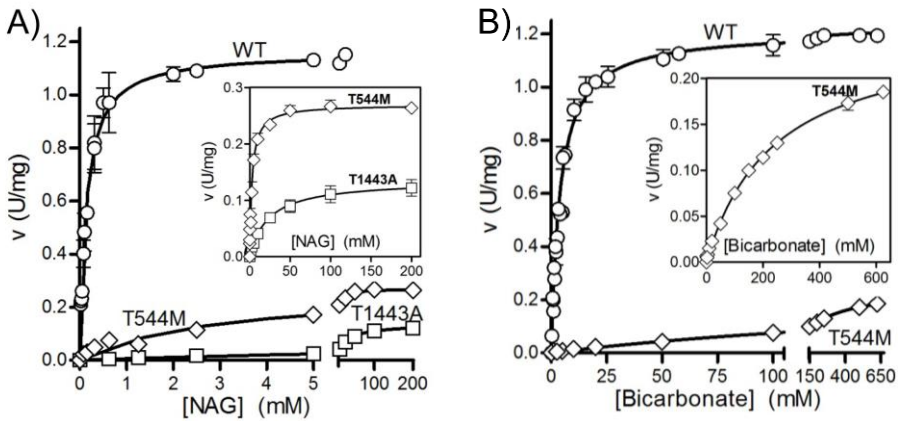


Figure 2. Dependence of CPS1 activity on the concentrations of NAG and bicarbonate for non-mutated (WT) CPS1 and for the indicated mutant forms. The curves fitted to the data are hyperbolae for the kinetic constants given in Table 3. The insets expand the curves for the mutants, to demonstrate the hyperbolic kinetics despite the decreased activity. A: NAG varied. In the case of the *p.Thr544Met* mutant the concentration of bicarbonate was fixed at 0.5 M. B: bicarbonate varied.

2.1.4.3 Glycerol partially replaces NAG in the activation of human CPS1.

We also investigated potentially crucial traits of hCPS1 that had not been amenable to investigation until now. Thus, we show here (Fig. 3B, right panel) that hrCPS1 is activated by glycerol in the absence of NAG, whereas in the presence of NAG is inhibited by increasing concentrations of glycerol. Similar observations had been made with rat liver CPS1 (Fig. 2B, left panel and [Britton et al., 1981; Rubio et al., 1983b]). The substantial activation attained with glycerol and the fact that this polyol appears to activate CPS1 without specifically binding to the NAG site [Rubio et al., 1983b] make conceivable the possibility of developing treatments for CPS1D patients with a damaged NAG site, in which CPS1 could be activated by compounds that do not bind to the NAG site.

2.1.4.4 Limited proteolysis reveals the hCPS1 domain organization.

The domain composition of hCPS1 had not been made amenable to investigation until now. The structure of *E. coli* CPS [Thoden et al., 1997] had revealed a multidomain organization that was anticipated [Rubio, 1993] by the

results of limited proteolysis and of other studies with rodent and *E. coli* CPSs [Cervera et al., 1993; Evans and Balon, 1988; Marshall and Fahien, 1988; Powers-Lee and Corina, 1986; Rodríguez-Aparicio et al., 1989; Rubio et al., 1991]. Limited proteolysis studies with rodent CPS1 revealed four points of preferential proteolytic cleavage that appear to correspond to exposed

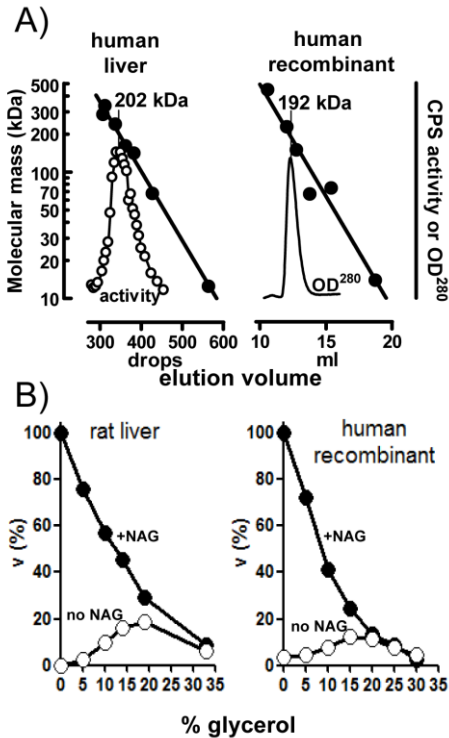


Figure 3. Recombinant human CPS1 replicates the natural enzyme in the oligomeric state and the ability to be activated and inhibited by glycerol. **A:** Gel exclusion chromatography analysis. The upper lines are the semilogarithmic plots of the masses of marker proteins (closed circles) versus their elution volumes. Peaks of CPS1 activity (left panel, open symbols) or of CPS1 protein (peak in the right panel) are shown below them. The mass of CPS1, estimated by interpolation, is shown above a vertical line emerging from the peak. Left panel, results with the enzyme purified from human liver [Rubio et al., 1981], using a conventional 0.9×56 cm column of Sephadex G-200 run at 23°C. For further details, including the list and masses of standards used, see [Rubio et al., 1981]. Right panel, present results with the recombinant human enzyme under essentially the same conditions as those in [Rubio et al., 1981] except for

the use of a Superdex 200HR (10/300) column mounted on an Äkta fast protein liquid chromatography system, at a flow rate of 0.5 ml/min, the application of 0.02 ml of a 3 mg/ml solution of human recombinant CPS1 or appropriate amounts of protein standards, and the continuous monitoring of the optical absorption at 280 nm. Protein standards for this panel were (masses are given in parenthesis, in kDa) bovine pancreatic ribonuclease A (13.7), bovine serum albumin (66.4), yeast alcohol dehydrogenase dimer (73.4) and tetramer (146.8), β -amylase (223.8) and ferritin (440). **B:** Effects of the addition of glycerol on CPS1 activity in the presence or in the absence of 10 mM NAG as indicated (closed and open circles, respectively). Velocities, measured as ADP production, are expressed as a percentage of the velocity in the presence of 10 mM NAG and in the absence of glycerol. Left and right panels, results with the rat liver enzyme and with the recombinant human enzyme, respectively.

RESULTS. Chapter 1

sequences linking adjacent globular domains, and which are differentially cleaved by elastase, trypsin and chymotrypsin [Powers-Lee and Corina, 1986; Marshall and Cohen, 1988] (Fig. 1). We observed limited tryptic or elastase digestion patterns with hrCPS1 that fully comply with the reported fragmentation patterns for rat liver CPS1 (Figs. 1B and C) [Marshall and Fahien, 1988; Powers-Lee and Corina, 1986]. These findings support the existence of the same domain organization and architecture in human and rat CPS1, also supporting the similarity of this architecture with that of *E. coli* CPS [Rubio et al., 1991; Thoden et al., 1997]. Interestingly, as with rat liver CPS1 [Marshall and Fahien, 1988], chymotrypsin inactivated rhCPS1 (Fig. 1D, plot) while decreasing very little CPS1 polypeptide size (Fig. 1D, gels), in agreement with prior experiments with rat liver CPS1 that showed that the enzyme is cleaved very close to its C-terminus [Marshall and Fahien, 1988].

2.1.4.5 Testing the effects of clinical CPS1D mutations on enzyme functionality

The present results show that the properties of rhCPS1 mirror those of natural CPS1, supporting the use of rhCPS1 for testing the impact of clinical mutations on enzyme function and stability. We demonstrate this use here with eight missense mutations found in seven neonatal CPS1D patients (Table 1 and Fig. 1A) [Eeds et al., 2006; Finckh et al., 1998; Kurokawa et al., 2007; Summar, 1998] and with two trivial polymorphisms (Table 1 and Fig. 1A) [Finckh et al., 1998; Summar et al., 2003]. Of these ten amino acid substitutions, two map in the bicarbonate phosphorylation domain (a catalytic domain) and are therefore likely to hamper activity. The other eight changes were selected because they map in CPS1 regions of unclear or unique function for which bacterial CPS would not be a good model. Thus, four of them map in the Glnase-like subdomain and another four in the C-terminal domain (Fig. 1A). In any case, *E. coli* CPS would not have been an optimal model for nine of these amino acid substitutions, since only one affected residue is strictly conserved in all CPSs (Table 1).

As expected, the polymorphisms p.Thr344Ala (c.1030A>G) and p.Gly1376Ser (c.4126G>A), mapping respectively in the Glnase-like and the NAG-binding domains (Fig. 1A), caused no negative effects on enzyme production (Fig. 4A),

Table 3. Influence of Clinical CPS1D Mutations on the Kinetic Parameters for the Substrates and for NAG

Enzyme form	Substrate or activator											
	NH ₄ ⁺			Bicarbonate			ATP			NAG		
	V ^{(Am)-∞} U/mg	K _m ^{app} mM	K _m ^{app} mM	V ^{(Bic)-∞} U/mg	K _m ^{app} mM	K _m ^{app} mM	V ^{(ATP)-∞} U/mg	K _m ^{app} mM	K _m ^{app} mM	V ^{(NAG)-∞} U/mg	K _m ^{app} mM	
Recombinant wild type form and forms with polymorphisms												
Wild type	1.19	1.0	4.0	1.23	4.0	4.0	1.22	0.47	0.47	1.15	0.14	
p.Thr344Ala	1.06	1.2 ± 0.1	4.3	1.04	4.3	4.3	1.06	0.45	0.45	1.02	0.17	
p.Gly1376Ser	1.16	1.1	4.5	1.28	4.5	4.5	1.22	0.45	0.45	1.17	0.12	
Recombinant forms carrying clinical missense mutations												
p.Asn355Asp	0.25	1.0	4.6 ± 0.7	0.26	4.6 ± 0.7	4.6 ± 0.7	0.28	0.57	0.57	0.26	0.34	
p.Tyr389Cys	0.73	1.1	4.2	0.79	4.2	4.2	0.73	0.43	0.43	0.70	0.12	
p.Ala438Pro	<0.01 ^a	-	-	<0.01 ^a	-	-	<0.01 ^a	-	-	<0.01 ^a	-	
p.Thr544Met	0.16	1.3	242	0.26	242	242	0.23	1.92	1.92	0.27	2.91 ± 0.38	
p.Ala1378Thr	0.93	0.9	4.3	0.98	4.3	4.3	1.01	0.90	0.90	0.97	0.21	
p.Thr1443Ala	0.04	1.0	3.6	0.034	3.6	3.6	0.04 ± 0.01	0.62	0.62	0.14	23.4 ± 5.2	

For clarity, standard errors are only given when they exceed 10% of the magnitude of the constant. The concentrations of NH₄⁺, bicarbonate, ATP and NAG were varied in the following mM ranges: 0.1–20, 0.3–625, 0.05–10 and 0–200, respectively. When not varied, the concentrations of these ligands were 35, 50, 5, and 10 mM, respectively (except for bicarbonate in the case of the p.T544M mutation, which was 0.5 M). Mg (as MgSO₄) was always in 15 mM excess over the ATP concentration.

^aDetection limit.

RESULTS. Chapter 1

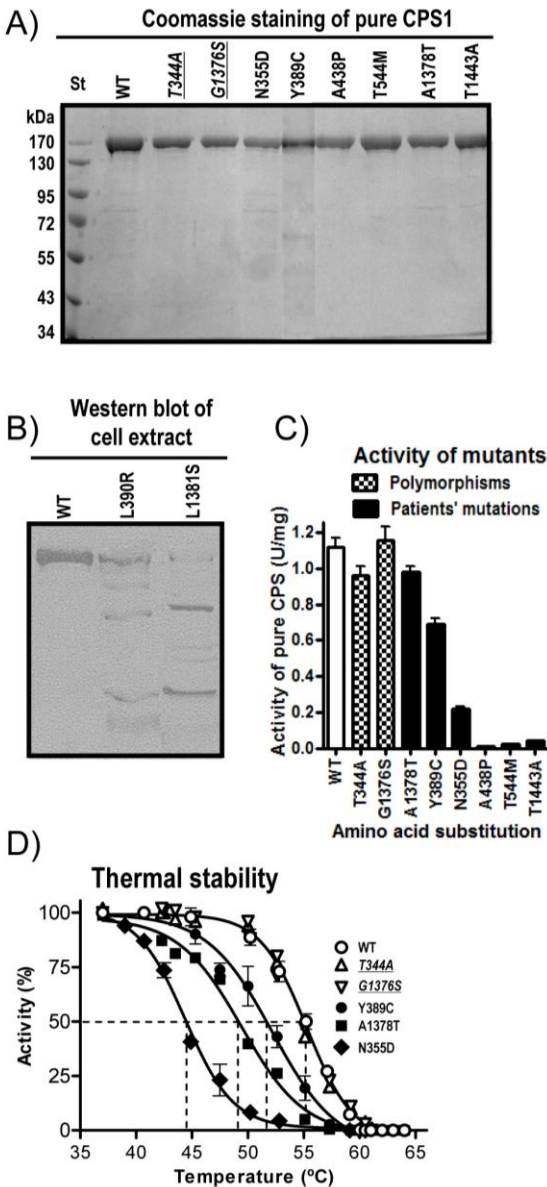


Figure 4. Production and properties of the enzyme forms with the amino acid substitutions studied here. **A:** SDS-PAGE (8% polyacrylamide, Coomassie staining) of purified human recombinant CPS1, either wild-type (WT) or carrying the indicated mutations. Polymorphisms are shown in *italic* script and underlined. St, protein markers with masses indicated on the side. **B:** Western blot of cell extracts (~20 μ g protein per well) of Sf9 cells infected with baculoviruses encoding either wild-type human CPS1 or its p.Leu390Arg or p.Leu1381Ser mutants. A polyclonal rabbit antiserum against rat liver CPS1 was used for immunostaining. **C:** Enzyme activity (standard assay conditions) of the purified wild-type or mutant human CPS1 forms. The bars for the wild-type enzyme, for the two polymorphisms and for the forms carrying clinical mutations are filled in white, checkerboard and black, respectively. Error bars give standard errors. **D:** Thermal stability of wild-type recombinant CPS1, and of the forms carrying either the polymorphisms (in *italic* and underlined), or the indicated clinical mutations. The enzyme, at 0.5–1 mg/ml in a solution of 50 mM glycyl-glycine pH 7.4, 20 mM KCl and 20% glycerol, was heated 15 min at the indicated temperature, then rapidly cooled to 0°C, and its activity determined immediately at 37°C. The horizontal dashed line marks 50% inactivation, whereas the vertical dashed lines cross the X-axis at the temperature at which 50% inactivation occurs for each enzyme form.

activity (Fig. 4C), kinetic parameters for each substrate or for NAG (Table 3), or on thermal stability (Fig. 4D). In contrast, two clinical mutations affecting respectively these same regions, p.Leu390Arg (c.1169T>G) and p.Leu1381Ser

(c.4142T>C) (Fig. 1A), were clearly disease-causing, since they induced strong CPS1 instability as revealed by western blotting of insect cell extracts that showed (Fig. 4B) proteolytic digestion bands instead of the clear-cut CPS1 band observed with the wild-type enzyme. Indeed, these two mutations, which replace hydrophobic residues (residues of hydrophobic nature are found at these positions in all known CPS sequences, Table 1) by polar residues of larger (p.Leu390Arg) or smaller (p.Leu1381Ser) size, were predicted by the MutPred server to be associated with loss of stability (Table 1). The other six mutations studied (Table 1) were sufficiently stable to allow purification (Fig. 4A). The p.Ala438Pro (c.1313G>C) and p.Thr544Met (c.1631C>T) mutations, which affect the bicarbonate-phosphorylation domain, and the p.Thr1443Ala (c.4327A>G) mutation, which affects the C-terminal domain, greatly decreased enzyme activity, to undetectable or nearly undetectable values (Fig. 4C), clearly indicating that they are disease-causing. Although having no detectable activity in the assay for the complete reaction (detection limit, 1% of the activity of wild-type rhCPS1), the p.Ala438Pro mutant catalyzed the partial reaction of ATP synthesis from ADP and CP (not shown) that is the reversal of the final step of the CPS1 reaction (a three-step reaction: 1- bicarbonate phosphorylation; 2- carbamate production from carboxyphosphate and ammonia; and 3- carbamate phosphorylation) but, as expected from the domain that is affected by the mutation, it failed to catalyze the bicarbonate-dependent ATPase partial reaction that reflects the bicarbonate phosphorylation step [Metzenberg et al., 1958; Rubio et al. 1981].

The large decrease in the activity of the p.Thr544Met mutant was shown to be due (Table 3 and Fig. 2) to the combination of a ~60-fold increase in the apparent K_m for bicarbonate, a ~20-fold increase in the K_a for NAG and a ~4-fold respective decrease and increase in the apparent V_{max} and K_m for ammonia. Except the increase in the K_a^{NAG} , these kinetic aberrations stem from the fact that this mutation affects a domain that catalyzes the initial two steps of the reaction, involving as substrates ATP, bicarbonate (step 1) and ammonia (step 2) [Rubio 1993; Thoden et al., 1997]. The increased K_a^{NAG} cannot be attributed to direct changes in the NAG site, which sits on another enzyme domain [Rodríguez-Aparicio et al. 1989], but to the hampering of the cross-talk between the bicarbonate phosphorylation domain and the NAG binding domain that results in a large increase in affinity for NAG when both ATP and bicarbonate are bound [Alonso and Rubio, 1983].

The decrease in enzyme activity caused by the p.Thr1443Ala mutation, (Fig. 4C) is accounted by a nearly 200-fold increase in the K_a^{NAG} and by a nearly 10-fold decrease in the apparent V_{max} (Table 3 and Fig. 2A). The localization of the NAG-binding domain justifies these effects if the mutation hampers NAG

RESULTS. Chapter 1

binding and the transmission of the regulatory signal from the NAG site to both phosphorylation domains. Indeed, the two partial reactions of the enzyme, which reflect the two phosphorylation steps, were undetectable in this mutant (results not shown).

The other three mutations examined here, p.Asn355Asp (c.1063A>G) and p.Tyr389Cys (c.1166A>G) which affect the Glnase-like domain and coexist in patient 1, and p.Ala1378Thr (c.4132G>A), which maps on the NAG-binding domain, appear to have too little an effect on enzyme activity or on kinetic parameters to justify the neonatal deficiency (Fig. 4C, Table 3). The largest changes observed with these mutations were a ~4-fold decrease in V_{\max} and a nearly 3-fold increase in the K_a for NAG, occurring with the p.Asn355Asp mutation (Table 3). However, thermal inactivation assays (Fig. 4D) revealed that these mutations substantially decreased the thermal stability of rhCPS1, particularly p.Asn355Asp, which lowered ~11°C the mid-inactivation temperature. In contrast, the temperature dependence of enzyme inactivation was identical for the two polymorphisms and for the wild-type enzyme (Fig. 4D). The combined effects of the decrease in V_{\max} and the modest increase in K_a^{NAG} with the p.Asn355Asp mutant, together with the decreased enzyme stability, may result in enzyme deficiency. This may also be the case with the p.Tyr389Cys mutation, which decreased ~40% enzyme activity (Fig. 4C) and caused a substantial, although less drastic effect on thermal inactivation. Finally, with p.Ala1378Thr, the deficiency could be due to the combination of the decreased stability and a twofold increase in apparent K_m for ATP (Table 3). The fact that these last two mutations are, respectively, only one and three positions away from Leu390 and Leu1381, two residues for which their p.Leu390Arg and p.Leu1381Ser mutations were found to cause dramatic loss of enzyme stability (see above), lends further support to the view that these mutations may hamper sufficiently enzyme stability "*in vivo*" to cause enzyme deficiency.

2.1.4.6 Influence of the substrates and of NAG on the resistance of human CPS1 to proteolytic or thermal inactivation

From all of the above, enzyme destabilization appears a crucial element in the causation of CPS1D with five of the eight missense mutations studied here (p.Asn355Asp, p.Tyr389Cys, p.Leu390Arg, p.Ala1378Thr and p.Leu1381Ser). Therefore, enzyme stabilization by ligands acting as chemical chaperones could be a useful future treatment of CPS1D. Similarly to the observation that chaperoning of phenylalanine hydroxylase by its essential cofactor tetrahydrobiopterin is clinically useful in phenylketonuria [Erlandsen et al.,

2004], the CPS1 substrates and particularly the CPS1 allosteric activator NAG and its pharmacological deacylase-resistant analog and orphan drug NCG could afford some protection of CPS1. We tested the effects of these ligands on the sensitivity of rhCPS1 to proteolytic inactivation by elastase. These experiments are based on previous data on the sensitivity of the rodent liver enzyme to elastase and on the effects on that sensitivity of enzyme ligands [Evans & Balon, 1988; Guadalajara et al., 1983; Marshall & Fahien, 1988; Powers-Lee & Corina, 1986]. Figs. 5 A and B reveal a strong influence of NAG and NCG on elastase inactivation of rhCPS1. Whereas these two compounds alone accelerated rhCPS1 inactivation in this system (Fig. 5A), when they were added together with MgATP, the enzyme became nearly entirely protected against even very large elastase concentrations (Fig. 5B). MgATP alone did not cause such large protection. Similarly, the combination of MgATP and NAG or MgATP and NCG was highly effective in protecting the enzyme from thermal inactivation (Fig. 5C). In this case MgATP also substantially protected the enzyme, although the highest degree of protection was attained with MgATP together with NAG or NCG. We also showed in these experiments that glycerol at a concentration (20%) at which it causes maximal NAG-independent CPS1 activation (Fig. 3D) also protected substantially the enzyme from thermal inactivation (Fig. 5C). Interestingly, the protecting effects of NAG/MgATP, NCG/MgATP, MgATP and glycerol were also patent when a mutant (p.Ala1378Thr) exhibiting reduced thermal stability was studied (Fig. 5D).

In the matrix of liver mitochondria, where CPS1 is localized, MgATP is likely to be abundant under most circumstances, but NAG may not, particularly under conditions of protein restriction as when a urea cycle deficiency is suspected. Therefore, under these circumstances the administration of NCG might help protect the enzyme from thermal inactivation or from proteolytic degradation. Therefore, studies on the effects of NCG on CPS1 stability “*in vivo*” are warranted. Indeed, a genetically demonstrated CPS1D patient has been documented to respond to NCG administration [Williams et al., 2010].

2.1.4.7 Final comments

rhCPS1 is shown here to mirror the natural human enzyme or the rat liver enzyme in all aspects analyzed, including substrate and activator kinetics, oligomeric form, domain composition, sensitivity to proteases and protection

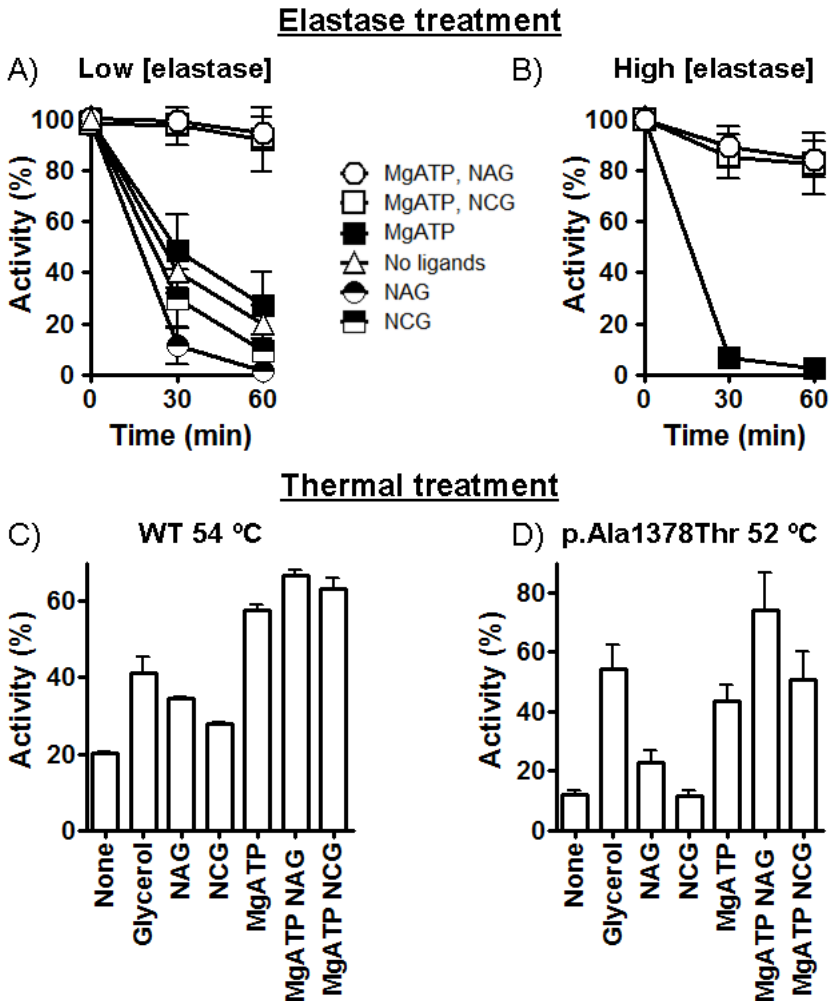


Figure 5. Effects of ligands on the inactivation of rhCPS1 by elastase or by heating. When indicated, NAG, NCG and ATP were added, at 10 mM concentrations. When ATP was added, MgSO₄ was also added at a concentration of 20 mM. Incubations were terminated by dilution in the continuous enzyme activity assay, monitoring ADP production at 37°C. Results are given as percentages of the activity not having undergone the corresponding proteolytic or heating treatment. A and B: Digestions of rhCPS1 with respective elastase concentrations of 10 µg/ml or 50 µg/ml. Other conditions were as in Fig. 1D. C and D: Thermal inactivation of the wild-type or the p.Ala1378Thr mutant forms of rhCPS1 (both used at 0.1 mg/ml concentrations) after 15-min incubation at the indicated temperatures in a solution of 50 mM glycyl-glycine pH 7.4, 20 mM KCl, 1 mM DTT with the indicated ligands or with 20% glycerol.

thereof by ligands, and ability of glycerol to replace NAG as a CPS1 activator. Our confirmation that the domain organization is that reported for other CPSs supports the present structural rationalizations of the effects of missense changes in *CPS1* [Häberle et al., 2011; Martínez et al., 2010]. Since recombinant production of hCPS1 permits the introduction of amino acid changes at will, the way is now open for testing the functional impact of missense mutations found in CPS1D. It would be desirable to compare the present system and even to complement it with the one using *Schyzosaccharomyces pombe* as the expression host of hCPS1 [Ahuja and Powers-Lee, 2008]. Although data on this system are very limited, the simultaneous use of both expression systems may prove desirable for maximizing the number of clinical mutants that can be tested. After all, both expression systems are heterologous with respect to hCPS1 and therefore the possibility cannot be excluded that some mutant forms can be expressed in one system but not in the other.

With the present system, clear-cut correlations between given missense mutations and specific molecular phenotypes can be established, hopefully shedding light on the degree of severity of different mutations. Indeed, the lack of detectable effects of the two polymorphisms studied here, and the severity of the effects demonstrated for six of the eight mutations analyzed, clearly indicate that the experimental studies with rhCPS1 mutants expressed “*in vitro*” identify disease-causing mutations. Even with the two mutations that caused the less drastic effects, they triggered some negative changes on enzyme activity and/or stability that were not observed with the two polymorphisms. In any case, the extension of the present pilot study to larger series of clinical CPS1D mutations should permit deeper ascertaining of the sensitivity of our approach to identify disease-causing mutations and for estimating their actual severity.

Our results shed also light on the role of the Glnase-like domain of CPS1 and on the reasons for its preservation despite the fact that CPS1 does not use glutamine [Rubio et al., 1981]. The observations that the amino acid substitutions p.Asn355Asp and p.Tyr398Cys, mapping in the Glnase-like domain of CPS1 (Fig. 1A), do not inactivate CPS1 or cause dramatic changes in K_m values for the substrates (Table 3), agrees with the general belief that this domain is not directly involved in the enzyme reaction. However, these mutations, as well as another two mutations (p.S123F and p.H337R) introduced previously in the N-terminal region of rat CPS1 [Pekkala et al., 2010] resulted in 40-75% reduction in enzyme activity (see for the present mutations Fig. 4C). These results also agree with a study [Ahuja and Powers-Lee, 2008] in which hCPS1 lacking the entire N-terminal region exhibited a 700-fold reduction in enzyme activity, although the very drastic change of deleting 25% of the

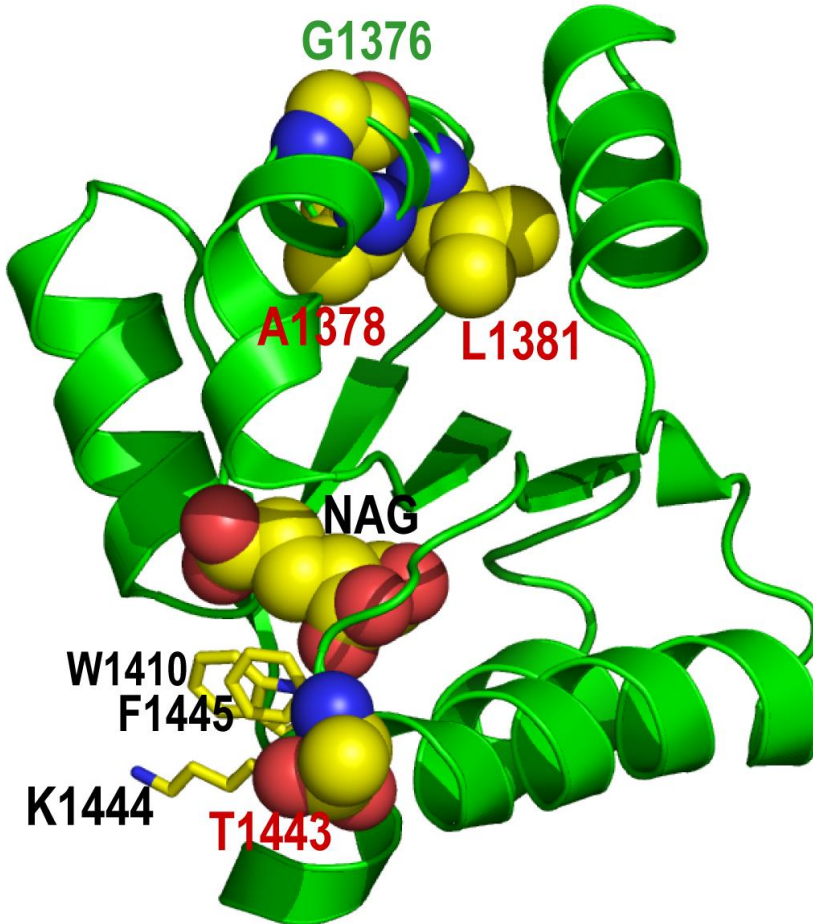
RESULTS. Chapter 1

protein molecule may render difficult the interpretation of this extreme degree of inactivation. In any case, the changes revealed by our present studies and by earlier studies with single amino acid substitutions affecting this CPS1 domain clearly support an activating role of this N-terminal region on CP synthesis, which is catalyzed by the C-terminal moiety of the enzyme [Cervera et al., 1993; Rubio, 1993]. Such activation appears reminiscent of the one caused by the small subunit of *E. coli* CPS on the catalysis by the large subunit of the reaction from ammonia [Meister, 1989]. The Glnase-like domain also stabilizes the enzyme, since the p.Leu390Arg, p.Asn355Asp and p.Tyr389Cys mutations decrease CPS1 stability. Again this role is reminiscent of the strong stabilization triggered in *E. coli* CPS upon association of the small and large subunits [Cervera et al., 1993]. The activation of CP synthesis, and the enhanced enzyme stability, may be sufficiently important advantages to warrant retention of the N-terminal region in CPS1.

Our present findings also shed light on the roles of the C-terminal domain. We localized "*in silico*" the NAG site in the crystal structure of this domain [Pekkala et al., 2009], providing as experimental support for this localization the results of photoaffinity labelling with N-chloroacetyl-L-glutamate and of site-directed mutagenesis of rat liver CPS1 [Pekkala et al., 2009]. We now provide even more direct proof for such localization with hCPS1, the enzyme for which the crystal structure of the C-terminal domain was determined. Thus, among all the mutations studied here, the one mapping closest to the proposed NAG site, p.Thr1443Ala, produces by far the most drastic decrease in the apparent affinity for NAG (two orders of magnitude decrease) (Table 3). In our proposed NAG binding site [Pekkala et al., 2009], Thr1443 is close to the bound activator, adjacent to a residue of the NAG site and to one of the three lid residues that cover the bound NAG molecule (Supp. Fig. S2). The importance of Thr1443 for effector regulation of CPS1 is highlighted also by the observation that phosphorylation of the equivalent residue in hamster CPSII (a component of CAD, the trifunctional enzyme involved in pyrimidine biosynthesis), Ser1406 [Simmer et al., 1990], hampers CPSII allosteric regulation by its negative effector UTP [Carrey et al., 1985]. As NAG in CPS1, UTP binds to the C-terminal domain of CPSII [Liu et al. 1994].

The present data also evidence that the C-terminal domain is an important determinant for CPS1 stability, since two mutations mapping in this domain, p.Ala1378Thr and p.Leu1381Ser, which affect the inner face of a helix from the outer layer of the $\alpha\beta\alpha$ sandwich conforming this domain (Supp. Fig. S2), substantially or very drastically destabilized the enzyme. This important impact on enzyme stability clearly supports a high degree of integration of the C-terminal domain in the CPS1 architecture. Thus, despite the multidomain

character of CPS1, the architecture of this enzyme would appear to be highly cooperative, explaining the influence of NAG binding on events that occur far away from the C-terminal domain, such as the activation of both phosphorylation domains [Rubio et al., 1983a].



Supplementary Figure S2. Location in the structure of the C-terminal domain of the amino acids whose substitutions are studied here. The protein is in ribbon representation. Bound NAG [Pekkala et al., 2009] and the amino acids that are replaced are represented in spheres and are labeled whereas the NAG site lid residues W1410, F1445 and K1444 are shown in sticks representation. Green and red labeling correspond, respectively, to the polymorphism and to clinical mutations. NAG, and the amino acids forming the NAG site lid, are labeled in black. In NAG and in the highlighted residues, C, N and O atoms are colored yellow, blue and red, respectively. This figure was prepared with PyMOL (<http://www.pymol.org>).

2.1.5 Acknowledgements

[These acknowledgements corresponds to the editorial office, not to the authors]

We thank Jackie Senior for editorial advice and Marian Stevens-Kroef for critical comments on earlier drafts of this manuscript.

Disclosure statement

The authors declare no conflict of interest

2.1.6 References

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods* 7:248-249.
- Ahuja V, Powers-Lee SG. 2008. Human carbamoyl-phosphate synthetase: insight into N-acetylglutamate interaction and the functional effects of a common single nucleotide polymorphism. *J Inherit Metab Dis* 31:481-491.
- Alonso E, Girbés J, García-España A, Rubio V. 1989. Changes in urea cycle-related metabolites in the mouse after combined administration of valproic acid and an amino acid load. *Arch Biochem Biophys* 272:267-273.
- Alonso E, Cervera J, García-España 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. 1983. Binding of N-acetyl-L-glutamate to rat liver carbamoyl phosphate synthetase (ammonia). *Eur J Biochem* 135:331-337.
- Amero SA, James TC, Elgin SC. 1994. Production of antibodies using proteins in gel bands. *Methods Mol Biol* 32:401-406.
- Barcelona-Andrés B, Marina A, Rubio V. 2002. Gene structure, organization, expression, and potential regulatory mechanisms of arginine catabolism in *Enterococcus faecalis*. *J Bacteriol* 184:6289-6300.
- Bradford MM. 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.
- Britton HG, Rubio V, Grisolia S. 1981. Synthesis of carbamoyl phosphate by carbamoyl phosphate synthetase I in the absence of acetylglutamate. Activation of the enzyme by cryoprotectants. *Biochem Biophys Res Commun* 99:1131-1137.

- Brusilow SW, Horwich AL. 2001. Urea cycle enzymes. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors; Child B, Kinzler KW, Vogelstein B, associated editors. *The Metabolic and Molecular Bases of Inherited Disease*, 8e. New York: McGraw-Hill. Vol 2, p 1909-1963.
- Carrey EA, Campbell DG, Hardie DG. 1985. Phosphorylation and activation of hamster carbamyl phosphate synthetase II by cAMP-dependent protein kinase. A novel mechanism for regulation of pyrimidine nucleotide biosynthesis. *EMBO J* 4:3735-3742.
- Cervera J, Conejero-Lara F, Ruiz-Sanz J, Galisteo ML, Mateo PL, Lusty CJ, 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.
- Eeds AM, Hall LD, Yadav M, Willis A, Summar S, Putnam A, Barr F, Summar ML. 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.
- Erlandsen H, Pey AL, Gámez A, Pérez B, Desviat LR, Aguado C, Koch R, Surendran S, Tyring S, Matalon R, Scriver CR, Ugarte, Martínez A, Stevens RC. 2004. Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. *Proc Natl Acad Sci USA*. 101:16903-16908.
- Evans DR, Balon MA. 1988. Controlled proteolysis of ammonia-dependent carbamoyl-phosphate synthetase I from Syrian hamster liver. *Biochim Biophys Acta* 953:185-196.
- Finckh U, Kohlschütter A, Schafer H, Spherhake K, Colombo JP, Gal A. 1998. Prenatal diagnosis of carbamoyl-phosphate synthetase I deficiency by identification of a missense mutation in *CPS1*. *Hum Mutat* 12:206-211.
- Guadalajara AM, Rubio V, Grisolia S. 1983. Inactivation of carbamoyl phosphate synthetase (ammonia) by elastase as a probe to investigate binding of the substrates. *Biochem Biophys Res Commun*. 117:238-244.
- Guthöhrlein G, Knappe J. 1968. Structure and function of carbamoylphosphate synthase. I. Transitions between two catalytically inactive forms and the active form. *Eur J Biochem* 7:119-127.
- Häberle J, Shchelochkov OA, Wang J, Katsonis P, Hall L, Reiss S, Eeds A, Willis A, Yadav M, Summar S, Urea Cycle Disorders Consortium, Lichtarge O, Rubio V, Wong LJ, Summar M. 2011. Molecular defects in human carbamoyl phosphate synthetase I: mutational spectrum, diagnostic and protein structure considerations. *Hum Mutat* 32:579-585.
- Häberle J, Boddaert N, Burlina A, Chakrapani A, Dixon M, Huemer M, Karall D, Martinelli D, Sanjurjo Crespo P, Santer R, Servais A, Valayannopoulos V, Lindner M, Rubio V, Dionisi-Vici C. 2012. Suggested Guidelines for the

RESULTS. Chapter 1

- Diagnosis and Management of Urea Cycle Disorders. *Orphanet J Rare Dis* 7:32.
- 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.
- Kurokawa K, Yorifuji T, Kawai M, Momoi T, Nagasaka H, Takayanagi M, Kobayashi K, Yoshino M, Kosho T, Adachi M, Otsuka H, Yamamoto S, Murata T, Suenaga A, Ishii T, Terada K, Shimura N, Kiwaki K, Shintaku H, Yamakawa M, Nakabayashi H, Wakutani Y, Nakahata T. 2007. Molecular and clinical analyses of Japanese patients with carbamoylphosphate synthetase 1 (CPS1) deficiency. *J Hum Genet* 52:349-354.
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, Mooney SD, Radivojac P. 2009. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 25: 2744-2750.
- Liu X, Guy HI, Evans DR. 1994. Identification of the regulatory domain of the mammalian multifunctional protein CAD by the construction of an *Escherichia coli* hamster hybrid carbamyl-phosphate synthetase. *J Biol Chem* 269:27747-27755.
- Marshall M, Cohen PP. 1972. Ornithine transcarbamylase from *Streptococcus faecalis* and bovine liver. I. Isolation and subunit structure. *J Biol Chem* 247:1641-1653.
- Marshall M, Fahien LA. 1985. Proximate sulfhydryl groups in the acetylglutamate complex of rat carbamylphosphate synthetase I: their reaction with the affinity reagent 5'-p-fluorosulfonylbenzoyladenine. *Arch Biochem Biophys* 241:200-214.
- Marshall M, Fahien LA. 1988. Proteolysis as a probe of ligand-associated conformational changes in rat carbamyl phosphate synthetase I. *Arch Biochem Biophys* 262:455-470.
- Martínez AI, 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.
- 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.
- Metzenberg RL, Marshall M, Cohen PP. 1958. Carbamyl phosphate synthetase: studies on the mechanism of action. *J Biol Chem* 233:1560-1564.
- Nyunoya H, Broglie KE, Widgren EE, Lusty CJ. 1985. Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. *J Biol Chem* 260:9346-9356.

- Pearson DL, Dawling S, Walsh WF, Haines JL, Christman BW, Bazyk A, Scott N, Summar ML. 2001. Neonatal pulmonary hypertension--urea-cycle intermediates, nitric oxide production, and carbamoyl-phosphate synthetase function. *N Engl J Med.* 344:1832-1838.
- Pekkala S, Martinez AI, 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 synthetase1. *Biochem J* 424:211-220.
- Pekkala S, Martínez AI, Barcelona B, Yefimenko I, Finckh U, Rubio V, Cervera J. 2010. Understanding carbamoyl-phosphate synthetase I (CPS1) deficiency by using expression studies and structure-based analysis. *Hum Mutat* 31:801-808.
- Pierson DL, Brien JM. 1980. Human carbamylphosphate synthetase I. Stabilization, purification, and partial characterization of the enzyme from human liver. *J Biol Chem* 255:7891-7895.
- Powers-Lee SG, Corina K. 1986. Domain structure of rat liver carbamoyl phosphate synthetase I. *J Biol Chem* 261:15349-15352.
- Rajzman L, Jones ME. 1976. Purification, composition, and some properties of rat liver carbamyl phosphate synthetase (ammonia). *Arch Biochem Biophys* 175:270-278.
- Rodriguez-Aparicio LB, Guadalajara AM, 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. 1993. Structure-function studies in carbamoyl phosphate synthetases. *Biochem Soc Trans* 21:198-202.
- Rubio V, Britton HG, Grisolia S. 1983a. 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, Britton HG, Grisolia S. 1983b. Activation of carbamoyl phosphate synthetase by cryoprotectants. *Mol Cell Biochem* 53-54:279-298.
- Rubio V, Cervera J, Lusty CJ, Bendala E, Britton HG. 1991. Domain structure of the large subunit of *Escherichia coli* carbamoyl phosphate synthetase. Location of the binding site for the allosteric inhibitor UMP in the COOH-terminal domain. *Biochemistry* 30:1068-1075.
- 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.
- Ryall J, Nguyen M, Bendayan M, Shore GC. 1985. Expression of nuclear genes encoding the urea cycle enzymes, carbamoyl-phosphate synthetase I and

RESULTS. Chapter 1

- ornithine carbamoyl transferase, in rat liver and intestinal mucosa. *Eur J Biochem* 152:287-292.
- Shih VE. 1976 Congenital hyperammonemic syndromes. *Clin Perinatol* 3:3-14.
- Simmer JP, Kelly RE, Rinker AG Jr, Scully JL, Evans DR. 1990. Mammalian carbamyl phosphate synthetase (CPS). DNA sequence and evolution of the CPS domain of the Syrian hamster multifunctional protein CAD. *J Biol Chem* 265:10395-10402.
- Summar ML. 1998. Molecular genetic research into carbamoyl-phosphate synthase I: molecular defects and linkage markers. *J Inherit Metab Dis* 21 Suppl 1:30-39.
- Summar ML, Dasouki MJ, Schofield PJ, Krishnamani MR, Vnencak-Jones C, Tuchman M, Mao J, Phillips JA 3rd. 1995. Physical and linkage mapping of human carbamyl phosphate synthetase I (*CPS1*) and reassignment from 2p to 2q35. *Cytogenet Cell Genet* 71:266-267.
- Summar ML, Hall LD, Eeds AM, Hutcheson HB, Kuo AN, Willis AS, Rubio V, Arvin MK, Schofield JP, Dawson EP. 2003. Characterization of genomic structure and polymorphisms in the human carbamyl phosphate synthetase I gene. *Gene* 311:51-57.
- Thoden JB, Holden HM, Wesenberg G, Raushel FM, Rayment I. 1997. Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. *Biochemistry* 36:6305-6316.
- Williams M, Huijmans JGM, van Diggelen OP, van der Low EJTM, de Klerk JBC, Haerberle J (2010) Carbamoyl phosphate synthetase I (CPS 1) deficiency: treatment with carnitine (Carbaglu). *J Inherit Metab Dis* 33 (Suppl 1):S118
- Yefimenko I, Fresquet V, Marco-Marín C, Rubio V, Cervera J. 2005. Understanding carbamoyl-phosphate synthetase deficiency: impact of clinical mutations on enzyme functionality. *J Mol Biol* 349:127-141.

RESULTS

2.2 CHAPTER 2

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

by

Carmen Díez-Fernández^a, Liyan Hu^b, Javier Cervera^{a,c}, Johannes Häberle^{b,*} and Vicente Rubio.^{a,c,*}

^aInstituto de Biomedicina de Valencia of the CSIC, Valencia, Spain; ^bUniversity Children's Hospital Zurich and Children's Research Center, Zurich, Switzerland; ^cGroup 739 of the Centro de Investigación Biomédica en Red sobre Enfermedades Raras (CIBERER) del Instituto de Salud Carlos III, Spain

Published in: Molecular Genetics and Metabolism (2014) 112:123-32

*Correspondence to:

J. Häberle
Kinderspital Zürich
Steinwiesstrasse 75
8032 Zürich, Switzerland
Phone: +41 442667342.
Fax: +41 442667167
johannes.haerberle@kispi.uzh.ch

V. Rubio
Instituto de Biomedicina de Valencia (IBV-CSIC)
Jaume Roig 11
46010 Valencia, Spain
Phone: +34 96 3391772.
Fax: +34 96 3690800
rubio@ibv.csic.es

2.2.1 Abstract: Carbamoyl phosphate synthetase 1 deficiency (CPS1D) is an inborn error of the urea cycle that is due to mutations in the *CPS1* gene. In the first large repertory of mutations found in CPS1D, a small CPS1 domain of unknown function (called the UFSD) was found to host missense changes with high frequency, despite the fact that this domain does not host substrate-binding or catalytic machinery. We investigate here by *in vitro* expression studies using baculovirus/insect cells the reasons for the prominence of the UFSD in CPS1D, as well as the disease-causing roles and pathogenic mechanisms of the mutations affecting this domain. All but three of the 18 missense changes found thus far mapping in this domain in CPS1D patients drastically decreased the yield of pure CPS1, mainly because of decreased enzyme solubility, strongly suggesting misfolding as a major determinant of the mutations negative effects. In addition, the majority of the mutations also decreased from modestly to very drastically the specific activity of the fraction of the enzyme that remained soluble and that could be purified, apparently because they decreased V_{\max} . Substantial although not dramatic increases in K_m values for the substrates or for N-acetyl-L-glutamate were observed for only five mutations. Similarly, important thermal stability decreases were observed for three mutations. The results indicate a disease-causing role for all the mutations, due in most cases to the combined effects of the low enzyme level and the decreased activity. Our data strongly support the value of the present expression system for ascertaining the disease-causing potential of CPS1 mutations, provided that the CPS1 yield is monitored. The observed effects of the mutations have been rationalized on the basis of an existing structural model of CPS1. This model shows that the UFSD, which is in the middle of the 1462-residue multidomain CPS1 protein, plays a key integrating role for creating the CPS1 multidomain architecture leading us to propose here a denomination of "Integrating Domain" for this CPS1 region. The majority of these 18 mutations distort the interaction of this domain with other CPS1 domains, in many cases by causing improper folding of structural elements of the Integrating Domain that play key roles in these interactions.

Key words: urea cycle diseases; CPS1 deficiency; hyperammonemia; inborn errors; CPS 1 structure; site-directed mutagenesis

2.2.2 Introduction

Primary CPS1 deficiency (CPS1D; MIM #237300), a recessively inherited urea cycle disease leading to frequently fatal hyperammonemia [1,2], is due to mutations in the *CPS1* gene. This gene, located in 2q35 [3] and being composed of 38 exons and 37 introns [4-6], with 4,500 coding nucleotides, encodes a 1500-residue proenzyme [7] that is synthesized in hepatocytes and enterocytes [8,9], and which, upon internalization to the mitochondrial matrix, yields after cleavage of its N-terminal 38 amino acids, the mature 1,462-amino acid multidomain (Fig. 1A) CPS1 protein [10-12] [E.C. 6.3.4.16].

CPS1 catalyzes the first step of the urea cycle ($2\text{ATP} + \text{NH}_3 + \text{HCO}_3^- \rightarrow 2\text{ADP} + \text{HPO}_4^{2-} + \text{NH}_2\text{CO}_2\text{PO}_3^{2-}$) [13], converting ammonia to carbamoyl phosphate (CP), a compound that is utilized by ornithine transcarbamylase (OTC) to make citrulline in the second reaction of the urea cycle. The three-step CPS1 reaction (Fig. 1B) includes two analogous ATP-dependent phosphorylations, of bicarbonate and carbamate, and an intervening step of carbamate synthesis from carboxyphosphate and ammonia [13]. To be active, CPS1 requires the presence of an essential allosteric activator, N-acetyl-L-glutamate (NAG) [10,14,15], made by NAG synthase from glutamate and acetyl-coenzyme A [16]. The rate of NAG synthesis heavily depends on glutamate concentration [17], and therefore NAG represents a switch for CPS1 activity, which is turned off when glutamate levels decrease, thus preventing excessive nitrogen draining by the urea cycle from an already low amino acid pool [18,19]. NAG activation is a unique property of CPS1 (and to a lesser extent of the piscine CPS1 homologue CPSIII) [20], not being shared by other CPSs, all of which are active in the absence of effectors and are insensitive to NAG [13,21,22]. NAG activation involves a functionally crucial cross-talk between NAG and phosphorylation sites. Thus, NAG vastly increases the apparent affinity of CPS1 for ATP and for its ionic activators K^+ and Mg^{2+} [15] and, conversely, ATP, K^+ and Mg^{2+} greatly increase CPS1 affinity for NAG [23].

The complexities of the CPS1 reaction and regulation are backed by a sophisticated multidomain protein machinery [13], that is still imperfectly characterized, since the crystal structure of CPS1 has not been determined (except for its ~15 kDa C-terminal domain, representing only ~10% of the entire molecule) [24,25]. Nevertheless, a number of approaches including the use of limited proteolysis [26] support a domain composition for CPS1 that mirrors that of *Escherichia coli* CPS, the only CPS for which the structure is known [27], notwithstanding the fact that these two enzymes have very important differences [13,21]. These differences include the chain composition (a single chain in CPS1; two subunits in *E. coli* CPS), the already indicated

RESULTS. Chapter 2

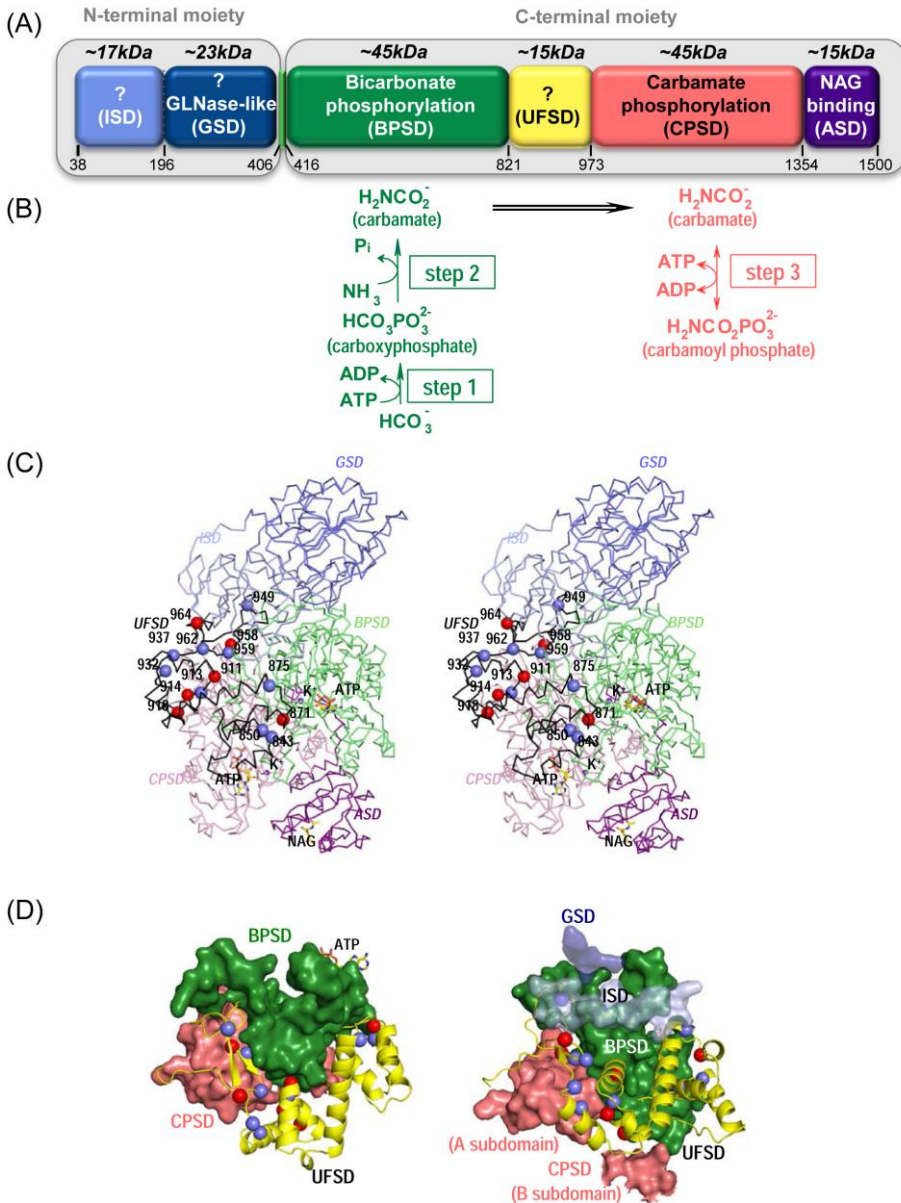


Figure 1. Domain composition, function and architecture of CPS1. (A) Schematic linear representation in the CPS1 polypeptide of the different domains. The grey-shaded background bars schematize the 40-kDa N-terminal and the 120-kDa C-terminal moieties of the enzyme that correspond to the small and large subunits of *E. coli* CPS, respectively. The colored bars represent the different (cont'd next page)

differences in the requirement for NAG, and the fact that CPS1 uses ammonia with high efficiency and cannot utilize glutamine as an ammonia source, whereas bacterial CPS uses ammonia poorly but utilizes glutamine as an internal ammonia source. The small subunit of the bacterial enzyme binds and cleaves the glutamine, channeling the resulting ammonia to the large subunit, where the entire carbamoyl phosphate synthesis reaction from ammonia takes place [21].

As already mentioned, limited proteolysis and other studies support the similarity in domain composition and function of bacterial CPS and CPS1 (Fig. 1A). These studies evidenced [13,26,28] that CPS1 is composed of N- and C-terminal moieties of, respectively, ~40 and ~120 kDa, that correspond to the small and large subunits of *E. coli* CPS [12]. Since CPS1 uses no glutamine [10], no functions for the N-terminal ~40-kDa moiety are known other than possibly some enzyme stabilization and activation [26,29]. In contrast, the ~120 kDa moiety is known to host the catalytic and NAG-regulatory machinery [13,21], being composed of two ~60-kDa halves, each one of them consisting of an N-terminal ~45-kDa phosphorylation domain followed by a ~15 kDa domain

Fig. 1, cont'd: *CPS1 domains as defined from sequence alignment with E. coli CPS. Their approximate masses, in kDa, and the domain start/end residue numbers are given, respectively, above and below each domain. The N-terminal mitochondrial targeting peptide that is removed upon internalization in the organelle is not represented. The domains are defined by their function (when known; ?, function unknown), showing below each of them the corresponding acronym (ISD, GSD, BPSD, UFSD, CPSD and ASD) as defined in [2]. (B) Reactional steps of the CPS1 reaction, shown under the domain catalyzing it, colored as the domain. The thick empty horizontal black arrow denotes the migration of carbamate from the bicarbonate phosphorylation site to the carbamate phosphorylation site. (C) Stereo view of structural model of C α trace of human CPS1 [42]. Each domain is colored differently and labeled in the same color. In the UFSD, the spheres mark the residues (numbered) hosting extremely drastic (red) or less drastic (blue) mutations found in CPS1D patients. Both ATP molecules and essential K⁺ ions are placed in their sites by superimposition of the E. coli CPS structure (Protein Databank file IBXR; [48]). (D) Relations of the UFSD with other domains. Two views of the domain (cartoon representation, in yellow) are shown, highlighting the residues hosting the mutations (red and blue spheres as above) illustrating that the mutations cluster at the regions of contact with the other domains (labeled). Only the parts of these other domains that contact the UFSD are shown, in surface representation, colored as in (A). In the right panel the ISD is shown as transparent surface to allow visualization of Ala949, which is clamped between the ISD, the GSD and the BPSD.*

RESULTS. Chapter 2

[13]. Both phosphorylation domains are homologous [12]. The one in the N-half phosphorylates bicarbonate and the other in the C-half phosphorylates carbamate (Fig. 1B) [30,31]. The C-terminal ~15-kDa domain (residues 1354-1500) of the C-half (the C-terminal domain of the enzyme) binds NAG and is crucial for activation, having been called the allosteric domain (ASD) [25,26,32], not hosting substrate sites or catalytic machinery. The other ~15 kDa domain (precise mass, 17.7 kDa; residues 822-973) connects both phosphorylation domains and its function is unclear since it has no substrate sites or catalytic components, having been called unknown function subdomain (UFSD) [26]. Interestingly, the UFSD has been found to host missense mutations in CPS1D patients (Fig. 1C,D) with rather high frequency [2], suggesting an important and until now unclear role of this domain.

CPS1D has been associated with a relatively large number of different, generally "private" mutations, which occur in single families with very little recurrence [2]. Many of the >130 missense mutations reported in CPS1D remain to be proven responsible for the deficiency. The fact that they appear to be distributed non-homogeneously among the different gene exons even after correcting for the presence of CpG islands, suggests that some enzyme regions have a more important role on enzyme stability, folding or functionality than other regions, and that, therefore, mutations falling on these more important regions have higher repercussion than those falling in less critical regions of the protein. The UFSD domain, particularly its C-terminal half, may be such an important region since the relative occurrence of missense mutations (normalized per 100 nucleotides) in exons 22 and 23, which encode most of the C-terminal half of the UFSD and which have no CpG islands, approximately doubles the relative frequency of these mutations for the entire coding sequence of the enzyme ($p < 0.01$; χ^2 test) [2].

We exploit here our recent ability to produce pure mature human CPS1, either wild-type or with the desired mutations, in a baculovirus/insect cell system [26], to examine the disease-causing potential of all known UFSD missense mutations ($n=18$) found in CPS1D patients [2,4,33-37] and to clarify the role and importance of the UFSD. When expression was possible, we studied the properties of the purified mutant enzyme forms, comparing them with wild-type human CPS1. Our findings support the disease-causing role of the mutations reported to affect the UFSD, revealing a key role of the UFSD for proper enzyme folding and for the regulatory cross-talk between NAG and phosphorylation sites.

2.2.3 Materials and Methods

2.2.3.1 Patients and CPS1 Mutations

Table 1 lists 18 amino acid substitutions found in patients with CPS1D, which are localized in the UFSD, giving some detail on the type of the clinical presentation in the patients carrying these mutations. Seventeen of the mutations were reported already [2,4,33-37], whereas one (mutation 13) is a novel change identified by one of us (J.H.) in a late onset CPS1D patient. We also report a patient with a neonatal presentation carrying an already reported mutation (p.Asp914His), a mutation for which there was no clinical information on the previously reported patient. In contrast, the new patient was known to have a neonatal presentation, indicating that the mutation was probably severe. The missense mutations in these two new patients were initially identified by mRNA studies in phytohemagglutinin-stimulated lymphocytes and then were confirmed by studies on genomic DNA as previously reported [37]. Samples were obtained with full informed consent of those entitled to give it, to perform molecular genetic diagnostics for clinical purposes. The PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>) [38] and MutPred (<http://mutpred.mutdb.org/>) servers [39] were used to assess in silico the disease-causing potential of the different mutations. Amino-acid conservation was determined by ClustalW sequence alignment of either CPS1, CPSIII or other CPSs from 14, 6 and 24 species, respectively.

2.2.3.2 Production of CPS1 carrying the desired mutation

The indicated mutations were introduced into pFastBac-CPS1 [26] using the overlapping extension method (Quickchange kit from Stratagene) and the forward and reverse primers given in Supp. Table S1. The correctness of the constructs, the presence of the desired mutation, and the absence of unwanted mutations were corroborated by sequencing.

Human mature liver CPS1, either wild-type or carrying the desired mutation, with the N-terminal mitochondrial targeting sequence replaced by the 28-residue N-terminal His₆-tag, MSYYHHHHHHDYDIPTTENLYFQGAMDP, was expressed in a baculovirus/insect cell system and purified as previously described [26]. The same purification procedure (cell centrifugation, cell lysis, centrifugal clarification of the extract, Ni-affinity chromatography and centrifugal ultrafiltrative concentration [26]) was used for wild-type CPS1 and for the mutant enzyme forms. Protein in the final enzyme preparation was determined according to Bradford [40]. The CPS1 yield was determined from

RESULTS. Chapter 2

Table S1. Synthetic oligonucleotides used for site-directed mutagenesis

Mutation	Direction	Sequence 5'-3'^a
p.L843S	Forward	CTTAGAAAAGAG T CGTCTGAACC
p.L843S	Reverse	GGTTCAGAC G ACTCTTTTCTAAG
p.R850C	Forward	CCAAGCAGCACG T GATCTATGCCATTG
p.R850C	Reverse	CAATGGCATAGATA C ACGTGCTGCTGG
p.R850H	Forward	GAACCAAGCAGCACG C ATATCTATGCCATTG
p.R850H	Reverse	CAATGGCATAGATA T GCGTGCTGCTGGTTC
p.T871P	Forward	GATTGAGAAGCTCC C ATACATTG
p.T871P	Reverse	CAATGTAT G GGAGCTTCTCAATC
p.K875E	Forward	GCTCACATACATTGACG A GTGGTTTTTG
p.K875E	Reverse	CAAAAACCA C TGTCATGTATGTGAGC
p.G911E	Forward	CAAAGGAGATT G AGTTCTC
p.G911E	Reverse	GAGAA C TCAATCTCCTTTG
p.G911V	Forward	CAAAGGAGATT G TGTTCTC
p.G911V	Reverse	GAGAA C ACAATCTCCTTTG
p.S913L	Forward	GATTGGGTT C TTAGATAA
p.S913L	Reverse	TTATCT A AGAACCCAATC
p.D914H	Forward	GGGTCTCACATAAGCAGATT T C
p.D914H	Reverse	GAAATCTGCTT A TGTGAGAACCC
p.D914G	Forward	GGGTCTCAG G TAAGCAGATTTC
p.D914G	Reverse	GAAATCTGCTT A CTGAGAACCC
p.S918P	Forward	GATAAGCAGATT C AAAATGCC
p.S918P	Reverse	GGCATT T TGGAATCTGCTTATC
p.R932T	Forward	CAGACAAGGGAGCTG A CGTTAAAG
p.R932T	Reverse	CTTTAA C GTCAGCTCCCTTGCTG
p.I937N	Forward	GTTAAAGAAAA A CAACCC
p.I937N	Reverse	GGGT G TTGTTTTTCTTTAAC
p.A949T	Forward	GATACACTGGCT A CAGAATAC
p.A949T	Reverse	GTATT C TGTAGCCAGTGTATC
p.L958P	Forward	AAACTATCCCTATGTTAC C TAC
p.L958P	Reverse	GTAGGTAACATAG G GATAGTTT
p.Y959C	Forward	CAAAC T ATCTCTGTGTTACCTAC
p.Y959C	Reverse	GTAGGTAAC A CAGAGATAGTTTG
p.Y962C	Forward	CTATGTTAC C TGCAATGGTC
p.Y962C	Reverse	GAC C ATTGCAGGTAACATAG
p.G964D	Forward	GTTACCTACAAT G ATCAGGAGC
p.G964D	Reverse	GCTCCTG A TCATTGTAGGTAAC

^aBold type indicates base substitutions to introduce the desired mutation

the total amount of protein in the enzyme preparation and the fraction of the protein corresponding to pure CPS1. The latter fraction was determined densitometrically from the band of approximately 163 kDa observed in Coomassie-stained 8% polyacrylamide gels after SDS-PAGE [41]. For this

purpose, stained gel images were collected with a Fujifilm LAS-3000 Imager using transmitted white light, and the fraction of the protein migrating in the CPS1 band of ~163 kDa was determined with the Multi Gauge quantification software (from Fuji Film).

2.2.3.3 Enzyme activity assays

CPS1 activity was assayed at 37°C as carbamoyl phosphate production, monitored as citrulline by using OTC [26]. The standard assay mixture contained 50 mM glycyl-glycine pH 7.4, 70 mM KCl, 1 mM dithiothreitol, 20 mM MgSO₄, 5 mM ATP, 35 mM NH₄Cl, 50 mM KHCO₃, 10 mM NAG, 5 mM L-ornithine and 4 U/ml OTC. Specific activities refer to the pure enzyme (determined densitometrically, see section 2.2). When the concentration of a substrate was varied, other substrates were kept at the concentrations given above, with MgSO₄ being in 20 mM excess over ATP. Data for variable substrate or NAG concentration were fitted to hyperbolae using the GraphPad Prism program (GraphPad Software, San Diego, CA). One enzyme unit makes 1 μmol citrulline per minute. Specific activities refer to pure CPS1, estimated by densitometry (see above).

2.2.3.4 Other assays

The thermal stability of CPS1 was monitored by incubating for 15 min at the specified temperature 0.5 mg protein/ml of the indicated CPS1 form (either wild-type or mutant), in a solution containing 50 mM glycyl-glycine pH 7.4, 2 mM dithiothreitol, 10% glycerol, 0.5 M NaCl and 20 mM imidazole. At the end of the incubation the solution was rapidly cooled at 0°C and enzyme activity was determined immediately in the standard assay at 37°C.

Western blotting was carried out after SDS-PAGE (section 2.2) as reported [26] using immunoluminescent detection (Super Signal West Pico Chemiluminiscent Substrate, Thermo Scientific).

The previously reported [42] atomic structure model of CPS1 (based on the experimental crystal structure of *E. coli* CPS) was kindly provided by B. Barcelona (Instituto de Biomedicina de Valencia) and is used here for structural analysis of the mutations effects. Pymol (DeLano Scientific; <http://www.pymol.org>) was used for visual analysis, for structural superimposition and for depicting protein structures.

Table 1
CPS1 missense mutations mapping in the UFSD, found in CPS1D patients.

Mutation #	Amino acid change ^a	Presentation	Report	Amino acid in CPS			PolyPhen-2 prediction	MutPred prediction
				I	III	Other		
1	p.Leu843Ser ^b	Neonatal	[4]	L/M	L	L/A	Probably damaging	0.92
2	p.Arg850Cys	Neonatal	[35]	R	R	R	Probably damaging	0.98
3	p.Arg850His	Neonatal	[2,33,37]				Probably damaging	0.98
4	p.Thr871Pro	Late, 2nd allele: p.E1194D	[37]	T	T	T/S	Probably damaging	0.86
5	p.Lys875Glu ^b	Neonatal	[4]	K	K	Variab.	Possibly damaging	0.84
6	p.Gly911Glu	Neonatal	[2]	G	G	G	Probably damaging	0.87
7	p.Gly911Val	Neonatal	[34]				Probably damaging	0.97
8	p.Ser913Leu	Neonatal	[2]	S	S	S/a/g/d	Probably damaging	0.79
9	p.Asp914His	Patient 1; Unknown	[2]	D	D	D	Probably damaging	0.95
10	p.Asp914Cly	Patient 2 ^c ; Neonatal Hyperammonemia, seizures, Homozygous ^d	Present [2]				Probably damaging	0.92
11	p.Ser918Pro	Neonatal	[33]	S/G	G	A/s/e	Benign	0.93
12	p.Arg932Thr	Neonatal	[2]	R	R	R/k	Probably damaging	0.93
13	p.Ile937Asn	Late ^e	Present	I	I	I/V/L	Probably damaging	0.83
14	p.Ala949Thr	Late, ^a 2nd allele: p.Y89 [*]	[2]	A	A	A/G	Probably damaging	0.89
15	p.Leu958Pro	Neonatal	[34]	L	L	L/M/F/Y	Probably damaging	0.94
16	p.Tyr959Cys	Unknown, Alive ^d	[2]	Y	Y	Y	Probably damaging	0.94
17	p.Tyr962Cys	2nd allele: p.P1462R Unknown, Alive ^d	[2]	Y	Y	Y/f	Probably damaging	0.92
18	p.Gly964Asp	Neonatal	[36]	G	G	Variab.	Probably damaging	0.37

Unless indicated below, cDNA changes are to be found in [4,5,7]. Amino acids are shown in three-letter code for the mutant allele found in the UFSD, and in one-letter code for the second allele and for occurrence in the various CPSs, being in low case when found with low frequency. *Variab.* denotes the occurrence at a given position, in the indicated groups of CPSs, of >4 types of amino acids with no constant chemical characteristics (polar, apolar, charged, etc.). PolyPhen-2 grades the probability of a damaging effect of an amino-acid substitution, from higher to lower, as *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.

^a Translation 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.
^b The changes p.Leu843Ser and p.Lys875Glu coexisted within the same allele.
^c Novel patient. The cDNA change is c.2740G > C, the same found in the previously reported patient with this amino acid change [4].
^d Unknown when originally reported. New data gathered on the patient.
^e Novel patient. The cDNA change is c.2810 T > A. The changes p.Ile937Asn and p.Gly401Arg (c.1201G > C) coexisted within the same allele. The second allele carries the mutation c.3337-3 T > A/c.3337-2A > T, which should lead to exon 27 skipping, frameshift and premature termination.

2.2.4 Results

2.2.4.1 CPS1D mutations that affect the UFSD

Table 1 summarizes the 17 already reported [2,4,33-37] mutations reported in CPS1D patients that affect the UFSD, plus one more mutation (#13) falling in this domain which is reported here in a late onset patient. Thirteen of these mutations were found in patients with neonatal or with very severe presentations and thus might be expected to have drastic negative effects on CPS1 level or activity. Only mutations 4, 13, 14, 16 and 17 were found exclusively in late onset patients or in patients without appropriate clinical information to judge about their severity. The fact that two mutations affecting the same residue were reported for Arg850, Gly911 and Asp914, three residues that are invariant among all CPSs, strongly supports a damaging role of the corresponding six mutations (p.Arg850Cys, p.Arg850His, p.Gly911Glu, p.Gly911Val, p.Asp914His, and p.Asp914Gly), in agreement with the predictions of the Polyphen-2 and MutPred servers (Table 1). In another patient with a neonatal presentation, the p.Leu843Ser and p.Lys875Glu mutations, both mapping in the UFSD, coexisted in the same allele [4]. Of these, the p.Leu843Ser mutation appears more likely to be disease-causing, because of the high conservation of Leu843 and because Polyphen-2 and MutPred give p.Leu843Ser a higher degree of probability and a higher score for having a damaging role than they do for p.Lys875Glu. The high residue conservation and the Polyphen-2 and MutPred server predictions also support the disease-causing nature of the p.Arg932Thr and p.Leu958Pro mutations, found in another two patients with neonatal presentations, whereas a disease-causing role appears less certain for the p.Ser913Leu, p.Ser918Pro and p.Gly964Asp mutations, found in other neonatal patients, since the residues affected are less conserved and the Polyphen-2 and MutPred servers are non-unanimous in anticipating a disease-causing role (Table 1). The lack of unanimity in the predictions by both servers is possible, since they do not utilize the same traits [38,39] for assessing the disease-causing role of a given mutation.

The mutations p.Thr871Pro, p.Ile937Asn, p.Ala949Thr, p.Tyr959Cys and p.Tyr962Cys, reported in CPS1D patients with late presentations or with unreported presentation but with patient survival (suggesting less severe phenotype), affect invariant or virtually invariant residues and give high probability scores for being damaging in the Polyphen-2 and MutPred servers. Since the p.Ala949Thr mutation coexists with a null allele (p.Tyr89*) in the patient [2], it cannot abolish enzyme activity since otherwise the patient would have had a neonatal presentation. In principle, the same can be said for the p.Ile937Asn mutation, given the fact that the other *CPS1* allele found in the

RESULTS. Chapter 2

same patient corresponds to a very severe splice site aberration that should lead to exon 27 skipping, with frameshift and truncation. Since in this patient the *CPS1* allele hosting the p.Ile937Asn mutation also hosts a second missense change, p.Gly401Arg (Table 1), this second change cannot be expected to be highly deleterious either. In contrast, a drastic effect is likely for the p.Thr871Pro mutation since it coexists in the patient with the expectedly mild mutation (given the mildness of the substitution), p.Glu1194Asp [37].

2.2.4.2 UFSD mutations negatively influence the yield of CPS1

We previously used the baculovirus/insect cell system utilized here for producing twenty CPS1 forms carrying different missense changes identified in CPS1D patients and spread all over the CPS1 protein except the UFSD [26,43]. We also reported [25] the effects of another five non-CPS1D-associated but functionally important mutations affecting the C-terminal domain (ASD). Furthermore, in unpublished studies, we have monitored the effects of ten more mutations (nine in the ASD) of which five had been identified in patients. Only three among these 35 experimentally tested mutations compromised drastically recombinant CPS1 production ([26], and data not shown). The present UFSD mutations yield an entirely different picture (Fig. 2A). Recombinant production of CPS1 protein was virtually abolished by six mutations and it was decreased 85-95% and ~75% by five and four mutations, respectively. Only for three of the eighteen mutations studied here (p.Arg850Cys, p.Arg850His and p.Ala949Thr) the yield of pure CPS1 was comparable to that of the wild-type enzyme. Interestingly, each one of the two mutations found in the same allele of one patient, p.Leu843Ser and p.Lys875Glu (Table 1), reduced the yield of CPS1 by nearly 90% (Fig. 2A), raising the possibility that their joint presence in the same protein molecule might reduce further CPS1 production.

The low yields observed with most mutants appear to be due to poor CPS1 polypeptide production and to gross misfolding. Thus, SDS-PAGE of the pellet (Fig. 2B, *P*) and the supernatant (Fig. 2B, *S*) obtained by centrifugation of the initial cell extract not only revealed for most mutants decreased CPS1 levels, but also showed that the CPS1 present tended to appear in the precipitate. Furthermore, a multiplicity of bands of lower molecular weight than complete CPS1, corresponding to CPS1 degradation products, were revealed by western blotting in the precipitates of these low-yield mutants (Fig. 2C).

In summary, the majority of the mutations found in CPS1D patients that affect the UFSD appear to disturb CPS1 folding and to enhance CPS1 degradation. The drastic decrease in enzyme production should be disease-causing at least

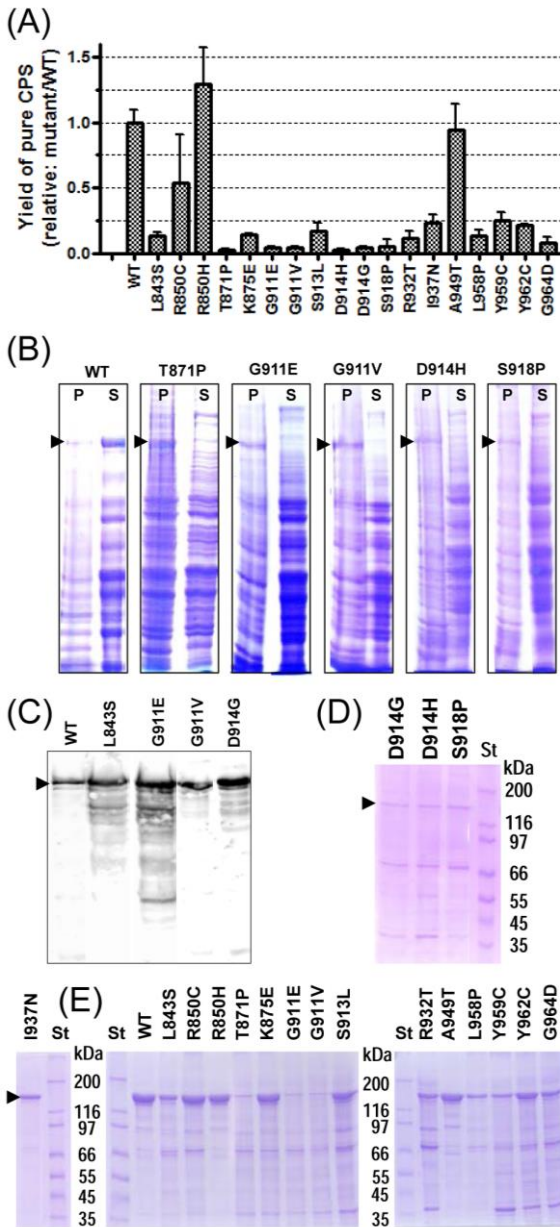


Figure 2. Effects of UFSD domain mutations on CPS1 production. WT, wild-type recombinant enzyme. Mutations are represented in single-letter code. Arrowheads signal the position of the CPS1 band. (A) Yield of pure CPS1 protein (total protein \times CPS1 purity determined densitometrically from Coomassie-stained SDS) per liter of cell culture, relative to the mean yield of pure wild-type CPS1 (2.7 mg/L cell culture). (B) Distribution of protein between the insoluble, P, and the soluble, S, fractions of the initial extracts of cells expressing the wild-type enzyme or the indicated mutants, analyzed by SDS-PAGE and Coomassie staining. The samples applied to each track correspond to identical volumes of original cell extract. (C) Western blotting and CPS1 immunostaining in the precipitates from the cell extracts for the indicated enzyme forms. (D, E) Coomassie-stained SDS-PAGE of the preparations obtained after purification of the wild-type and the different mutant enzyme forms as indicated. St, protein markers, with masses, in kDa, indicated at the side.

for the seven mutations that reduce enzyme production by >90% (p.Thr871Pro, p.Gly911Glu, p.Gly911Val, p.Asp914His, p.Asp914Gly, p.Ser918Pro and p.Gly964Asp; Fig. 2A). Indeed, six of these mutations were found in patients

RESULTS. Chapter 2

with neonatal or very severe presentations, and the other mutation (p.Thr871Pro), although found in a late onset patient, coexisted with an expectedly mild second allele and is likely to have by itself a severe effect (Table 1). On the other hand, the two mutations found in late onset patients that were anticipated to be mild (see section 3.1), p.Ala949Thr and p.Ile937Asn, were associated with substantial or normal yield of CPS1 (Fig. 2A).

2.2.4.3 Eight UFSD mutations strongly impair CPS1 activity

Although with most mutations little soluble CPS1 was produced, we always carried out the complete CPS1 purification protocol, observing by SDS-PAGE of the final preparation, with all the mutants (Figs. 2D,E), a 163 kDa-band corresponding to soluble CPS1 (confirmed by western blotting, data not shown). However, the purity of CPS1 decreased with the yield, being <10% of the protein (densitometric estimation) for the mutants with the lowest yield (Figs. 2D,E). Despite the presence of the CPS1 band, CPS1 activity was very low in the final preparations of the mutants p.Thr871Pro, p.Gly911Glu, p.Gly911Val, p.Asp914His, p.Asp914Gly, p.Ser918Pro, p.Leu958Pro and p.Gly964Asp. The estimated specific activity of pure soluble CPS1 was $\leq 6\%$ of wild-type for six of these mutations and $\leq 12\%$ for the other two (Fig. 3A). Assuming that the mutation-induced misfolding causes similar decreases in the production of soluble CPS1 protein in insect cells and in the liver, the low yields (Fig. 2A), combined with the decreased activities (Fig. 3A), lead to an expectation of very low residual activities in the liver (<1% of normal for homozygosity) for any of these eight mutant forms. The low residual activity further supports the disease-causing nature of these mutations and agrees with their observation in neonatal or very severely affected patients (Table 1). Only the p.Thr871Pro mutation was observed in a late onset patient, but, as already indicated, the second allele in this patient was expectedly mild (p.Glu1194Asp).

2.2.4.4 Activity and stability changes with another eight UFSD mutations

Eight of the ten mutations not dealt with in section 3.3 generally were associated with less drastic but substantial decreases in specific activity, and/or with decreased thermal stability of the enzyme (Fig. 3). Among these mutations, p.Ile937Asn caused the largest specific activity decrease (~90% decrease) (Fig. 3A), which combined with the observed 75% reduction in yield (Fig. 2A), would result (making the already indicated assumption that the yield in insect cells grossly corresponds to the yield in the liver) in ~3% residual activity for homozygosity, supporting disease-causation. Nevertheless, the role

of the other mutation found in the same allele (Table 1), p.Gly401Arg, remains to be determined, although this mutation maps in the CPS1 N-terminal region

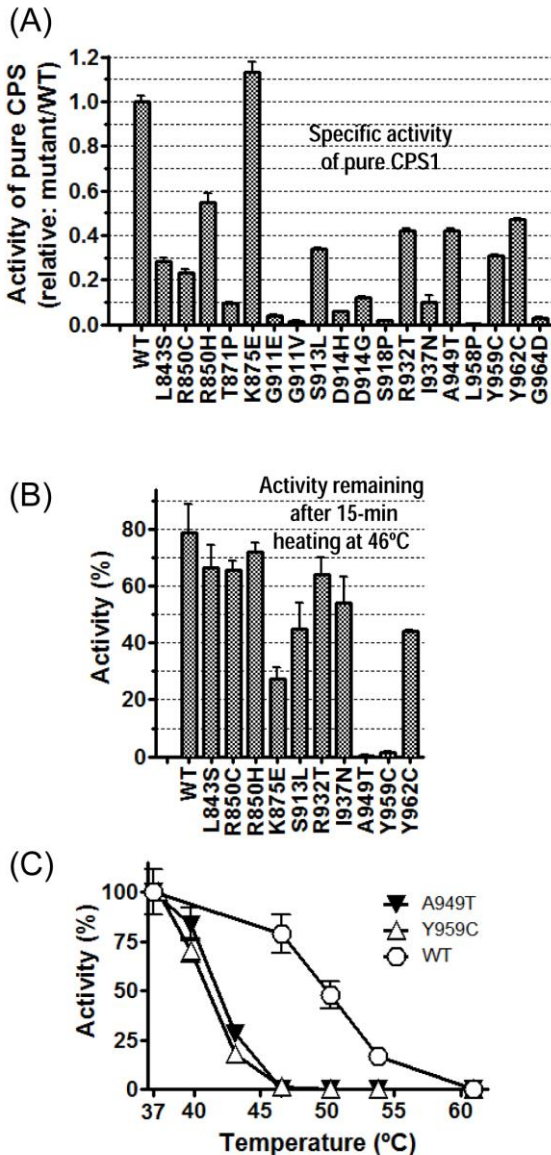


Figure 3. Enzyme activity (A) and thermal stability (B, C) of wild-type or mutant forms of CPS1. The activity per mg of pure CPS1 is expressed relative to the corresponding activity of the pure wild-type enzyme (2.3 U/mg). (B,C) Fraction (as percentage) of the activity remaining for the indicated enzyme form after 15-min heating either at 46°C (B) or at the indicated temperatures (C).

of unknown function (Fig. 1) which hosts few CPS1D mutations [2], and it is predicted by Polyphen-2 and MutPred as benign or as having little probability of being disease-causing (not shown).

RESULTS. Chapter 2

In the other case (Table 1) in which the same allele carried two missense mutations, p.Leu843Ser and p.Lys875Glu, the first of these mutations decreased activity by ~70% (Fig. 3A) and both mutations reduced ~90% the yield of CPS1, leading to the assumption of a liver activity of ~3% of normal for a null second allele (the case in the patient carrying these mutations [2]). This estimate might be optimistic considering the already mentioned possibility that the joint presence of both mutations in the same CPS1 polypeptide chain could reduce even further proper folding and CPS1 production, and also because the p.Lys875Glu mutation substantially decreased CPS1 thermal stability (Fig. 3B). In summary, the data for this patient appear to account fully for its neonatal presentation.

Decreased thermal stability appears a major determinant of the effects of the p.Ala949Thr and p.Tyr959Cys mutations (Figs. 3B,C), being the only important aberration observed for the first of these two mutations. Thus, p.Ala949Thr was not associated with decreased yield (Fig. 2A) or with a large specific activity decrease (Fig. 3A), but it lowered the half-inactivation temperature of the enzyme by ~8°C (Fig. 3C), bringing it down to nearly 40°C, which is close to the physiological body temperature, leading to the expectation of an important reduction in enzyme half-life and thus of enzyme level in the tissue. Nevertheless, some residual activity should be expected, agreeing with the late onset presentation in the patient carrying this mutation (the other allele carried a null mutation, Table 1). A similar degree of reduction in thermal stability was also observed for the p.Tyr959Cys mutation (Fig. 3B,C). However, in this case the mutation also decreased importantly CPS1 yield (by 75%; Fig. 2A) and specific activity (by 70%; Fig. 3A), with the resultant final activity in the tissue possibly being decreased further by the accelerated thermal inactivation. The late onset presentation in this patient (Table 1) could be due to the residual activity from this mutation, together with that resulting from the mutation carried in the second *CPS1* allele (p.Pro1462Arg), which is not inactivating (our own unpublished data).

Although the p.Tyr962Cys mutation only reduced ~50% the specific activity of the soluble enzyme (Fig. 3A) without substantially decreasing thermal stability (Fig. 3B), it decreased ~75% CPS1 yield (Fig. 2A). The combination of these two detrimental effects might result in an activity in the tissue of ~6% of normal, provided that the second *CPS1* allele found in the patient carrying this mutation (p.Ile632Arg; #17, Table 1), which affects the bicarbonate phosphorylation domain, is inactivating. This residual activity would explain that patient #17 is alive.

The patients with the p.Ser913Leu or p.Arg932Thr mutations had neonatal presentations (Table 1) that must reflect a cumulative effect of modest to drastic but not extreme changes in yield or specific activity of CPS1 caused by these mutations. Thus, the p.Ser913Leu mutation decreased CPS1 yield and specific activity, respectively, to ~20% and ~30% of normal, and the p.Arg932Thr mutation caused respective decreases to ~12% and 40% of normal (Figs. 2A and 3A). A residual liver activity of ~3% of normal is conceivable for both mutations, if they coexist in the patients with null second *CPS1* alleles.

2.2.4.5 K_m effects are the major changes associated with mutations affecting Arg850

Changes in yield (Fig. 2A), specific activity (Fig. 3A) or thermal stability (Fig. 3B) cannot account for the neonatal presentations observed in the two patients that carried mutations affecting Arg850, since for the p.Arg850His and p.Arg850Cys mutations thermal stability was essentially normal and the yields and specific activities were at least ~50% and ~25% of normal, respectively. However, since the specific activity assay used here utilizes saturating (for the wild-type enzyme) substrate concentrations, there might be room for K_m or K_a effects that could have negative consequences on enzyme activity at the generally low substrate concentrations present in the tissue [44,45]. This is certainly the case for the p.Arg850Cys mutation, which exhibits ~5-fold and ~3-fold increases in the K_m values for ATP and bicarbonate, respectively, and, even more importantly, a ~18-fold increase in the K_a value for NAG (Fig. 4A-C,E,G). Similarly, in the case of the p.Arg850His mutation the K_m for ATP and the K_a for NAG were increased 3-fold and 10-fold, respectively (Fig. 4A,B,F). Although these kinetic constant changes are not large enough to decrease importantly the activity observed in the standard assay, they can lead to drastic reductions in the CPS1 activity *in vivo*, since, for example, NAG may be far from saturating even for the wild-type enzyme [17-19,44,45], and thus, much less saturating for these mutant enzyme forms.

Among the ten UFSD mutants in which CPS1 yield and activity permitted kinetic analysis (Figs. 4A-D), substantial K_m or K_a changes were rarely observed. In addition to the already indicated changes in the two mutants of Arg850, the p.Ile937Asn mutation modestly increased the K_a value for NAG (Fig. 4A) and the K_m for ammonia (Figs. 4D,H), and the p.Leu843Ser and p.Ser913Leu mutations increased the K_a for NAG (Fig. 4A). It is interesting that the K_a for NAG was increased in all cases that exhibited a kinetic change, suggesting a role of the UFSD in the cross-talk known to exist between the NAG site and the catalytic machinery of the enzyme [15,23]. In summary,

RESULTS. Chapter 2

while K_m or K_a changes do not appear to play a paramount role in decreasing enzyme activity with most UFSD mutants, they certainly do so with the two mutations affecting Arg850, for which they appear major determinants of the severe deficiency observed in the patients carrying these mutations.

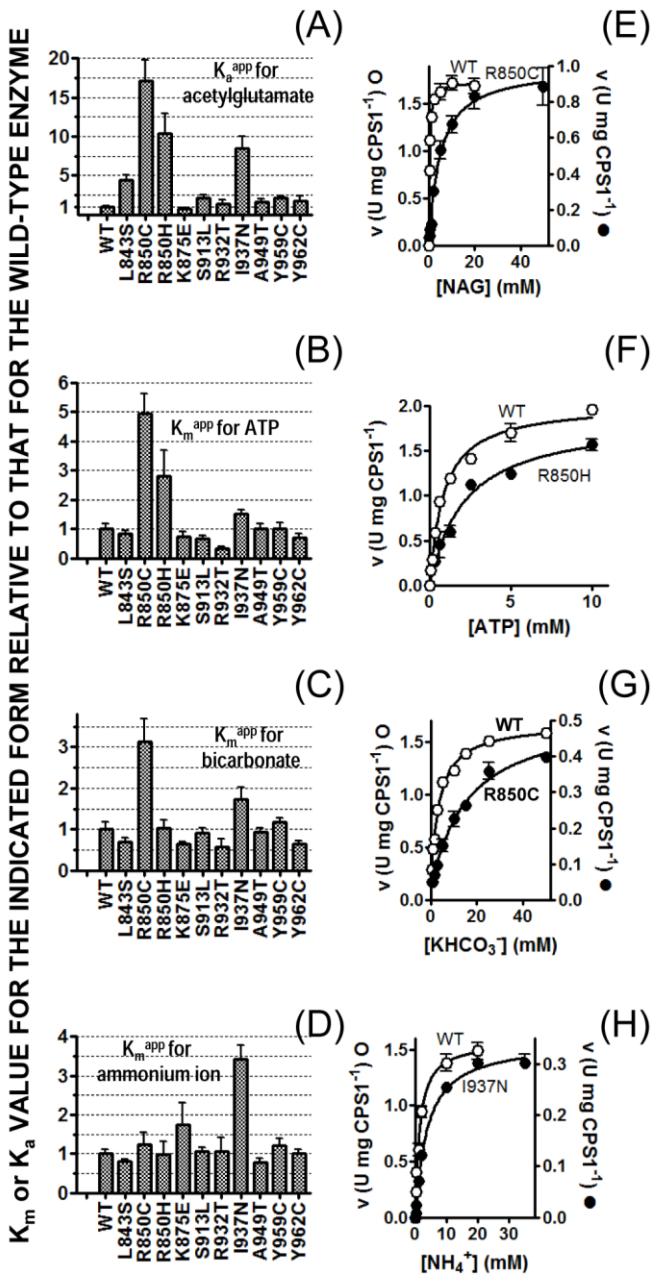


Figure 4. Effects of UFSD mutations on the K_m values of CPSI for its three substrates and on the K_a value for NAG. (A-D) Histograms illustrating the changes in K_m or K_a values for the indicated CPSI mutants and substrates or NAG, relative to the corresponding values for the wild-type enzyme. The curves on the right panels (E-H) illustrate the dependency of the reaction velocity on the concentration of each substrate or NAG for the wild-type enzyme or for the indicated mutants.

2.2.4.6 UFSD mutations generally decrease V_{\max} of the enzyme

With those mutations in which kinetic analysis was possible, it was observed that, except in the case of p.Lys875Glu, the mutations modestly decreased V_{\max} (Fig. 5A) in a proportion that was similar to the decrease in specific activity (Fig. 3A). This was to be expected if the mutations affect mainly k_{cat} without producing very large K_m effects, since the standard activity assay utilizes saturating (for the wild-type enzyme) concentrations of the substrates and of NAG. A plot of specific activity versus V_{\max} (Fig. 5B) could be reasonably adjusted to linear regression passing through zero. This suggests that the very low activity of the mutants that were very little expressed and that, therefore, were not amenable to kinetic analysis, could be due to a V_{\max} effect. Thus, the UFSD domain appears to influence the rate at which the enzyme catalyzes its complex, three-step reaction.

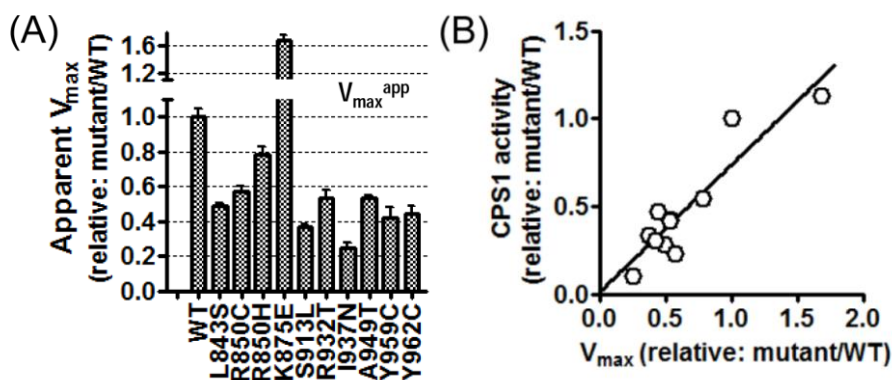


Figure 5. Influence of UFSD mutations on V_{\max} values for CPS1 activity. (A) Changes in V_{\max} values for the different mutants, as a fraction of the value for wild-type CPS1. (B) Plot of the specific activity for each mutant as a function of the V_{\max} for this same mutant. Relative values with respect to wild-type are given for both parameters. The regression line goes virtually through zero and gives a value of $r^2=0.842$.

2.2.5 Discussion

The present results validate the baculovirus/insect cell expression system [46] for monitoring the consequences of CPS1D mutations. Such system is already the standard for the same task in Gaucher's disease [47,48] and has been used with other diseases [49-52]. It associates the abundant expression and possibility of purification that characterizes *E. coli* expression, with a high compatibility with human proteins, which frequently cannot be expressed in bacteria [46]. In the case of Gaucher's disease, the large experience (exemplified in [47, 48]) attests the concordance of the results in this expression system with the clinical phenotype. This concordance was prospectively proven with two polymorphisms of methylenetetrahydrofolate reductase [50], since the one (p.Ala222Val) shown to have negative effects on the enzyme when using baculovirus/insect cells, was later on proven by epidemiological and genome-wide association studies to be linked to increased plasma homocysteine levels [53]. Our experience with CPS1 also supports the faithfulness of this system in revealing the impact of *CPS1* mutations. Thus, in our prior studies [26] two *CPS1* polymorphisms believed to be trivial had no substantial effect on any CPS1 trait investigated (activity, stability, kinetic constants, gross estimation of yield), whereas two mutations affecting a catalytic domain and used as positive controls were strongly detrimental. Furthermore, the results with three CPS1D-associated mutations affecting the glutaminase-like subdomain (GSD) (Fig. 1A) of unknown function and of three additional mutations affecting the allosteric domain (ASD, Fig. 1A) accounted for disease-causation because of impaired activity, stability or NAG activation. These results, which followed earlier pilot studies using baculovirus/insect cell-expressed rat CPS1 (a surrogate of human CPS1) [43], are extended now with those for 18 mutations affecting the UFSD.

The present results confirm that the seventeen mutations reported earlier [2,4,33-37] and the one described here that affect the UFSD of CPS1 in patients with CPS1D are disease-causing. The observed effects of the mutations on the production, activity, thermal stability and kinetic constants of the recombinant enzyme match in most cases the severity of the clinical presentation. Our data illustrate in one case the fact that each one of two missense changes in the same allele (p.Leu843Ser and p.Lys875Glu) has a detrimental effect on the enzyme, raising the possibility that their combined presence in the same protein molecule could be even more detrimental.

A clear conclusion of our studies is that, at least for this domain and with this eukaryotic expression system, the determination of the yield of CPS1 protein is a key element in judging the disease-causing role of each mutation. We clearly show that the abolition or the drastic decrease of CPS1 production is a crucial

mechanism of disease production by many missense mutations affecting the UFSD. Our findings point to decreased efficiency of proper folding as a key determinant of the poor production of the soluble protein. Although there are few *in vitro* expression studies for CPS1 [25,26,29,43], our prior investigations with 35 missense mutations mapping in other domains (with a predominance of mutations affecting the C-terminal ASD) do not highlight poor CPS1 production as a key determining element of the effects of the mutations affecting these other protein regions, since only with ~10% of the mutations was CPS1 production hampered or abolished. It therefore appears that a decrease in the efficiency of proper CPS1 folding is a characteristic and remarkable trait of CPS1D-associated mutations affecting the UFSD. Such trait may be the reason for the prominence of this domain in CPS1D, where missense mutations have been found in relatively high density [2].

The central position in the enzyme architecture and the intimate relations of the UFSD with other CPS1 domains [2,27,42] may account for the particular impact of the mutations affecting this domain on CPS1 folding. In the existing structural CPS1 models [2,42] based on the *E. coli* CPS structure [27], the UFSD makes very extensive contacts with both phosphorylation domains and, to a lesser extent, with the N-terminal moiety of the enzyme (the small subunit-like region) (Figs. 1C,D). This L-shaped domain (Fig. S1A) which embraces between its two arms the C subdomain of the bicarbonate phosphorylation domain (Fig. S1D), is sandwiched between the small subunit-like N-terminal moiety and the carbamate phosphorylation domain, two domains that lie respectively on top and below the plane defined by this L (Fig. 1D, right panel). Therefore, mutation-triggered UFSD misfolding can be expected to result in distorted relations between these other protein domains. The CPS1D-associated mutations affecting this domain cluster in the regions of interaction of the UFSD with other domains rather than in the exposed regions (Fig. 1D) that are involved in intermolecular interactions [27]. Therefore, the distribution of the mutations does not support the possibility of an increase in the interactions between the UFSDs of different enzyme molecules as the cause for the decreased solubility observed with many of these mutants. This agrees with our observation with *E. coli* CPS that the UFSD is not involved in dimer formation [54], and with our finding that human CPS1 exists mainly as monomers [10,26].

A detailed structural rationalization of the effects of the mutations studied here is included as Supplementary material. In summary, our results highlight a paramount integrating role of the UFSD for building the highly complex CPS1 architecture, revealing the key organizing function of this domain, and exemplifying the importance of this core structural element despite its lack of

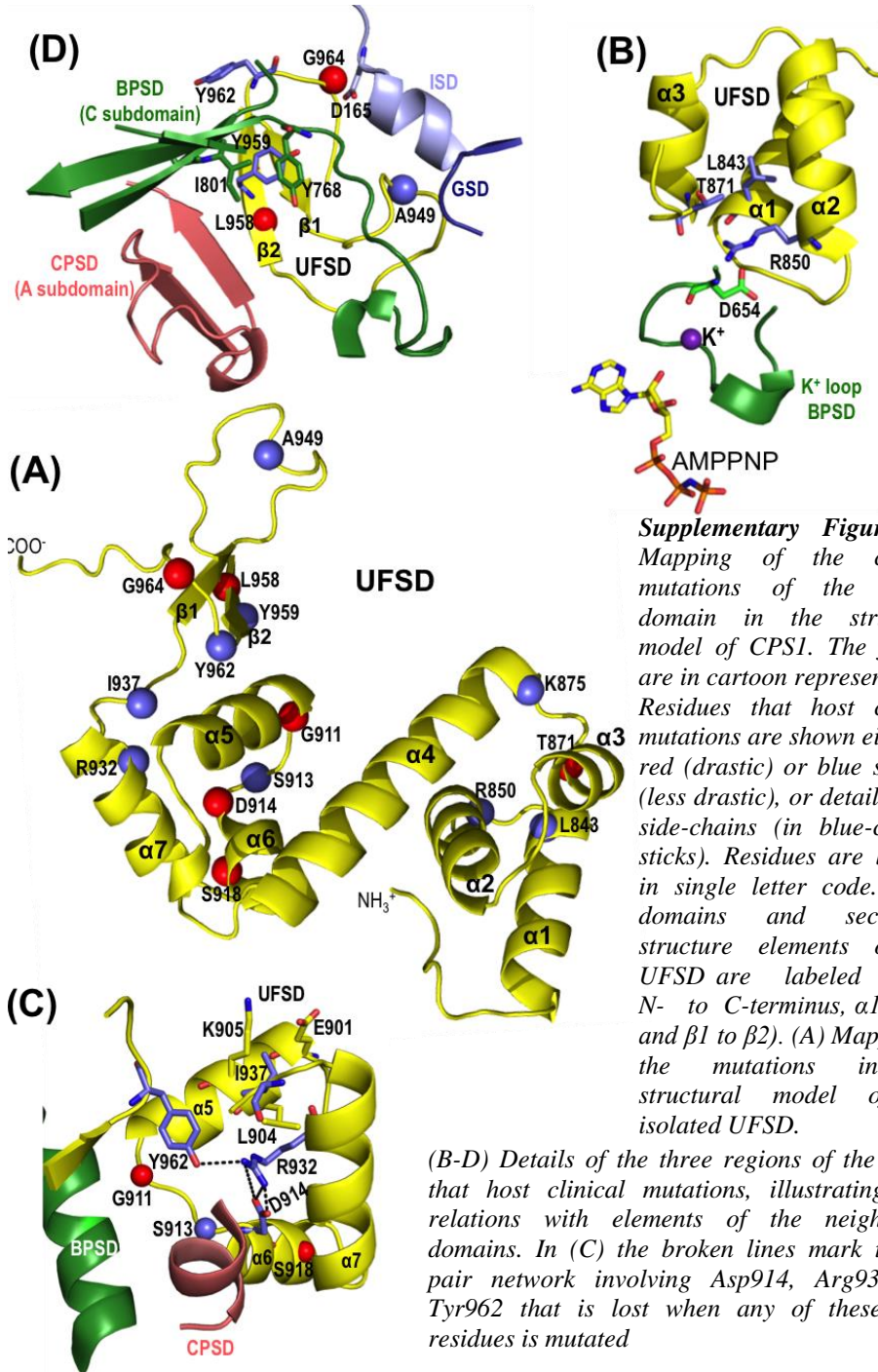
RESULTS. Chapter 2

substrate binding and catalytic machinery. Interestingly, a recent *in silico* study already proposed for this domain a key structural role in CPS1 [55]. On the basis of such key function of this domain, which is conserved in all forms of CPS (including the archaeal types, in which the C-terminal moiety is split into two complementary regions encoded by different genes), we would propose to call it in the future the "Integrating Domain" of CPS1. Our data fully account for the important representation of this domain in the database of missense mutations found in CPS1D [2], revealing that the main effects of the mutations in this domain are to negatively affect global CPS1 folding and architecture. We also document for some mutations and propose for other mutations that they have effects of greater or lower magnitude on the V_{\max} of the enzyme. This effect could be due to gross defects on domain architecture, that, for example, could block the conduit of the activating allosteric NAG signal, or the tunnel [27] through which intermediates must flow between both phosphorylation centers; or it could derive from more subtle changes such as the hampering of the concerted opening [56,57] of the B subdomains of both phosphorylation domains to allow product release. Indeed, mutations in the UFSD could be in the signaling path between both phosphorylation centers since the UFSD contacts the B domain of the carbamate phosphorylation domain, which is believed to trigger the concerted opening [57], and it also contacts the active site of the bicarbonate phosphorylation domain, which is to be opened (Fig. 1D). Similarly, the UFSD appears to be involved in the cross-talk between the NAG site and the catalytic centers, since some UFSD mutations increase the K_a for NAG. Full understanding of this involvement will have to await the determination of the structural mechanism of NAG activation of CPS1.

2.2.6 Supplementary material

Structural rationalization of the mutations effects.

Among the 18 mutations studied here, five affect four residues at the free end of one arm of the L-shaped UFSD, nine map at the other end of this arm, at the vertex of the L, and the remaining four mutations affect the distal half of the other arm (Fig. S1A). The core structure of the arm that hosts the larger number of mutations is folded in *E. coli* CPS [27] and is predicted to be folded in the structural models of CPS1 [2, 42] as two four-helix bundles connected by a long helix that participates in both bundles. Five mutations affecting four residues map in the more N-terminal four-helix bundle (Fig. S1A,B). Of these, three (p.Leu843Ser, p.Thr871Pro and p.Lys875Glu) likely compromise bundle folding. The other two mutations (p.Arg850Cys and p.Arg850His) affect a protruding residue that interacts with the active center K^+ loop of the adjacent



Supplementary Figure S1. Mapping of the clinical mutations of the UFSD domain in the structural model of CPS1. The figures are in cartoon representation. Residues that host clinical mutations are shown either as red (drastic) or blue spheres (less drastic), or detailing the side-chains (in blue-colored sticks). Residues are labeled in single letter code. CPS1 domains and secondary structure elements of the UFSD are labeled (from N- to C-terminus, $\alpha 1$ to $\alpha 7$; and $\beta 1$ to $\beta 2$). (A) Mapping of the mutations in the structural model of the isolated UFSD.

(B-D) Details of the three regions of the UFSD that host clinical mutations, illustrating their relations with elements of the neighboring domains. In (C) the broken lines mark the ion pair network involving Asp914, Arg932 and Tyr962 that is lost when any of these three residues is mutated

RESULTS. Chapter 2

bicarbonate phosphorylation domain, which sits next to this bundle (Fig. S1B), increasing the K_m value for ATP and, in the most drastic of the two mutations, for bicarbonate, without importantly hampering enzyme folding. This part of the UFSD may be involved in the as yet structurally unclarified cross-talk known to exist between this ATP site and the NAG site in the ASD [15, 23], because both mutations affecting Arg850 as well as another mutation in this four-helix bundle, p.Leu843Ser, substantially increased the K_a for NAG.

Six of the nine mutations at the vertex of the L (Fig. S1A,C) most likely affect the proper folding of the second four-helix bundle, with one (p.Ser918Pro) simply breaking a helix, and the other five hampering the gluing of the helices into the bundle either by disturbing an interhelical hydrophobic nest (the case for p.Ile937Asn), or by abolishing an ion pair network (the case for p.Asp914Gly/His, p.Arg932Thr and p.Tyr962Cys; Fig. S1C). The importance of the interaction of this bundle with the adjacent bicarbonate and carbamate phosphorylation domains is made patent by the important effects of the mutations p.Gly911Glu, p.Gly911Val, and p.Ser913Leu (Figs. 2A and 3A of the main text), which should disturb the interactions with two approximately parallel helices, each one from a neighboring domain. The fact that the p.Ser913Leu and p.Ile937Asn mutations increase the K_a for NAG (Fig. 4A of the main text) suggests that they are in the as yet unidentified route by which the cross-talk between the NAG site in the ASD and the catalytic machinery of the enzyme takes place.

The remaining four mutations affect the other arm of the L (Fig. S1A), a C-terminal β hairpin stem-loop terminated in a string that folds back towards the β strands. This hairpin is inserted as a gluing, extended element, between the interaction domain (ISD) and the glutaminase-like domains (GSD) of the small subunit-like N-terminal region (Fig. 1A of the main text) and the respective C and A subdomains of the bicarbonate phosphorylation and the carbamate phosphorylation domains (Fig. 1D of the main text, and Fig. S1D). This part of the UFSD is so integrated structurally with the other domains that its two β strands and its final string extend the central β sheet of the carbamate phosphorylation A subdomain, conforming with it nearly a β barrel that encircles one helix of this A subdomain (Fig. S1D). Two of these mutations should either break (p.Leu958Pro) or displace (p.Tyr959Cys) the β strand that continues the A subdomain β sheet, whereas another mutation falls in the C-terminal string (p.Gly964Asp) at a point where this string glues together the ISD and the already mentioned helix of the C subdomain of the bicarbonate phosphorylation domain. The extremely drastic effects of the p.Leu958Pro and p.Gly964Asp mutations (Figs. 2A and 3A, main text) attest to the structural importance of this region for proper CPS1 domain architecture. The

p.Ala949Thr mutation affects a residue at the tip of the hairpin loop which was reported in previous modeling studies of the UFSD [55] to interact with Arg172, belonging to the ISD. Actually, this residue is clamped between the ISD, the GSD, and elements of the C subdomain of the bicarbonate phosphorylation domain (Fig. S1D). The major negative effect of this mutation was on the thermal stability of the enzyme (Figs. 3B,C, main text). The other mutation having a similar impact on thermal stability (p.Tyr959Cys) affects a nearby residue that is also involved in interactions with the ISD and with the same structural elements of the bicarbonate phosphorylation domain, suggesting that this region and these domain-domain interactions have particular importance for ensuring good enzyme stability. In fact, the only other mutation (p.Lys875Glu) shown here to affect importantly enzyme stability also interacts with the ISD (not shown), supporting our previous conclusion [26] that one major function of the ISD is to increase CPS1 stability.

2.2.7 Acknowledgements

We thank Belén Barcelona (IBV-CSIC, Valencia) for her structural CPS1 model, and María Pilar Albero (Centro de Investigación Príncipe Felipe, Valencia) for initial data on p.Arg850Cys and p.Arg850His mutants, and Matthias Baumgartner (University Children's Hospital, Zurich) for critical reading of the manuscript. The mutation analysis of CPS1D for the two new patients reported here was kindly supported by Orphan Europe. This does not involve any conflict of interest. This work was supported by grants from the Fundación Alicia Koplowitz, the Valencian and Spanish governments (Prometeo 2009/051 and BFU2011-30407, respectively) and the Swiss National Science Foundation (grant 310030_127184). C.D-F was a FPU fellow of the Spanish Government and received from that Government a bursary for short-time work in Zurich

Conflict of interest statement

The authors state that there is no conflict of interest.

2.2.8 References

- [1] S. Brusilow, A.L. Horwich, Urea Cycle Enzymes, in: C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle (Eds.), *The Metabolic & Molecular Bases of Inherited Disease*, 8th Edition, McGraw-Hill, New York, 2001, pp. 1909-1963.

RESULTS. Chapter 2

- [2] J. Häberle, O.A. Shchelochkov, J. Wand, P. Katsonis, L. Hall, S. Reiss, A. Eeds, A. Willis, M. Yadav, S. Summar, O. Lichtarge, V. Rubio, L.J. Wong, M. Summar, Molecular defects in human carbamoyl phosphate synthetase I: mutational spectrum, diagnostic and protein structure considerations, *Hum Mutat.* 32 (2011) 579-589.
- [3] M.L. Summar, M.J. Dasouki, P.J. Schofield, M.R. Krishnamani, C. Vnencak-Jones, M. Tuchman, J. Mao, J.A Phillips 3rd, Physical and linkage mapping of human carbamyl phosphate synthetase I (CPS1) and reassignment from 2p to 2q35, *Cytogenet Cell Genet.* 71 (1995) 266-267.
- [4] J. Häberle, E. Schmidt, S. Pauli, B. Rapp, E. Christensen, B. Wermuth, H.G. Koch, Gene structure of human carbamylphosphate synthetase 1 and novel mutations in patients with neonatal onset, *Hum Mutat.* 21 (2003) 444.
- [5] M.L. Summar, L.D. Hall, A.M. Eeds, H.B. Hutcheson, A.N. Kuo, A.S. Willis, V. Rubio, M.K. Arvin, J.P Schofield, E.P. Dawson, Characterization of genomic structure and polymorphisms in the human carbamyl phosphate synthetase I gene, *Gene.* 311 (2003) 51-57.
- [6] S. Funghini, M.A. Donati, E. Pasquini, E. Zammarchi, A. Morrone, Structural organization of the human carbamyl phosphate synthetase I gene (CPS1) and identification of two novel genetic lesions, *Hum Mutat.* 22 (2003) 340-341.
- [7] Y. Haraguchi, T. Uchino, M. Takiguchi, F. Endo, M. Mori, I. Matsuda, Cloning and sequence of a cDNA encoding human carbamyl phosphate synthetase I: molecular analysis of hyperammonemia, *Gene.* 107 (1991) 335-340.
- [8] E.H. Van Beers, E.H. Rings, G. Posthuma, M.A. Dingemans, J.A. Taminiou, H.S. Heymans, A.W. Einerhand, H.A Büller, J. Dekker, Intestinal carbamoyl phosphate synthase I in human and rat. Expression during development shows species differences and mosaic expression in duodenum of both species, *J Histochem Cytochem.* 46 (1998) 231-240.
- [9] M.A. Neill, J. Aschner, F. Barr, M.L. Summar, Quantitative RT-PCR comparison of the urea and nitric oxide cycle gene transcripts in adult human tissues, *Mol Genet Metab.* 97 (2009) 121-127.
- [10] V. Rubio, G. Ramponi, S. Grisolia, Carbamoyl phosphate synthetase I of human liver. Purification, some properties and immunological cross-reactivity with the rat liver enzyme, *Biochim Biophys Acta.* 659 (1981) 150-160.
- [11] M. Mori, S. Miura, T. Morita, M. Takiguchi, M. Tatibana, Synthesis, intracellular transport and processing of mitochondrial urea cycle enzymes, *Adv Enzyme Regul.* 21 (1982) 121-132.
- [12] H. Nyunoya, K.E. Broglie, E.E. Widgren, C.J Lusty, Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat, *J Biol Chem.* 260 (1985) 9346-9356.

- [13] V. Rubio, Structure-function studies in carbamoyl phosphate synthetases, *Biochem Soc Trans.* 21 (1993) 198-202.
- [14] L.M. Hall, R.L. Metzzenberg, P.P. Cohen, Isolation and characterization of a naturally occurring cofactor of carbamyl phosphate biosynthesis, *J Biol Chem.* 230 (1958) 1013-1021.
- [15] V. Rubio, H.G. Britton, S. Grisolia, Mitochondrial carbamoyl phosphate synthetase activity in the absence of N-acetyl-L-glutamate. Mechanism of activation by this cofactor, *Eur J Biochem.* 134 (1983) 337-343.
- [16] C. Bachmann, S. Krähenbühl, J.P. Colombo, Purification and properties of acetyl-CoA: L-glutamate N-acetyltransferase from human liver, *Biochem J.* 205 (1982) 123-127.
- [17] M. Tatibana, S. Kawamoto, T. Sonoda, M. Mori, Enzyme regulation of N-acetylglutamate synthesis in mouse and rat liver, *Adv Exp Med Biol.* 153 (1982) 207-216.
- [18] K. Shigesada, K. Aoyagi, M. Tatibana, 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 (1978) 385-391.
- [19] P.M. Stewart, M. Walser, Short term regulation of ureagenesis, *J Biol Chem.* 255 (1980) 5270-5280.
- [20] P.M. Anderson, A glutamine- and N-acetyl-L-glutamate-dependent carbamyl phosphate synthetase activity in the teleost *Micropterus salmoides*, *Comp Biochem Physiol B.* 54 (1976) 261-263.
- [21] A. Meister, Mechanism and regulation of the glutamine-dependent carbamyl phosphate synthetase of *Escherichia coli*, *Adv Enzymol Relat Areas Mol Biol.* 62 (1989) 315-374.
- [22] X. Liu, H.I. Guy, D.R. Evans, Identification of the regulatory domain of the mammalian multifunctional protein CAD by the construction of an *Escherichia coli* hamster hybrid carbamyl-phosphate synthetase, *J Biol Chem.* 269 (1994) 27747-27755.
- [23] E. Alonso, V. Rubio, Binding of N-acetyl-L-glutamate to rat liver carbamoyl phosphate synthetase (ammonia), *Eur J Biochem.* 135 (1983) 331-337.
- [24] Y. Xie, K. Ihsanawati, S. Kishishita, K. Murayama, C. Takemoto, M. Shirozu, RIKEN Structural Genomics/Proteomics Initiative, Crystal structure of MGS domain of carbamoyl-phosphate synthetase from *Homo sapiens*. (February-2009). Protein DataBank file 2YVQ, <http://www.rcsb.org/pdb>.
- [25] S. Pekkala, A.I. Martinez, B. Barcelona, J. Gallego, E. Bendala, I. Yefimenko, V. Rubio, J. Cervera, Structural insight on the control of urea synthesis: identification of the binding site for N-acetyl-L-glutamate, the

RESULTS. Chapter 2

- essential allosteric activator of mitochondrial carbamoyl phosphate synthetase, *Biochem J.* 424 (2009) 211-220.
- [26] C. Diez-Fernandez, A.I. Martinez, S. Pekkala, B. Barcelona, I. Perez-Arellano, A. Guadalajara, M. Summar, J. Cervera, V. Rubio, Molecular characterization of carbamoyl-phosphate synthetase (CPS1) deficiency using human recombinant CPS1 as a key tool, *Hum Mutat.* 8 (2013) 1149-1159.
- [27] J.B. Thoden, H.M. Holden, G. Wesenberg, F.M. Raushel, I. Rayment, Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product, *Biochemistry.* 36 (1997) 6305-6316.
- [28] S.G. Powers-Lee, K. Corina, Domain structure of rat liver carbamoyl phosphate synthetase I, *J Biol Chem.* 261(1986) 15349-15352.
- [29] V. Ahuja, S.G. Powers-Lee, Human carbamoyl-phosphate synthetase: insight into N-acetylglutamate interaction and the functional effects of a common single nucleotide polymorphism, *J Inherit Metab Dis.* 31 (2008) 481-491.
- [30] E. Alonso, J. Cervera, A. García-España, E. Bendala, V. Rubio, 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 (1992) 4524-4532.
- [31] E. Alonso, V. Rubio, 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 (1995) 377-384.
- [32] L.B. Rodriguez-Aparicio, A.M. Guadalajara, V. Rubio, 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 (1989) 3070-3074.
- [33] Y. Wakutani, H. Nakayasu, T. Takeshima, M. Adachi, M. Kawataki, K. Kihira, H. Sawada, M. Bonno, H. Yamamoto, K. Nakashima, Mutational analysis of carbamoylphosphate synthetase I deficiency in three Japanese patients, *J. Inherit. Metab. Dis.* 27 (2004) 787-788.
- [34] A.M. Eeds, L.D. Hall, M. Yadav, A. Willis, S. Summar, A. Putnam, F. Barr, M.L. Summar, The frequent observation of evidence for nonsense-mediated decay in RNA from patients with carbamyl phosphate synthetase I deficiency, *Mol. Genet. Metab.* 89 (2006) 80-86.
- [35] K. Kurokawa, T. Yorifuji, M. Kawai, T. Momoi, H. Nagasaka, M. Takayanagi, K. Kobayashi, M. Yoshino, T. Kosho, M. Adachi, H. Otsuka, S. Yamamoto, T. Murata, A. Suenaga, T. Ishii, K. Terada, N. Shimura, K. Kiwaki, H. Shintaku, M. Yamakawa, H. Nakabayashi, Y. Wakutani, T. Nakahata, Molecular and clinical analyses of Japanese patients with carbamoylphosphate synthetase 1 (CPS1) deficiency, *J. Hum. Genet.* 52 (2007) 349-354.

- [36] S. Funghini, J. Thusberg, M. Spada, S. Gasperini, R. Parini, L. Ventura, C. Meli, L. De Cosmo, M. Sibilio, S.D. Mooney, R. Guerrini, M.A. Donati, A. Morrone, Carbamoyl phosphate synthetase 1 deficiency in Italy: clinical and genetic findings in a heterogeneous cohort, *Gene*. 493 (2012) 228-234.
- [37] R. Kretz, L. Hu, V. Wettstein, D. Leiteritz, J. Häberle, Phytohemagglutinin stimulation of lymphocytes improves mutation analysis of carbamoylphosphate synthetase 1, *Mol Genet Metab*. 106 (2012) 375-388.
- [38] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, *Nat Methods*. 7 (2010) 248-249.
- [39] B. Li, V.G. Krishnan, M.E. Mort, F. Xin, K.K. Kamati, D.N. Cooper, S.D. Mooney, P. Radivojac, Automated interference of molecular mechanisms of disease from amino acid substitutions, *Bioinformatics*. 25 (2009) 2744-2750.
- [40] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal Biochem*. 72 (1976) 248-54.
- [41] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*. 227 (1970) 680-685.
- [42] A.I. Martínez, I. Pérez-Arellano, S. Pekkala, B. Barcelona, J. Cervera, Genetic, structural and biochemical basis of carbamoyl phosphate synthetase 1 deficiency, *Mol Genet Metab*. 101 (2010) 311-323.
- [43] S. Pekkala, A.I. Martinez, B. Barcelona, I. Yefimenko, U. Finckh, V. Rubio, J. Cervera, Understanding carbamoyl-phosphate synthetase I (CPS1) deficiency by using expression studies and structure-based analysis, *Hum Mutat*. 31 (2010) 801-808.
- [44] E. Alonso, J. Girbés, A. García-España, V. Rubio, Changes in urea cycle-related metabolites in the mouse after combined administration of valproic acid and an amino acid load, *Arch Biochem Biophys*. 272 (1989) 267-273.
- [45] A. García-España, E. Alonso, V. Rubio, Influence of anions on the activation of carbamoyl phosphate synthetase (ammonia) by acetylglutamate: implications for the activation of the enzyme in the mitochondria, *Arch Biochem Biophys*. 288 (1991) 414-420.
- [46] T.A. Kost, J.P. Condreay, D.L. Jarvis, Baculovirus as versatile vectors for protein expression in insect and mammalian cells, *Nat Biotechnol*. 23 (2005) 567-575.
- [47] M. Montfort, A. Chabás, L. Vilageliu, D. Grinberg, Functional analysis of 13 GBA mutant alleles identified in Gaucher disease patients: Pathogenic changes and "modifier" polymorphisms, *Hum Mutat*. 23 (2004) 567-575.
- [48] B. Liou, A. Kazimierczuk, M. Zhang, C.R. Scott, R.S. Hegde, G.A. Grabowski, Analyses of variant acid beta-glucosidases: effects of Gaucher disease mutations, *J Biol Chem*. 281 (2006) 4242-4253.

RESULTS. Chapter 2

- [49] A.K. Das, J.Y. Lu, S.L. Hofmann, Biochemical analysis of mutations in palmitoyl-protein thioesterase causing infantile and late-onset forms of neuronal ceroid lipofuscinosis, *Hum Mol Genet.* 10 (2001) 1431-1439.
- [50] K. Yamada, Z. Chen, R. Rozen, R.G. Matthews, Effects of common polymorphisms on the properties of recombinant human ethylenetetrahydrofolate reductase, *Proc Natl Acad Sci USA.* 98 (2001) 14853-14858.
- [51] T. Grau, N.O. Artemyev, T. Rosenberg, H. Dollfus, O.H. Haugen, E. Cumhur Sener, B. Jurklies, S. Andreasson, C. Kernstock, M. Larsen, E. Zrenner, B. Wissinger, S. Kohl, Decreased catalytic activity and altered activation properties of PDE6C mutants associated with autosomal recessive achromatopsia, *Hum Mol Genet.* 20 (2011) 719-730.
- [52] E. Ostergaard, R.J. Rodenburg, M. van den Brand, L.L. Thomsen, M. Duno, M. Batbayli, F. Wibrand, L. Nijtmans, Respiratory chain complex I deficiency due to NDUFA12 mutations as a new cause of Leigh syndrome, *J Med Genet.* 48 (2011) 737-740.
- [53] K. Miyaki, Genetic polymorphisms in homocysteine metabolism and response to folate intake: a comprehensive strategy to elucidate useful genetic information. *J Epidemiol.* 20 (2010) 266-270.
- [54] P. Mora, V. Rubio, J. Cervera, Mechanism of oligomerization of *Escherichia coli* carbamoyl phosphate synthetase and modulation by the allosteric effectors. A site-directed mutagenesis study, *FEBS Lett.* 511 (2002) 6-10.
- [55] M. Lopes-Marques, G. Igrejas, A. Amorim, L. Azevedo. Human carbamoyl phosphate synthetase I (CPSI): insights on the structural role of the unknown function domains. *Biochem Biophys Res Commun.* 421 (2012) 409-412.
- [56] J.B. Thoden, G. Wesenberg, F.M. Raushel, H.M. Holden, Carbamoyl phosphate synthetase: closure of the B-domain as a result of nucleotide binding, *Biochemistry.* 38 (1999) 2347-2357.
- [57] I. Yefimenko, V. Fresquet, C. Marco-Marín, V. Rubio, J. Cervera, Understanding carbamoyl phosphate synthetase deficiency: impact of clinical mutations on enzyme functionality, *J Mol Biol.* 349 (2005) 127-141

RESULTS

2.3 CHAPTER 3

EXPERIMENTAL STUDIES ON THE INHERITED METABOLIC ERROR CARBAMOYL PHOSPHATE SYNTHETASE 1 DEFICIENCY SHED LIGHT ON THE MECHANISM FOR SWITCHING ON/OFF THE UREA CYCLE

by

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

^aInstituto de Biomedicina de Valencia of the Consejo Superior de Investigaciones Científicas (IBV-CSIC), Valencia, Spain; ^bFacultad de Medicina, Universidad Católica de Valencia, Spain; ^cUniversity Children's Hospital Zurich and Children's Research Center, Zurich, Switzerland; ^dGroup 739 of the Centro de Investigación Biomédica en Red sobre Enfermedades Raras (CIBERER) del Instituto de Salud Carlos III, Spain

Submitted to: Journal of Genetics and Genomics

*Correspondence to:

V. Rubio
Instituto de Biomedicina de Valencia (IBV-CSIC)
Jaume Roig 11
46010 Valencia, Spain
Phone: +34 96 3391772.
Fax: +34 96 3690800
rubio@ibv.csic.es

2.3.1 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. This last observation raises the possibility of using in these patients NAG site saturation therapy with the NAG analog and registered drug N-carbamyl-L-glutamate. Furthermore, these data, 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. Restrained molecular dynamics simulations based on the observed mutations effects are consistent with this proposal, providing further backing for this structurally plausible signaling mechanism by which NAG could trigger urea cycle activation via CPS1.

Key words: Urea cycle diseases, inborn errors, hyperammonemia, site-directed mutagenesis; restrained molecular dynamics; allosteric regulation; carbamoyl phosphate synthetase 1; enzyme.

2.3.2 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, CPS1 deficiency 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 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 (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>) 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 CPS reaction (Alonso et al., 1992; Alonso and Rubio, 1995).

Our present work deals with the analysis of the effects of CPS1D-associated mutations (called also from here on 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

RESULTS. Chapter 3

glutamate by an enzyme (NAG synthase) exhibiting a high K_m for glutamate (Sonoda and Tatibana, 1983).

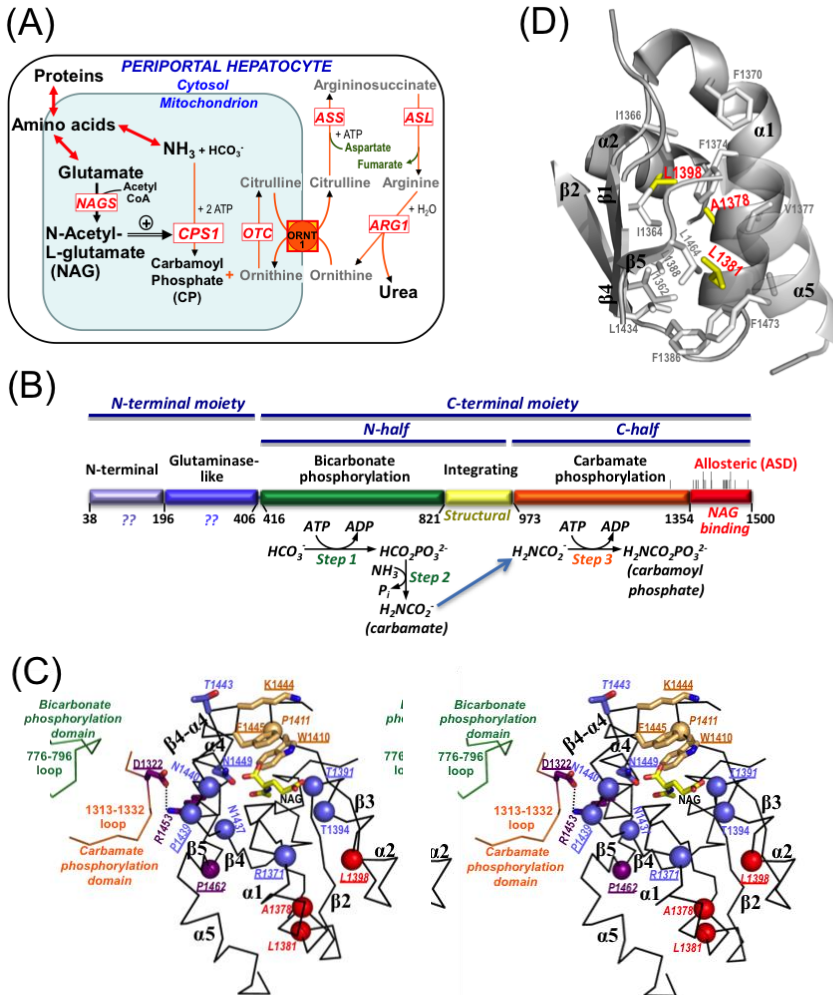


Figure 1. The CPS1-NAG switch for urea cycle control. (A) Simplified view of the urea cycle to highlight its control at the level of CPS1. The double arrows denote bidirectional fluxes between the elements that are linked. Enzymes are boxed and abbreviated as follows: NAGS, NAG synthase; OTC, ornithine transcarbamylase; ASS, argininosuccinate synthetase; ASL, argininosuccinate lyase; and 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, which is where the switch mechanism operates, is highlighted. (Cont'd next page)

To explore the effects of ASD-mapping missense clinical mutations (abbreviated ASD clinical mutations) we utilize 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 Table 1). 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 we define better the NAG site. We had previously localized this 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

Fig. 1, cont'd: (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 Table 1). The longer lines indicate that two different mutations affect the same residue. **(C)** Stereo view 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, being identified by labeling (*italics*, residues hosting CPS1D-associated mutations; residues underlined are those hosting mutations studied experimentally here). 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 $^\alpha$ 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 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 prepared from PDB file 2YVQ.

RESULTS. Chapter 3

(Protein Databank file 2YVQ; Xie et al., 2007). We now have found that some ASD mutations have effects not expected in 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 β 4- α 4 loop could be the initial signal in NAG activation.

2.3.3 Results

2.3.3.1 Clinical ASD domain mutations studied

The twelve reported (Häberle et al., 2011) CPS1D-associated missense mutations mapping in the ASD are listed in Table 1A. 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 (Adzhubei et al., 2010; Li et al., 2009) (Table 1A).

Table 1A highlights in bold-face the clinical mutations that were studied experimentally here. The study begun 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 (Figs. 2C and 2D) are referred 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 Figs. 2D and 2E.

2.3.3.2 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 1A) decreased importantly CPS1 production or purification (Fig. 2A). Only with the L1398V mutation there was some reduction (~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

~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 1A). 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 β_1].

2.3.3.3 Effects of clinical ASD mutations on CPS1 activity and on 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 (~97%) inactivation. Previously, the clinical mutations R1453W and R1453Q were found to inactivate the otherwise apparently well folded mutant CPS1 (Table 1A), whereas T1443A and Y1491H decreased by ~20-fold and ~4-fold, respectively, the specific activity of the enzyme (Pekkala et al., 2010; Díez-Fernández et al., 2013). Thus, eight among the twelve reported ASD clinical mutations (Häberle et al., 2011) importantly decrease enzyme specific activity. Of the patients carrying these activity-decreasing mutations, those for which the clinical condition was known presented severe deficiency (Table 1A).

The study of the NAG activation kinetics of the purified mutant proteins (Figs. 2D and 2E) 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 ~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 Holzknecht, 1990). The same should be the case for the T1443A and Y1491H clinical mutations, which had been found earlier (Pekkala et al., 2010; Díez-Fernández et al. 2013) to increase K_a^{NAG} ~160-fold and ~50-fold, respectively (Table 1A). Overall, in five out of the twelve ASD clinical mutations the K_a^{NAG} was importantly increased.

The large decrease in the observed enzyme activity caused by the P1462R mutation was largely due to decreased velocity at saturation of NAG ($V^{\text{[NAG]}=\infty}$;

RESULTS. Chapter 3

Table 1. CPS1 allosteric domain mutations either found in CPS1 deficiency patients or designed to test functional features, and their effects.

<i>A. Mutations found in patients with CPS1 deficiency ("clinical mutations")</i>								
Amino acid change ^a	Amino acid in CPS			Severity			Proposed residue role (new NAG model ^b)	Major effects on recombinantly expressed CPS1
	I	III	Other	Observed in patients	Server-predicted			
					PolyPhen-2	MutPred		
							g-score	
R1371L ^c	R	R/q	Variable	Unknown	Benign	0.51	Ion pair NAG γ -COO ⁻ (restraint)	$\uparrow K_a^{NAG}$ ~100-fold
A1378T ^d	A	A	Apolar	Severe	Prob. damag.	0.92	α_3/β_5 hydrophobic nest	\downarrow thermal stability ^e
L1381S ^f	L	L/f	L/F/v	Severe	Prob. damag.	0.92	α_3/β_5 hydrophobic nest	Degradation ^e
T1391M ^c	T	T	Variable	Severe ^g	Prob. damag.	0.80	H-bond with NAG NH (restraint)	$\uparrow K_a^{NAG}$ ~200-fold
L1398V ^c	L	L	L/Y/f/w/i	Severe ^g	Prob. damag.	0.72	α_3/β_5 hydrophobic nest	\downarrow thermal stability
P1411L ^{d,f}	P	P	Variable	Mild	Prob. damag.	0.84	β_3 - α_3 loop, near NAG site	Modest V_{max} effect ^h
P1439L ^c	P	P	Variable	Severe ^g	Prob. damag.	0.66	β_4 - α_4 loop conformation	$\uparrow K_a^{NAG}$ ~15-fold
T1443A ^d	T	T/S	Variable	Severe	Poss. damag.	0.74	H-bond with NAG α -COO ⁻ (restraint) ⁱ	$\uparrow K_a^{NAG}$ ~160-fold ^e
R1453W ^{c,h}	R	R	R/l/i	Severe	Prob. damag.	0.82	NAG signal transmission to the catalytic domains	Inactivation ^h
R1453Q ^h	R	R	R/l/i	Severe	Prob. damag.	0.89	NAG signal transmission to the catalytic domains	Inactivation ^h
P1462R ^c	P/A	P	Variable	Unknown	Poss. damag.	0.85	Conformation of signal transmission region	$\downarrow V_i^{[NAG]_{\infty}}$ ~30-fold $\uparrow K_M^{ATP}$ ~10-fold
Y1491H ^f	Y/F	Y/h	Variable	Unknown	Benign	0.9	Interdomain signal transmission	$\uparrow K_a^{NAG}$ ~50-fold ^h

<i>B. Mutations introduced on rational bases</i>						
Amino acid change ^a	Amino acid in CPS			Proposed role of the residue (new NAG model ^b)	Major effect on recombinant expressed CPS1	
	I	III	Other			
D1322L ^j	D	D	D/y	NAG signal transmission	Inactivation	
T1391V	T	T	Variable	H-bond with NAG NH (restraint)	$\uparrow K_a^{NAG}$ ~300-fold ^k	
T1394A	T	T	T/d/y	H-bond with NAG γ -COO ⁻ (restraint)	$\uparrow K_a^{NAG}$ ~600-fold ^k	
W1410A	W	W	Variable	β_3 - α_3 loop, near NAG site ^l	No substantial effect	
W1410K	W	W	Variable	β_3 - α_3 loop, near NAG site ^l	$\uparrow K_a^{NAG}$ ~40-fold ^k	
N1437D	N	N	Variable	H-bond with NAG O _{Ac} (restraint)	$\uparrow K_a^{NAG}$ ~50-fold ^k	
N1440D	N	N	Variable	H-bond with NAG α -COO ⁻ (restraint)	$\uparrow K_a^{NAG}$ ~90-fold ^k	
K1444A	K/r	K/R	Variable	Part of β_4 - α_4 loop ^m	$\uparrow K_a^{NAG}$ ~3-fold	
F1445A	F/y/h	F/q/h/y	Variable	Part of β_4 - α_4 loop ^m	No substantial effect	
N1449A	N	N	Variable	H-bond with NAG α -COO ⁻	$\uparrow K_a^{NAG}$ ~10-fold	

Mutations studied experimentally here are highlighted in bold-type. CPSIII is the piscine form of CPS, which is partially dependent on NAG (Hong et al., 1994). The severity of the clinical mutations is based on the clinical presentation, being graded as severe for neonatal presentations and/or death. P1411L was considered mild because it coexisted with a null second CPS1 allele (Q478*) and yet the deficiency had a late clinical onset indicating substantial residual CPS1 activity. No clinical information was available for the patient with the R1371L mutation, and, therefore, its severity is unknown. The patient carrying the P1462R mutation is alive, but it is uncertain whether this is due to a mild effect of this mutation or to residual activity associated with its second CPS1 allele, which was the missense change Y959C. The patient carrying the

Y1491H mutation had a late onset presentation, but the second allele was not identified, and therefore the possibility cannot be excluded that the residual activity was due to a non-diagnosed mild second allele. When reporting their occurrence in the various CPSs, amino acids are in low case when found with low frequency. Variable denotes the occurrence at a given position, in the indicated groups of CPSs, of >4 types of amino acids with no constant chemical characteristics (polar, apolar, charged, etc.). PolyPhen-2 grades the probability of a damaging effect of an amino-acid substitution, from higher to lower, as 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. For further details, see Materials and Methods.

^aTranslation 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.

^bWhen the contact was used as a restraint for model generation, it is specified.

^cHäberle et al., 2011.

^dEeds et al., 2006.

^eDíez-Fernández et al., 2013.

^fSummar 1998.

^gUnknown when originally reported. New data gathered on the patient.

^hPekkala et al., 2010.

ⁱNot present in previous docking model (Pekkala et al., 2009).

^jThis residue belongs to the carbamate phosphorylation domain. It is shown here because of the belief that it plays a key role in the transmission of the allosteric signal (see text).

^kPekkala et al., 2009.

^lNAG site lid residue and H-bond to NAG α -COO⁻ in previous docking model (Pekkala et al., 2009).

^mNAG site lid residue in previous docking model (Pekkala et al., 2009).

Fig. 2D), whereas the K_a^{NAG} remained essentially normal (Fig. 2E). Substrate kinetics for this mutant were normal for all substrates (not shown) except for ATP. Thus, the apparent K_m^{ATP} was increased ~10-fold (Fig. 2F), which should not substantially decrease enzyme activity in our ATP-rich standard assay, but which may contribute to the deficiency *in vivo*. Given the lack of substrate-binding or catalytic machinery in the ASD, the decreased $V^{\text{[NAG]}=\infty}$ 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 to be hampered. This may also be the reason for the inactivation caused by the ASD clinical mutations R1453W and R1453Q (Pekkala et al., 2009, 2010).

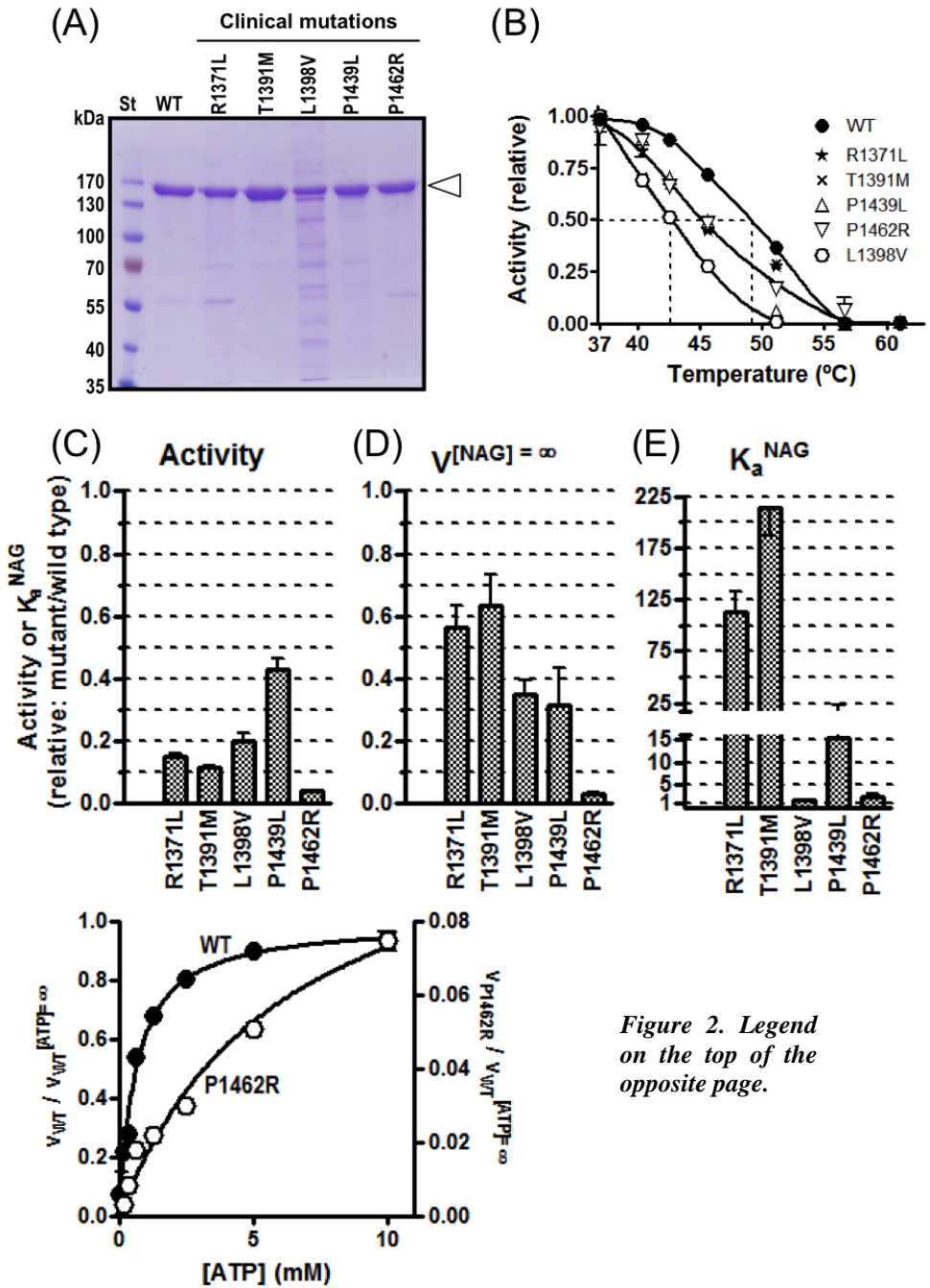


Figure 2. Legend on the top of the opposite page.

Figure 2. Production, stability, activity and kinetic properties of wild-type and CPSID-associated mutant CPSI forms. (A) SDS-PAGE (8% polyacrylamide, Coomassie staining) of purified human recombinant CPSI, either wild-type (WT) or carrying the indicated mutations. St, protein markers (PageRuler Prestained Protein Ladder, from Thermo Scientific), with masses indicated in kDa. The arrowhead signals the position of the CPSI band. (B) Inactivation upon 15-min heating at the indicated temperatures of recombinant human CPSI, either wild-type or carrying the indicated mutations. For clarity, a single line was fitted to the highly similar in terms of stability R1371L, T1391M, P1439L and P1462R mutants. The horizontal dashed line marks 50% inactivation, whereas the vertical dashed lines cross the X-axis at the temperature at which 50% inactivation occurs for the corresponding enzyme form. (C-E) Effects of CPSID-associated ASD mutations on enzyme activity (C), apparent V_{max} (D), K_a^{NAG} (E), and ATP substrate kinetics (F). For details on the assays see Materials and Methods. The activities and V_{max} values were always referred to 1 mg of protein. Results are given as fractions of the corresponding mean values for wild-type CPSI (Fernández-Díez et al., 2013). In (F), the curves are hyperbolae fitted with GraphPad Prism (GraphPad Software, San Diego, California). The value of I was the velocity extrapolated at infinite concentration of ATP for the wild-type enzyme.

2.3.3.4 The effects of ASD mutations shed further light 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, designed on rational bases) D1322L mutation (Table 1B; this part of the table lists rationally designed 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 CPSI 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 CPSI by abolishing transmission of the NAG signal to the catalytic domains, the D1322L mutation should also be inactivating. This was indeed the effect observed (Table 1B). Although the 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 CPSI (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

RESULTS. Chapter 3

traits of the site (Table 1B, in bold-type; and Fig. 3). The large effects on K_a^{NAG} of the N1437D, N1440D, and T1394A mutations as well as of the presently studied R1371L and T1391M clinical mutations (Fig. 2E) were consistent with

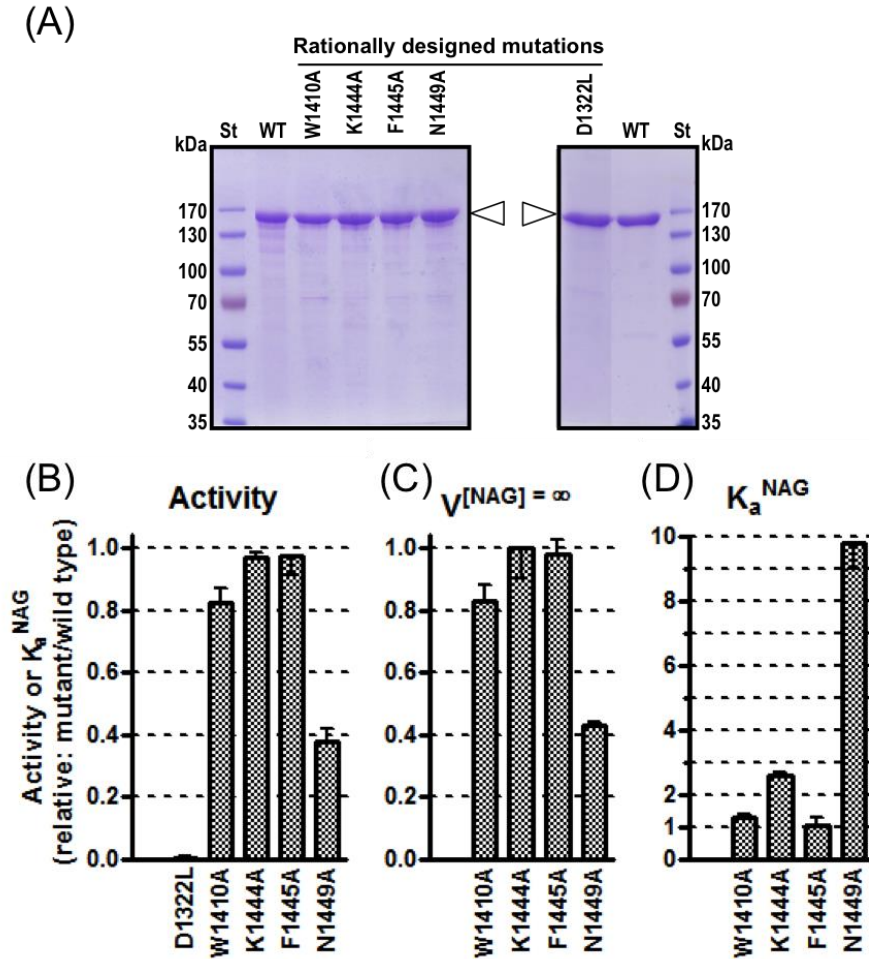


Figure 3. Production, activity and kinetic properties of ASD mutant forms not identified in CPSID. (A) SDS-PAGE (8% polyacrylamide, Coomassie staining) of the wild type enzyme (WT) and of the indicated mutants. St, protein markers (PageRuler Prestained Protein Ladder, from Thermo Scientific), with masses indicated in kDa. The arrowheads signal the position of the CPS1 band. (B-D) Effects of the ASD mutations tested on enzyme activity (B), apparent V_{max} (C), and K_a for NAG (D). Other details are as in Fig. 2.

the previously predicted interactions of R1371, T1391, T1394, N1437 and N1440 with NAG (Pekkala et al., 2009) (Fig. 4A). The ~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 $\beta 4$ - $\alpha 4$ loop (composed of residues 1438-1445), to which P1439 belongs, since this loop is an important part of the proposed NAG site (Figs. 1C and 4A). The localization of the NAG site was further supported by the results obtained here with the N1449A mutant (Fig. 3A). This mutant exhibited a ~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. 3A), which would alter residues forming a lid in the putative NAG site in our previous model (Fig. 1C), did not trigger important changes in specific activity or in NAG activation kinetics (Figs. 3B-D and Table 1B). Furthermore, the T1443A clinical mutation was reported to increase K_a^{NAG} ~160-fold (Table 1B) (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 suggested that in the NAG-bound form of the ASD, the final part of the $\beta 4$ - $\alpha 4$ loop changed its conformation relative to the NAG-free form (Xie et al., 2007), placing the side chain of T1433 near NAG and altering the location of the neighbouring W1410, K1444 and F1445 residues, which would fail to interact with the ligand (Fig. 4B).

2.3.3.5 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 molecular dynamics (MD) simulations of the CPS-NAG complex. The restraints were based on the K_a^{NAG} variations induced by the ASD mutations (Table 1) 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 [protein databank (PDB; www.rcsb.org) file 2YVQ] (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 (compare Figs. 4A and 4B). The γ -COO⁻ group of NAG interacts with the side chains of R1371 and T1394 in the more exposed region

RESULTS. Chapter 3

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 (Figs. 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 1).

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, r.m.s.d. ≤ 0.3 Å) (Fig. 4C). These poses generated by docking were very close to the energy-minimized conformation obtained by restrained MD (average r.m.s.d. 0.57 Å). 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).

2.3.4 Discussion

The present and earlier experimental results (Pekkala et al., 2010; Díez-Fernández et al., 2013 and 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 1A) (Pekkala et al., 2010). Three ASD clinical mutations destabilized CPS1 and eight hampered or abolished enzyme activity (Table 1A). 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 to 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 was already 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 β_4 - α_4 loop to accommodate NAG (Figs. 4A-D). The side-chains of F1445 and T1443 are those exhibiting the largest displacements (Figs. 4C and 4D), respectively pointing towards the NAG site and away from it in the NAG-free crystal form, and the reverse in the NAG-bound form (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 (Figs. 1C and 4E) carbamate phosphorylation domain, acting as an on/off switch for this catalytic domain, thus possibly being the first event in the NAG activation process.

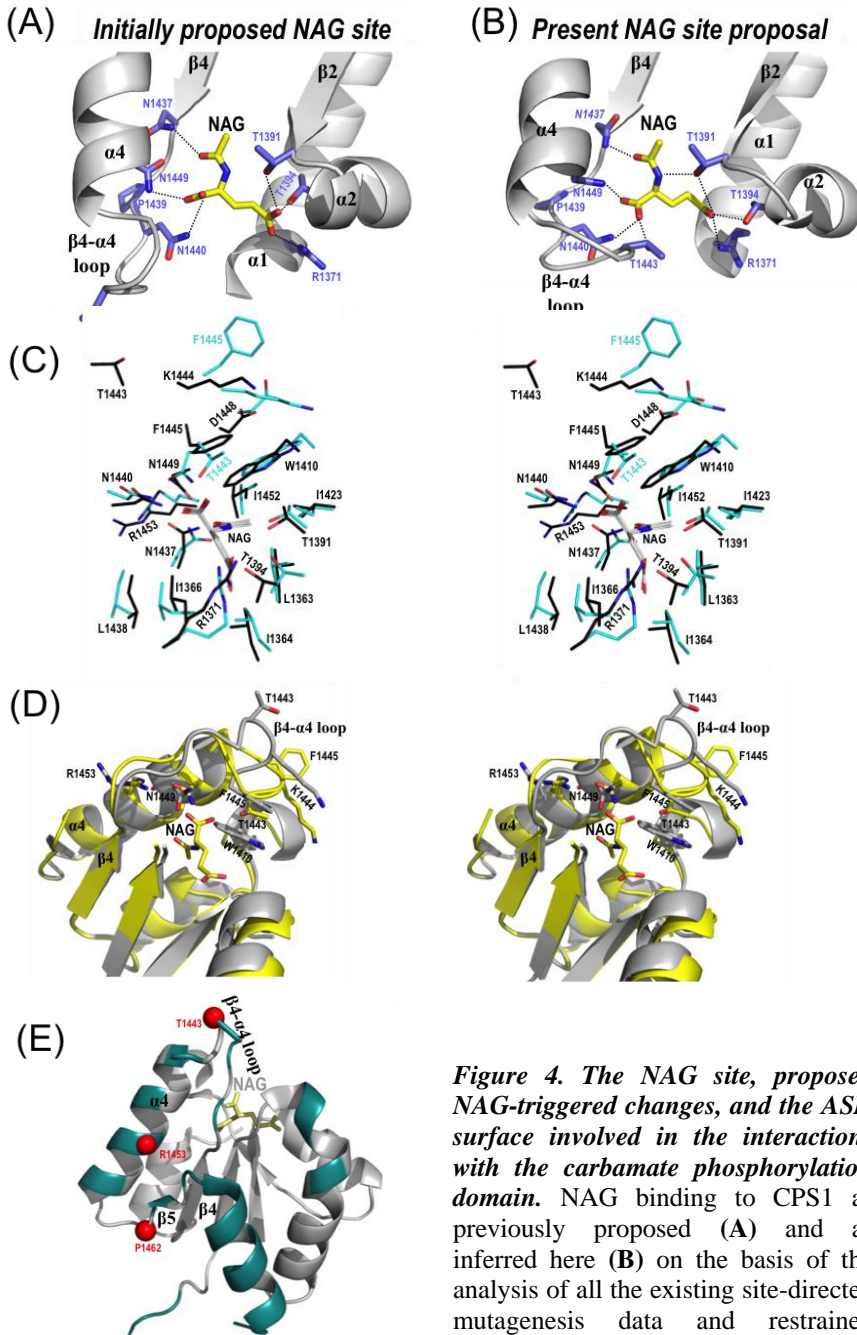


Figure 4. The NAG site, proposed NAG-triggered changes, and the ASD surface involved in the interactions with the carbamate phosphorylation domain. NAG binding to CPS1 as previously proposed (A) and as inferred here (B) on the basis of the analysis of all the existing site-directed mutagenesis data and restrained molecular dynamics (MD) simulations. (Cont'd on top of opposite page.

Figure 4 (cont'd). 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 of unrestrained docking of NAG to the hCPS1 protein structure obtained by restrained MD simulation. The 20 NAG binding poses (sticks with gray-colored carbon atoms) show 100% convergence. The amino acids side chains proposed to surround the NAG molecule are shown in thin sticks representation, in blue in the conformation resulting from restrained MD, superimposed on the conformation (in black) observed in the crystal structure of the NAG-free site (PDB file 2YVQ). (D) Stereo view 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.

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.

2.3.5 Materials and Methods

2.3.5.1 Human CPS1 production

Pure recombinant human mature liver CPS1 (CPS1) production with an N-terminal His₆-tag, and site-directed mutagenesis of the pFastBac-CPS1 vector encoding the enzyme to yield the desired human CPS1 mutant forms were performed as previously described (Díez-Fernández et al., 2013) using for the mutations appropriate oligonucleotides (Table 2). The same purification procedure (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 (1976).

Table S1. Synthetic oligonucleotides used in site-directed mutagenesis

p.D1322L	Forward	CCCCGGTTGAGGGATGCTCT CCCCATTCTGAG
	Reverse	CTCAGAATGGGG GAGAG CATCCCTCAACCGGGG
p.R1371L	Forward	CCAGCAATCATT CTGCCA AGAT
	Reverse	ATCTTGGCAGGAATGATTGCTGG
p.T1391M	Forward	GCTGTTTGCCAT GGAAGCC ACATC
	Reverse	GATGTGGCTTCCATGGCAAACAGC
p.L1398V	Forward	CCACATCAGACTGGG TCAAC GCC
	Reverse	GGCGTT GACCC AGTCTGATGTGG
p.W1410A	Forward	CACCCAGTGGCAG CGCCG TCTCAAGAAGG
	Reverse	CCTTCTTGAGACGG CGCTGCC ACTGGGGTGGC
p.P1439L	Forward	GTGATTAACCTTCTCAACAACAACAC
	Reverse	GTGTTGTTGTT GAGAAGG TTAATCAC
p.K1444A	Forward	CAACAACACT GCATTT GTCCATGATAATTATG
	Reverse	CATAATTATCATGGACAAAT GCAGT GTTGTTG
p.F1445A	Forward	CAACAACAACACTAAAG CTGTCC ATGATAATTATG
	Reverse	CATAATTATCATGGACAG CTTTA GTGTTGTTGTTG
p.N1449A	Forward	CTAAATTTGTCCATGAT GCTT ATGTGATTCGGAGG
	Reverse	CCTCCGAATCACATA AGCAT CATGGACAAATTTAG
p.P1462R	Forward	GTGGAATCC GTCTCCT ACTAATTTTC
	Reverse	GAAAATTAGTGAGGAG ACG GATTCCAC

^aBold type indicates base substitutions to introduce the desired mutation

2.3.5.2 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 mM glycyl-glycine pH 7.4, 70 mM KCl, 1 mM dithiothreitol, 20 mM MgSO₄, 5 mM ATP/5mM MgCl₂, 35 mM NH₄Cl, 50 mM KHCO₃, 10 mM NAG, 5 mM 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 mM MgSO₄ was present in the reaction mixture, Mg²⁺ was always in 20 mM 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 the means±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 mM glycyl-glycine pH 7.4, 10% glycerol, 0.5 M NaCl, 30 mM imidazole and 2 mM dithiothreitol were incubated 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 means±SE for at least duplicate assays.

2.3.5.3 Restrained molecular dynamics (MD) and docking calculations

We first used MD simulations with restraints based on the K_a^{NAG} variations induced by CPS mutations (Table 1) and on the contacts obtained in previous unrestrained docking runs (Pekkala et al., 2009) (see below). 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 ¹⁴¹⁵GQNPS¹⁴¹⁹ sequence missing in the crystallographic structure was modelled with MOE (CCG Inc.) using a method based on PDB

RESULTS. Chapter 3

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 $\beta 1$ - $\alpha 1$, $\beta 2$ - $\alpha 2$, $\beta 3$ and $\beta 4$ - $\alpha 4$ 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 Table 1 and the corresponding hydrogen-bonding contacts observed in unrestrained docking analyses (Pekkala et al., 2009): $O_{\square\square\text{COO}}^{\text{NAG}}\text{-HN}_{\eta}^{\text{R1371}}$, $\text{HN}^{\text{NAG}}\text{-O}_{\gamma}^{\text{T1391}}$, $O_{\gamma\square\text{COO}}^{\text{NAG}}\text{-HO}_{\gamma}^{\text{T1394}}$, $O_{\text{Ac}}^{\text{NAG}}\text{-HN}_{\delta}^{\text{N1437}}$ and $O_{\alpha\square\text{COO}}^{\text{NAG}}\text{-HN}_{\delta}^{\text{N1440}}$. An additional $O_{\alpha\square\text{COO}}^{\text{NAG}}\text{-HO}_{\gamma}^{\text{T1443}}$ restraint was imposed on the remaining $O_{\alpha\square\text{COO}}$ 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 hCPS-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).

2.3.5.4 Other techniques

The experimental crystal structure of the ASD (PDB file 2YVQ) (Xie et al., 2007) and the corresponding models for the NAG-bound form (Pekkala et al., 2009; and the present work) and the *E. coli* CPS structure-based CPS1 model (Martínez et al., 2010) have been used for analysis, superimpositions, and for structural representations, using Pymol (DeLano Scientific; <http://www.pymol.org>). 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. 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.

2.3.6 Acknowledgements

We thank Belén Barcelona (IBV-CSIC, Valencia) for the structural model of the complete enzyme and for help with Fig. 4D. This work was supported by grants from the Fundación Alicia Koplowitz, the Valencian (PrometeoII/2014/029 to V.R.) and Spanish governments (BFU2011-30407 to V.R. and BFU2012-30770 to J.G.), and the Swiss National Science Foundation (grant 310030_127184), to J.H. C.D-F. was a FPU fellow of the Spanish Government and received from that Government a bursary for short-time work in Zurich. The original mutation analysis of some CPS1D patients was kindly supported by Orphan Europe. This support does not involve any conflict of interest.

2.3.7 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 CR, Beaudet AL, Sly WS, Valle D, editors; Child B, Kinzler KW, Vogelstein B, associated editors. *The Metabolic and Molecular Bases of Inherited Disease*, 8e. New York: McGraw-Hill. Vol 2, p 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

RESULTS. Chapter 3

- 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. Chapter 4: Hyperammonemia and related disorders. In Blau N, Duran R, Gibson KM, Blaskovics M, Dionisi-Vici C, editors. *Physician's Guide to the diagnosis, treatment and follow-up of Inherited Metabolic Diseases*, Springer-Verlag Heidelberg, 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, editors. *The Urea Cycle*. John Wiley and Sons, New York 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

RESULTS. Chapter 3

- 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., European Registry and Network for Intoxication Type Metabolic Diseases (E-IMD). <http://www.e-imd.org/en/index.phtml>; Members of the Urea Cycle Disorders Consortium (UCDC). <http://rare diseasesnetwork.epi.usf.edu/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., Holzknacht, 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., 2007. RIKEN Structural Genomics/Proteomics Initiative. Crystal structure of MGS domain of carbamoyl-phosphate synthetase from *Homo sapiens*. Protein DataBank file 2YVQ, <http://www.rcsb.org/pdb>.

3. Discussion

DISCUSSION

Since the main body of this PhD dissertation consists of three publications having their independent discussions, in this global discussion we will try to avoid as much as possible repetition of topics already discussed, dealing with a limited number of relevant and untreated issues.

A point that we have not addressed in the Results section concerns the bases of the reported impact on vascular pathology of a widespread polymorphism, p.Thr1406Asn, that maps in the allosteric domain of the human enzyme [134-139]. We have worked here with the most frequent form of this polymorphism, the threonine variant (rs1047891, Ensemble database; allelic frequencies of Thr/Asn forms, 0.7/0.3; http://www.ensembl.org/Homo_sapiens/Variation/Explore?db=core;r=2:210675283-210676283;v=rs1047891;vdb=variation;vf=823866), since Powers-Lee's group, using the *S. pombe*-expressed enzyme, only reported minor effects of the p.Thr1406Asn substitution: ~20% decrease in k_{cat} and 1.2-fold, 1.7-fold and 1.5-fold increases in the values for the K_m^{ATP} , $K_m^{ammonia}$ and K_a^{NAG} , respectively [140]. Given the modesty of these changes and the large amount of extra time and effort needed, we decided to set aside the idea of introducing all the presently studied mutations in one and the other allelic forms of the enzyme for this polymorphism.

Another question not dealt with in the Results section is the path that led us to utilize the baculovirus/insect cell system for expression of CPS1. We needed a system capable of producing large amounts of the pure human protein and that allowed mutations to be introduced at will. These requirements were essential for proper characterization of the enzyme activity and, where necessary, of the partial activities of the enzyme (much less active than the complete reaction). It was also convenient to have a highly pure enzyme preparation for determining the kinetic parameters of even very poorly active enzyme mutants. In addition, it was anticipated that a robust system was needed to be able to isolate the tiny fraction of the enzyme that might be anticipated to be soluble with some of the mutants. Another goal of these studies was crystallographic, which was also expected to require large amounts of highly pure protein even with the miniaturization and automation of present-day crystallographic techniques.

Early attempts of recombinant expression in our group were carried out in Dr. Cervera's laboratory using cloned mature rat CPS1. *E. coli*, the methanol-consuming yeast *Pichia pastoris* and human embryonic kidney 293T (HEK293T) cells were used without real success, since the expression levels of soluble protein were always low (unpublished results). The finding by Powers-Lee [163] that the recombinant mature CPS1 from *Rana catesbeiana* could be expressed in *Schizosaccharomyces pombe* led the group of Dr. Cervera to begin

DISCUSSION

the process of attempting this expression in *S. pombe*, simultaneously with an attempt of using baculovirus and insect cells. The latter cells appeared promising: their animal nature, the strength of the promoters used and the reputation of having an excellent folding machinery (see for example http://www.embl.de/pepcore/pepcore_services/cloning/choice_expression_systems) made them good (although expensive and time-consuming) candidates for expressing animal CPS1. Good results were obtained with this system before the *S. pombe* system was ready in our hands, and this led us to adopt exclusively the baculovirus/insect cell expression system.

We feel that the utilization of this expression system has been rewarding concerning the understanding of the disease-causing role of CPS1D-associated mutations. Up to now our group has analyzed, either using the rat [14] or human recombinant CPS1 (Chapters 1-3 of the Results), 40 clinical missense mutations and 2 polymorphisms. While our studies proved minor or no differences relative to the wild-type in activity or stability of the two polymorphic forms, most of the clinical mutations studied revealed a disease-causing role either because of gross misfolding, important reduction in enzyme stability, inactivation of the enzyme without evidence for misfolding, reduced k_{cat} , or reduced apparent affinity for a substrate or for the essential CPS1 activator NAG. In some cases several of these effects coexisted for the same mutant as for many mutations of the integrating domain (Chapter 2 of the Results). For these mutants highly decreased amounts of soluble enzyme were produced, indicating an unfavorable partitioning of the mutant enzyme between the well-folded and misfolded forms during the folding process. However, our observations showed also that the soluble, apparently well-folded fraction of the enzyme exhibited in many cases important kinetic aberrations.

In any case, the overwhelming evidence extracted from our experiments is that most of the mutations reported in patients with CPS1D [12] that have been examined here were associated with important aberrations detected "*in vitro*", clearly indicating a disease-causing role. This evidence challenges our initial proposal that missense mutations found in patients are uncertain to have a disease-causing role. Since most reported missense mutations that we have examined "*in vitro*" are pathogenic, it is likely that the vast majority of all the mutations in the patient database are pathogenic. Indeed, evidence against their being polymorphisms (at least frequent polymorphisms) is growing with the growth of human genomes and exomes being analyzed, since otherwise it would be increasingly possible that they are found in homozygosis in some of the persons that have been fully sequenced already (see for example http://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=Graphics&list_uids=1373). Nevertheless, this tentative conclusion should be nuanced by the fact that

only two non-catalytic domains and a small selection of mutations in other domains has been studied "*in vitro*" thus far. Dr. Rubio's laboratory, in collaboration with Dr. Häberle, are continuing studies in other CPS1 domains to try to substantiate or refute this provisional conclusion. In any case, predictions based on *in vitro* results must be prudent, given the still unknown elements that may be behind some still rare but nevertheless real cases of variability in the clinical presentation associated with the same mutant alleles, as reported for two splicing mutations in two cases of CPS1 deficiency [209].

Apart from testing the disease-causing role of CPS1D mutations, the baculovirus/insect cell expression system used here enabled us to deepen into our understanding of the roles of CPS1 domains and of CPS1 activation by NAG. For example, although we have studied few mutations affecting the ~40 kDa N-terminal region of unknown function of CPS1 (Chapter 1 of the Results), the results with those mutations studied here fit the lack of catalytic or substrate and NAG-binding function of this region. In contrast, a clear function of this region on enzyme stabilization and on activation of catalysis (k_{cat} effect) has been documented here. The stabilizing role of this region recalls the strong stabilization triggered in *E. coli* CPS upon association of the small and large subunits of this bacterial enzyme [69]. The activating role of this region is also supported by prior studies of our laboratory [14] with the rat liver enzyme, in which the mutations p.Ser123Phe and p.His337Arg were found to importantly reduce k_{cat} without affecting the apparent affinities for the substrates or for NAG. This activating role is supported also by the previous finding of Powers-Lee's group [140] that the removal of the entire N-terminal region of the human enzyme decreased 700-fold enzyme activity. With *E. coli* CPS it had been found that the occupation of the glutamine site in the small subunit of the enzyme (which corresponds to the N-terminal region of CPS1) activates catalysis by the large subunit [6], an effect that might be reminiscent of the N-terminal region-dependent activation observed with CPS1. In summary, our present findings and prior observations of our laboratory and of Powers-Lee's group clearly exclude an absolute requirement of the N-terminal region for catalysis, but they provide strong reasons (stability and activation) for the retention by CPS1 of this N-terminal region despite the loss of the glutaminase activity associated to this region in all CPSs except CPS1 [1-3]. In any case, more experiments are required to further substantiate the role in CPS1 of the most N-terminal domain of this N-terminal region, since the glutaminase activity-attenuating function reported for this domain in the glutamine-utilizing CPSII [160] cannot operate in CPS1, which lacks glutaminase activity.

Concerning the phosphorylation domains of CPS1, our site-directed mutagenesis studies have been restricted to two mutations in the bicarbonate

DISCUSSION

phosphorylation domain. The reason for not exploring more these two homologous domains is the fact that, given the identity of function of these domains with the corresponding domains of other CPSs [1-3], less novel information could be gathered from mutations mapping in these domains. Actually, there is a long history of mutational studies with substitutions in these domains of other CPSs, particularly of *E. coli* CPS [92,153,154,210-213]. Furthermore, since these domains closely resemble biotin carboxylase [82], an enzyme component that also catalyzes bicarbonate phosphorylation [83,84], the mutations introduced into this last enzyme [214,215] also provided information on the functions of the corresponding residues of CPS1.

In any case, of the two mutations affecting the bicarbonate phosphorylation domain of human CPS1 that were introduced in the present studies (Chapter 1 of the Results), p.Ala438Pro inactivated the enzyme and p.Thr544Met had paramount effects on the apparent K_m value for bicarbonate (~60-fold increase) and on the K_a value for NAG (~25-fold increase), whereas the V_{max} was decreased ~five-fold and the K_m^{ATP} was increased ~4-fold. These results suggest a primary affectation of the site for bicarbonate with a secondary effect on the affinity for NAG stemming from the cross-talk between the bicarbonate phosphorylation site and the NAG site [40,61]. Interestingly, in pioneer work [87], our group introduced in *E.coli* CPS the p.Ala126Met mutation, aiming at replicating in this bacterial enzyme the p.Thr544Met mutation of human CPS1. The observed effects were similar to those found here with the human enzyme: the K_m values for bicarbonate and for ATP were increased 15-fold and 4-fold, respectively, and the V_{max} was decreased, in that case 15-fold. Furthermore, ornithine activation of the bicarbonate phosphorylation step was hampered (tested at a single concentration of ornithine), an effect that may be reminiscent of the negative effect of p.Thr544Met on the activation by NAG of human CPS1.

These findings with the p.Thr544Met and p.Ala126Met respective mutations of human and *E. coli* CPSs provide strong support for the utilization of *E. coli* CPS as a surrogate of human CPS1. Nevertheless, they clearly show that the effects of placing the same amino acid residue at a given location in the sequence may differ quantitatively depending on whether the enzyme is human CPS1 or bacterial CPS. In this work no other phosphorylation domains mutations have been examined for their effects on both the human and the bacterial enzymes. However, our group introduced [14] in rat CPS1 the human clinical mutations affecting the bicarbonate phosphorylation domain p.Thr471Asn, p.Gln678Pro and p.Pro774Leu, whereas the mutations corresponding to the last two of these mutations, p.Gln262Pro and p.Pro360Leu were introduced into the bacterial enzyme [87], studying their effects. Of these

mutations p.Gln678Pro and its bacterial counterpart p.Gln262Pro had as their major effects the dramatic destabilization of the corresponding enzymes. Human mutation p.Pro774Leu and the bacterial counterpart p.Pro360Leu dramatically decreased enzyme activity, in the first case not permitting activity detection, whereas in the second the detection was possible but it was heavily dependent on the presence of the activator ornithine. With the bacterial enzyme it was concluded [87] that the equilibrium between the inactive and active conformations of the enzyme that is displaced by ornithine towards the active form was heavily displaced in the mutant in the direction of the inactive conformation. Given the fact that CPS1 is much less active in the absence of the activator NAG than the *E. coli* enzyme in the absence of ornithine, the lack of activity of the rat enzyme with the mutation p.Pro774Leu might stem from a similar displacement of the equilibrium of the mutant CPS1 towards the inactive form. It is interesting that the p.Thr471Asn mutation also caused in the rat enzyme a large decrease in the affinity for NAG, highlighting the importance of the bicarbonate phosphorylation domain in the process of CPS1 activation by NAG, leading to propose that for this mutant the cross-talk between the phosphorylation site of the bicarbonate phosphorylation domain and the NAG binding site in the C-terminal domain of CPS1 is hampered. Therefore, the findings concerning mutants of the bicarbonate phosphorylation domain of CPS1 not only stress the expected role of this domain in bicarbonate binding and phosphorylation, but they also show the importance of this domain for ensuring a good affinity of the enzyme for NAG and good NAG activation. In addition, the comparison of the results with the corresponding mutations in the bacterial and rat enzymes further supports the value of the bacterial enzyme as a model of human CPS1, while highlighting that quantitative differences should be expected even for qualitatively similar effects with one and the other enzyme.

Concerning the carbamate phosphorylation domain, we have not introduced any mutation mapping therein in rat or human CPS1. However, our group introduced in *E. coli* CPS [87] three mutations, p.Val640Arg, p.Arg675Leu and p.Ser789Pro, that are the counterparts of the human CPS1 mutations affecting the carbamate phosphorylation domain p.Ile1054Arg, p.Arg1089Leu and p.Ser1203Pro. Of these *E. coli* CPS mutations, p.Ser789Pro selectively abolished catalysis of the carbamate phosphorylation step, whereas the other two mutations drastically decreased the activity for this step, also decreasing the affinity for the nucleotide and for carbamate (monitored as affinity for carbamoyl phosphate). Studies of the two partial reactions of the enzyme that reflect the individual steps of bicarbonate phosphorylation and carbamate phosphorylation showed that the p.Val640Arg mutation (and, by inference, the human CPS1 p.Ile1054Arg mutation) dramatically slows both phosphorylation

DISCUSSION

steps, leading to the conclusion [87] that this mutation, by primarily hampering the opening of the carbamate phosphorylation site to allow product release and nucleotide binding, also hampers the opening of the bicarbonate phosphorylation site. This supports the view (based on earlier mechanistic studies of the CPS1 reaction [52-55]) that this opening of both phosphorylation sites is a concerted process occurring simultaneously at both nucleotide sites. The effects of the p.Arg675Leu mutation (and of p.Arg1089Leu of human CPS1) points to selective near-abolition of catalysis at the carbamate phosphorylation site, whereas the effects of the p.Ser789Pro (and thus of p.Arg1089Leu of human CPS1) mutation were attributed to disturbance of the K⁺ loop of the bicarbonate phosphorylation domain, a loop that is considered essential for activity [77].

The only mutations affecting the carbamate phosphorylation domain that were studied directly on the recombinant human enzyme were the p.Cys1327Ala and p.Cys1337Ala mutations [164], two cysteine residues that change strongly their accessibility and even their ability to form an intramolecular disulfide bridge depending on whether NAG is bound or is not bound to the enzyme [213,214]. The p.Cys1327Ala mutation decreased 5-fold catalytic efficiency but it increased 7-fold the apparent affinity for NAG. This mutation was proposed to decrease protection of the enzyme against oxidative insults [164], something that could also apply to the p.Cys1327Arg mutation reported in a patient with CPS1D [12]. Nevertheless, the inactivating effect of the oxidation of these cysteines might be due to their cross-linking by a disulfide bridge [216,217], which probably compromises NAG activation of the enzyme (see below).

Given the extensive treatment and discussion in Chapter 2 of the Results of the numerous mutations introduced into the Integrating domain, we will make little mention here of this domain. As it might have been anticipated from the central location of this domain, which makes extensive contacts with both phosphorylation domains and which also contacts the N-terminal region of the enzyme (see Chapter 2 of the Results), a strong integrating and stabilizing function has been detected for this domain. Nevertheless, although misfolding and/or destabilization are the major effects of mutations falling on this domain, being the reason for the prominence of this domain in the mutational landscape of CPS1D patients [12], decreased V_{\max} and even decreased apparent affinity for NAG and ATP have been observed with some of the mutations mapping therein. These observations suggest that this domain may be involved directly or indirectly in the cross-talk between the NAG site and the catalytic centers and in the concerting of the opening of both active centers in the two phosphorylation domains [87, 88]. Furthermore, this domain contacts the active site of the bicarbonate phosphorylation domain, which is to be opened [88].

Similarly, we have dealt extensively in Chapter 3 of the Results with the 20 kDa C-terminal domain and thus we will make little mention of it here. This domain is known to harbor the binding sites for effectors in different CPSs [13,17,68-71,89-92]. The relatively low frequency of mutations mapping in this domain that cause misfolding possibly account for the lack of particular prominence of this domain in the clinical mutation landscape [12]. However, the frequency among the mutations affecting this domain of those that decrease the affinity for NAG endow this domain with particular therapeutic potential, rendering these mutations good candidates for using N-carbamylglutamate (NCG) for helping saturate in these cases the NAG site. It has to be remembered that, given the protein restriction that is generally applied to patients with urea cycle disorders [27,30], NAG levels might be low in these treated patients even if ammonia levels were increased. If this were the case, patients carrying NAG site-affecting mutations would be negatively affected by the low NAG levels, and in these cases NCG might be highly beneficial. Actually, it is to be considered whether mutations that do not affect the NAG site but that nevertheless increase the K_a for NAG because of the cross-talk between phosphorylation and allosteric domains might also benefit from NCG saturation therapy.

An important inference resulting from the present studies concerns the mechanism of NAG activation of CPS1. Clearly, a movement in the $\beta 4$ - $\alpha 4$ loop belonging to the NAG site but being at the interface with the carbamate phosphorylation domain appears to be a key early event in the process of activation (Chapter 3 of the Results and Fig. 23). A loop of the carbamate phosphorylation domain (residues 1318-1332) and an interacting surface of the allosteric domain have also been identified here as key elements in the transmission of the allosteric signal from this last domain to the carbamate phosphorylation domain. Interestingly, Cys1327 belongs to the 1318-1332 loop, explaining why the trapping of this loop in an inappropriate conformation due to disulfide bridge formation with the nearby Cys1337 might prevent activation, causing loss of enzyme activity [216,217]. Interestingly, a loop (residues 778-787) of the homologous bicarbonate phosphorylation domain has been proposed here (Chapter 3 of the Results) as the potential receiver of the allosteric signal in that phosphorylation domain, with passage of the signal via the 1318-1332 loop of the carbamate phosphorylation domain. Clearly, we do not know the details of this signal transmission process, but we have identified some crucial elements involved in such transmission. From our data and the structural model for human CPS1 it can be anticipated that NAG has long-range

DISCUSSION

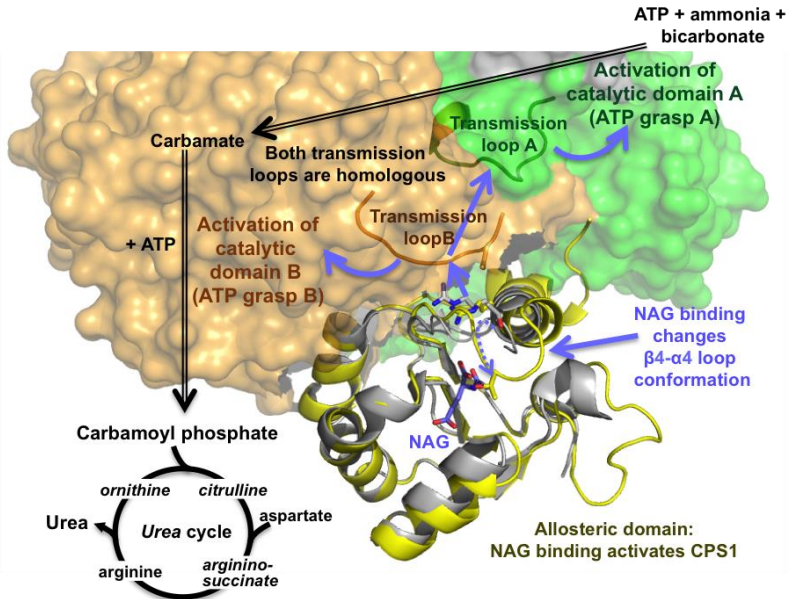


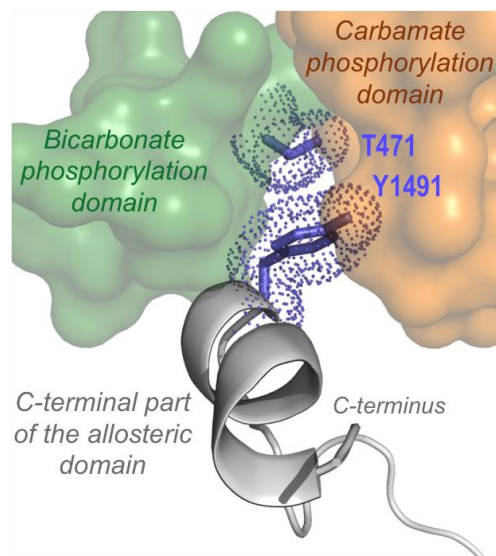
Figure 23. Graphic scheme of the CPS1/NAG switch for urea cycle control. Structural model of human CPS1 [97] in transparent surface representation, with the allosteric domain in cartoon representation and shown in the NAG-free form (gray, PDB 2YVQ) and in the modelled NAG- and ADP-bound form (in yellow). The bound NAG molecule and the side chains of Thr1443 and Arg1453 (not labelled for clarity) are shown in sticks representation. Solid blue arrows indicate the proposed allosteric signal transmission path, whereas the broken arrow indicates the β_4 - α_4 loop movement. Transmission loops are shown in backbone string representation and are labeled. Black empty arrows schematize the path of carbamate from the bicarbonate phosphorylation site to the carbamate phosphorylation site. Green and brown denote one and the other phosphorylation domains. A mini urea cycle is included to symbolize the cycle-controlling role of the CPS1 step

effects, with propagation of a movement in the peripheral allosteric domain to both phosphorylation centers. Given the mutual homology of both phosphorylation domains [1,74,77], the 1318-1332 and the 778-787 loops are homologous loops. Thus, these two loops may be expected to have similar interactions with other elements of the phosphorylation domain to which each loop belongs. Judged from the *E. coli* CPS structure [77], these two loops contact in both phosphorylation domains with the K^+ loop and with other elements of the nucleotide site that are key elements of the catalytic machinery of these domains. This gives further credence to the view that similar NAG-

DISCUSSION

triggered conformational changes in these two homologous loops could have similar effects on the corresponding phosphorylation domains. Thus, the 1318-1332 and the 778-787 loops can be called signal transmission loops (Fig. 23). Since they are also found in *E. coli* CPS [77], they may transmit the allosteric signals that operate in other CPSs. The identification of these two loops paves the way for further site-directed mutagenesis studies for exploration of the signal transmission process within the catalytic parts of the enzyme molecule.

Despite all the above, a picture in which the cross-talk between the NAG site and both phosphorylation domains occurs exclusively via the 1318-1332 and the 778-787 loops may be oversimplistic, given the ~ 50 -fold increase in K_a^{NAG} and ~ 4 -fold decrease in k_{cat} caused by the p.Tyr1491His mutation [14]. This mutation maps nine residues upstream from the C-terminus, in a region that when digested in rat liver CPS1 resulted in a very important decrease in the



*Figure 24. Model of the most C-terminal part of the allosteric domain (in grey) and its modelled (on the basis of the *E. coli* CPS structure) relations with both catalytic domains (green and orange, in surface representation of the relevant parts of these domains) [97]. The side chains of the residues reported to be mutated in patients with CPS1D, T471 and Y1491 [12], are shown in blue sticks, with blue dots representation of its Van der Waals sphere to highlight their clamping between both phosphorylation domains. Note the long grey string that connects this residue to the remainder of the allosteric domain (not shown) to suggest that the catalytic domains may exert traction on this domain if, because of conformational changes, they move away from it, bringing with them the clamped Y1491.*

DISCUSSION

affinity for NAG [66]. It appears highly unlikely that this region, not represented in the crystal structure of the allosteric domain of human CPS1 [93], directly participates in the NAG site. If similar to *E. coli* CPS [77], this C-terminal part of the enzyme would project away from the allosteric domain body, sitting on both phosphorylation domains, with the phenolic ring of Tyr1491 sandwiched between them (Fig. 24). In this respect, the already discussed bicarbonate phosphorylation domain mutation p.Thr471Asn [12] affects a residue that in the structure of *E. coli* CPS (Thr56) belongs to a hydrophobic cluster that also includes Val1065, the bacterial counterpart of human CPS1 Tyr1491. Since, as mentioned above, the p.Thr471Asn mutation dramatically decreased the affinity for NAG of rat CPS1 [14], this hydrophobic cluster appears a clear candidate to being involved in the passage of signals between the bicarbonate phosphorylation domain and the enzyme C-terminus. In this respect it is worth remembering that NAG binding studies [61] revealed a strong increase in CPS1 affinity for NAG when ATP was added. Thus, ATP binding to the phosphorylation domains enhances NAG binding by the allosteric domain. Perhaps, if the C-terminal region of CPS1 anchors the allosteric domain on both phosphorylation domains as in *E. coli* CPS, this region could transmit signals to the core of the allosteric domain by pulling from it if ATP binding alters the relative disposition of the catalytic domains with respect to the allosteric domain.

Summarizing all our work, we have provided a relatively easy instrumental system for testing effects of amino acid mutation and perhaps also of amino acid posttranslational modifications on the catalytic and controlling functions of CPS1. This system, by providing very large amounts of pure human CPS1, has opened the way to determining the crystal structure of human CPS1, a feat that is presently at reach. Thus, in collaboration with my work, other group members (Luis Mariano Polo and Sergio de Cima) have succeeded in obtaining well-diffracting crystals of CPS1 grown in the presence and in the absence of mixtures of NAG, ADP and Mg^{2+} (Fig. 25). Determination of these crystal structures is likely to provide a highly detailed description of the process of NAG activation, helping answer many unclarified questions on CPS1 (for example, which is the reason for the high affinity of CPS1 for ammonia). I will be an author in the expected manuscript reporting the anticipated structural findings, but for reasons of time and since I will not be the main author, this manuscript could not be included into the present PhD dissertation.

I would like to end this discussion by highlighting the future potential of the application of the present system to human CPS1. It may help determine the role of posttranslational modifications on CPS1 function and via CPS1, on urea cycle control. As already mentioned in the Introduction, proteomics studies are

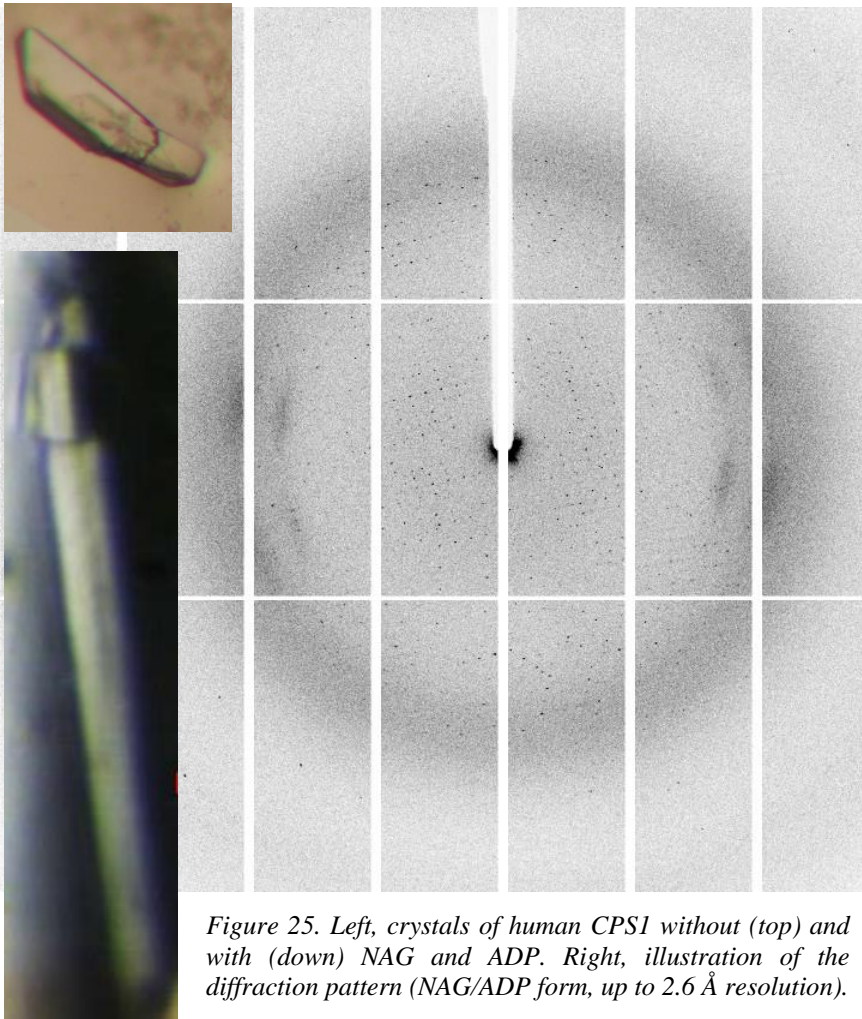


Figure 25. Left, crystals of human CPS1 without (top) and with (down) NAG and ADP. Right, illustration of the diffraction pattern (NAG/ADP form, up to 2.6 Å resolution).

expanding our view of the modifications observed in the whole human and animal proteome, and CPS1 has not escaped this trend, actually having been signalled by a number of studies [129-133] as the subject of changes having potentially high impact on urea cycle control and even linked to aging [129] and to secondary hyperammonemia [130]. Although the structure of CPS1 will eventually provide a reference frame for judging the potential impact of these modifications on human CPS1 activity and control, it would be essential to analyze each amino acid change in a clean system like the one used here to express CPS1. In a first approximation, this could be done by making amino acid replacements that could mimic the posttranslational modifications

DISCUSSION

(similarly to the substitution of phosphorylated serines by aspartate or glutamate) or even by exploiting new procedures for inserting in the sequence abnormal amino acids [218]. The long-term aim would be to reconstitute *in vitro* the whole posttranslational modifying/demodifying system, or to coexpress in the insect cells other elements of the modifying machinery.

In addition, our work has opened a way for future personalized treatment of CPS1D patients. In the case of phenylketonuria, a clinically useful achievement was the observation of phenylalanine hydroxylase stabilization by chaperoning by its essential cofactor tetrahydrobiopterin [32]. Likewise, in this work we proved the stabilization of CPS1 by its substrates and activators NAG and NCG (Chapter 1 of Results). However, the possibility of finding more efficient pharmacological chaperones can be now approached thanks to our expression system, and these could eventually be used to treat CPS1D patients with unstabilizing mutations. In the same way, as above mentioned, CPS1D patients with NAG site-affecting mutations and with mutations outside the NAG-binding domain but which are involved in the cross-talk between phosphorylation and allosteric domains could benefit from the commercially-available NCG saturation therapy. However, some mutations might totally hamper binding of activating molecules in the NAG-binding site. In these cases, other activating molecules could be tested, that, like glycerol, (Chapter 1 of Results) activate CPS1 to some extent by unknown mechanisms that do not appear to involve the NAG site. As a matter of fact, this new research line has already been started as part of my postdoctoral work supervised by Dr. Häberle.

4. Conclusions

CONCLUSIONS

CONCLUSIONS

1. Recombinant human CPS1 is efficiently produced in a commercial baculovirus/insect cell system. The recombinant enzyme mimics natural CPS1 in its kinetic and molecular properties and in needing the essential activator N-acetyl-L-glutamate (NAG) although cryoprotectants can partially replace NAG.
2. Limited proteolysis showed that recombinant CPS1 has the same domain composition as other CPSs, with interdomain hinges at 40 kDa and 20 kDa from the enzyme N-terminus and C-terminus, respectively.
3. This CPS1 production system is effective for testing the effects on the enzyme of the missense mutations found in patients with CPS1 deficiency (CPS1D). For patients with informative genotypes generally there was agreement between the severity of the effects *in vitro* and the seriousness of the patient phenotype. Lack of any *in vitro* effect of two trivial polymorphisms confirmed the high value of this system for discriminating between disease-causing and non-causing sequence variants of CPS1.
4. A disease-causing role has been proven for the majority of the missense mutations identified in CPS1D patients that were tested with this system. Some mutations caused gross misfolding or markedly decreased enzyme stability, and/or they inactivated the enzyme or triggered kinetic aberrations such as decreased apparent k_{cat} or decreased apparent affinity for one or the other of the three substrates or for the activator NAG.
5. The effects of the tested CPS1D mutations that mapped in the N-terminal 40-kDa region agree with the view that this region, despite having become catalytically inactive, has been preserved in CPS1 because of its stabilizing and enzyme-activating actions.
6. The effects of the tested mutations mapping in the bicarbonate phosphorylation domain (the more N-terminal domain of the 120-kDa C-terminal moiety of CPS1) affected bicarbonate binding and/or catalysis, but they also decreased the affinity for NAG, confirming the existence of cross-talk between the bicarbonate phosphorylation domain and the NAG-binding domain.

CONCLUSIONS

7. A small central domain of previously unknown function that is intercalated between both phosphorylation domains is concluded to have high mutational eloquence in CPS1D because its mutations frequently cause enzyme misfolding. Thus, this domain plays a paramount integrating role in the multidomain enzyme architecture, and it is called from here on the "Integrating Domain".
8. The low mutational eloquence in CPS1D of the C-terminal domain is possibly due to the fact that mutations mapping in this domain rarely destabilize importantly or cause gross misfolding of CPS1. Some of these mutations inactivate the soluble enzyme and most of them decrease enzyme activity, with many mutations strongly decreasing the affinity for NAG. Thus, the negative effects of most of these mutations are largely due to interference with NAG binding or to hampered transmission of the activating signal to the phosphorylation domains of the enzyme.
9. A mechanism for NAG activation based on the observation of the effects of missense mutations is proposed in which NAG triggers a movement of the $\beta 4$ - $\alpha 4$ loop of the C-terminal domain that drastically alters the interactions of this loop with the carbamate phosphorylation domain. Signal transmission involves Arg1453 from helix $\alpha 4$ and Asp1322 from the 1318-1332 loop of the carbamate phosphorylation domain. This loop, and the homologous 778-787 loop of the bicarbonate phosphorylation domain, are called here the signal transmission loops because they interact with the active centers of both phosphorylation domains and are proposed here to be the mediators of NAG activation.
10. The NAG analogue and orphan drug N-carbamyl-L-glutamate (NCG) is proposed to be potentially useful in the treatment of CPS1 deficiency due to mutations that decrease the affinity for NAG (saturation therapy) and of those that decrease enzyme stability (pharmacological chaperoning) since in the presence of ATP, NCG is found here to stabilize CPS1.

5. References (for the Introduction and the Discussion)

REFERENCES (Introduction and Discussion)

REFERENCES (Introduction and Discussion)

Note. These references are for the Introduction and General Discussion sections of this PhD dissertation. Individual chapters of the results section corresponding to papers published or submitted have their own reference lists

1. Rubio V. 1993. Structure-function studies in carbamoyl phosphate synthetases. *Biochem Soc Trans.* 21:198-202.
2. Rubio V. 1994. Carbon dioxide fixation and reduction in biological and model systems. In Bränden CI, Schneider G, editors. Oxford University Press, Oxford, pp. 249-264.
3. Rubio V, Cervera J. 1995. The carbamoyl-phosphate synthase family and carbamate kinase: structure-function studies. *Biochem Soc Trans* 23:879-883.
4. Makoff AJ, Radford A. 1978. Genetics and biochemistry of carbamoyl phosphate biosynthesis and its utilization in the pyrimidine biosynthetic pathway. *Microbiol Rev* 42:307-328.
5. Jones ME. 1980. Pyrimidine nucleotide biosynthesis in animals: genes, enzymes and regulation of UMP biosynthesis. *Annu Rev Biochem.* 49:253-279.
6. 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.
7. Ratner S. 1973. Enzymes of arginine and urea synthesis. *Adv Enzymol Relat Areas Mol Biol.* 39:1-90.
8. Cunin R, Glansdorff N, Pierard A, Stalon V. 1986. Biosynthesis and metabolism of arginine in bacteria. *Microbiol Rev.* 50:314-352.
9. Slocum RD. 2005. Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiol Biochem.* 43:729-745.
10. Grillo MA, Colombatto S. 2004. Arginine revisited: minireview article. *Amino Acids* 26:345-351.
11. Grisolia S, Rubio, V. 1986. Algunos aspectos del ciclo de la urea. In Ochoa S, Leloir LF, Oro J, Sols A, editors, Cornudella L coordinator. *Bioquímica y Biología Molecular. Temas de actualidad para graduados.* Salvat editores, Barcelona, pp 173-180.
12. Häberle J, Shchelochkov OA, Wand J, Katsonis P, Hall L, Reiss S, Eeds A, Willis A, Yadav M, Summar S, Consortium TUCD, Lichtarge O, Rubio V, Wong LJ, Summar M. 2011. Molecular defects in human carbamoyl phosphate synthetase I: Mutational spectrum, diagnostic and protein structure considerations. *Hum Mutat* 32:579-589.
13. Pekkala S, Martinez AI, 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

REFERENCES (Introduction and Discussion)

- essential allosteric activator of mitochondrial carbamoyl phosphate synthetase. *Biochem J.* 424:211–220.
14. Pekkala S, Martinez AI, Barcelona B, 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.
 15. Pierson DL, Brien JM. 1980. Human carbamylphosphate synthetase I. Stabilization, purification, and partial characterization of the enzyme from human liver. *J Biol Chem.* 255:7891-7895.
 16. 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.
 17. Rodriguez-Aparicio LB, Guadalajara AM, 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.
 18. Braissant O, McLin VA, Cudalbu C. 2013. Ammonia toxicity to the brain. *J Inherit Metab Dis.* 36:595-612.
 19. Chew SF, Ip YK. 2014. Excretory nitrogen metabolism and defence against ammonia toxicity in air-breathing fishes. *J Fish Biol.* 84:603-638.
 20. Needham, J. 1931. *Chemical embryology.* MacMillan, New York
 21. Campbell, JW. 1970. *Comparative Biochemistry of Nitrogen Metabolism. Volume 2. The Vertebrates.* Academic Press, New York.
 22. Kennedy J. 1976. Properties and chemistry of urea and related intermediates. In Grisolia S, Báguena R, Mayor F, editors. *The Urea Cycle.* John Wiley and Sons, New York, pp. 39-54
 23. Moffat DB. 1975. *The mammalian kidney.* Cambridge University Press, Cambridge.
 24. Earley LE. 1968. Effects of uremia on the cardiovascular and central nervous systems. *Surg Clin North Am.* 48:371-380.
 25. Krebs H, Henseleit K. 1932. Untersuchungen über die Harnstoffbildung im Tierkörper. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* 210:325-332.
 26. Krebs HA 1976. The discovery of the ornithine cycle. In Grisolia S, Báguena R, Mayor F, editors. *The Urea Cycle.* John Wiley and Sons, New York, pp. 1-12.
 27. Häberle J, Rubio V. 2014. Chapter 4: Hyperammonemia and related disorders. In Blau N, Duran R, Gibson KM, Blaskovics M, Dionisi-Vici C, editors. *Physician's Guide to the diagnosis, treatment and follow-up of Inherited Metabolic Diseases,* Springer-Verlag Heidelberg, pp. 67-72.
 28. Haussinger D, Gerok W, Sies H. 1984. Hepatic role in pH regulation: role of the intercellular glutamine cycle. *Trends Biochem Sci.* 9:300-302.

REFERENCES (Introduction and Discussion)

29. Nuzum CT, Snodgrass PJ. 1976. Multiple assays of the five urea cycle enzymes in human liver homogenates. In Grisolia S, Báguena R, Mayor F, editors. *The Urea Cycle*. John Wiley and Sons, New York pp. 325-349.
30. Häberle J, Boddaert N, Burlina A, Chakrapani A, Dixon M, Huemer M, Karall D, Martinelli D, Sanjurjo Crespo P, Santer R, Servais A, Valayannopoulos V, Lindner M, Rubio V, Dionisi-Vici C. 2012. Suggested Guidelines for the diagnosis and management of urea cycle disorders. *Orphanet J Rare Dis*. 7:32.
31. Rubio V, Grisolia S. 1981. Treating urea cycle defects. *Nature* 292:496.
32. Erlandsen H, Pey AL, Gámez A, Pérez B, Desviat LR, Aguado C, Koch R, Surendran S, Tying S, Matalon R, Scriver CR, Ugarte M, Martínez A, Stevens RC. 2004. Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. *Proc Natl Acad Sci USA*. 101:16903–16908.
33. 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.
34. Stewart PM, Walser M. 1980. Short term regulation of ureagenesis. *J Biol Chem*. 255:5270-5280
35. Bender DA. 2012. *Amino acid metabolism*. 3rd edition. Wiley-Blackwell (an imprint of John Wiley and Sons), Chichester, UK.
36. Shigesada K, Tatibana M. 1978. N-Acetylglutamate synthetase from rat-liver mitochondria. Partial purification and catalytic properties. *Eur J Biochem*. 84:285-291.
37. Morita T, Mori M, Tatibana M. 1982. Regulation of N-acetyl-L-glutamate degradation in mammalian liver. *J Biochem*. 91:563-569.
38. Anderson PM. 1980. Glutamine- and N-acetylglutamate-dependent carbamoyl phosphate synthetase in elasmobranchs. *Science* 208:291-293.
39. Guthöhrlein G, Knappe J. 1968. Structure and function of carbamoyl phosphate synthase. I. Transitions between two catalytically inactive forms and the active form. *Eur J Biochem*. 7:119-127.
40. Rubio V, Britton HG, 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.
41. Grisolia S, Cohen PP. 1951. Study of citrulline synthesis with soluble enzyme preparations. *J Biol Chem*. 191:189-202.
42. Jones ME, Lipmann F. 1960. Chemical and enzymatic synthesis of carbamyl phosphate. *Proc Natl Acad Sci USA*. 46:1194-1205.

REFERENCES (Introduction and Discussion)

43. Hall LM, Metzenberg RL, Cohen PP. 1958. Isolation and characterization of a naturally occurring cofactor of carbamyl phosphate biosynthesis. *J Biol Chem.* 230:1013-1021.
44. Metzenberg RL, Marshall M, Cohen PP. 1958. Purification of carbamyl phosphate synthetase from frog liver. *J Biol Chem.* 233:102-105.
45. Metzenberg RL, Marshall M, Cohen PP. 1958. Carbamyl phosphate synthetase: studies on the mechanism of action. *J Biol Chem.* 233:1560-1564.
46. Guadalajara A, Grisolia S, Rubio V. 1987. Limited proteolysis reveals low-affinity binding of N-acetyl-L-glutamate to rat-liver carbamoyl-phosphate synthetase (ammonia). *Eur J Biochem.* 165:163-169.
47. Jones ME. 1963. Carbamyl Phosphate: Many forms of life use this molecule to synthesize arginine, uracil, and adenosine triphosphate. *Science* 140:1373-1379.
48. Elliott KR, Tipton KF. 1974. Kinetic studies of bovine liver carbamoyl phosphate synthetase. *Biochem J.* 141:807-816.
49. Elliott KR, Tipton KF. 1974. Product inhibition studies on bovine liver carbamoyl phosphate synthetase. *Biochem J.* 141:817-824.
50. Elliott KR, Tipton KF. 1974. A kinetic analysis of enzyme systems involving four substrates. *Biochem J.* 141:789-805.
51. Clarke S. 1976. A major polypeptide component of rat liver mitochondria: carbamyl phosphate synthetase. *J Biol Chem.* 251:950-961.
52. Rubio V, Grisolia S. 1977. Mechanism of mitochondrial carbamoyl phosphate synthetase: synthesis and properties of active CO₂, precursor of carbamoyl phosphate. *Biochemistry* 16:321-329.
53. Rubio V, Britton HG, Grisolia S. 1979. Mechanism of carbamoyl-phosphate synthetase. Binding of ATP by the rat-liver mitochondrial enzyme. *Eur J Biochem* 93:245-256.
54. Britton HG, Rubio V, Grisolia S. 1979. Mechanism of carbamoyl-phosphate synthetase. Properties of the two binding sites for ATP. *Eur J Biochem.* 102: 521-530.
55. Rubio V, Britton HG, Grisolia S, Sproat BS, Lowe G. 1981. Mechanism of activation of bicarbonate ion by mitochondrial carbamoyl-phosphate synthetase: formation of enzyme-bound adenosine diphosphate from the adenosine triphosphate that yields inorganic phosphate. *Biochemistry* 20:1969-1974.
56. Powers SG, Meister A. 1978. Carbonic-phosphoric anhydride (carboxy phosphate). Significance in catalysis and regulation of glutamine-dependent carbamyl phosphate synthetase. *J Biol Chem.* 253:1258-1265.

REFERENCES (Introduction and Discussion)

57. Powers SG, Meister A. 1978. Mechanism of the reaction catalyzed by carbamyl phosphate synthetase. Binding of ATP to the two functionally different ATP sites. *J Biol Chem.* 253:800-803.
58. Rubio V, Britton HG, Rodríguez-Aparicio LB, Climent I. 1990. Carbamate synthases and kinases. In Aresta M, Schloss JV, editors. *Enzymatic and model carboxylation and reduction reactions for carbon dioxide utilization.* M Kluwer Academic Publishers, The Netherlands, pp. 221-228.
59. Wimmer MJ, Rose IA, Powers SG, Meister A. 1979. Evidence that carboxyphosphate is a kinetically competent intermediate in the carbamyl phosphate synthetase reaction. *J Biol Chem.* 254:1854-1859.
60. Raushel FM, Villafranca JJ. 1980. Phosphorus-31 nuclear magnetic resonance application to positional isotope exchange reactions catalyzed by *Escherichia coli* carbamoyl-phosphate synthetase: analysis of forward and reverse enzymatic reactions. *Biochemistry* 19:3170-3174.
61. Alonso E, Rubio V. 1983. Binding of N-acetyl-L-glutamate to rat liver carbamoyl phosphate synthetase (ammonia). *Eur J Biochem.* 135:331-337.
62. Britton HG, Rubio V, Grisolia S. 1981. Synthesis of carbamoyl phosphate by carbamoyl phosphate synthetase I in the absence of acetylglutamate. Activation of the enzyme by cryoprotectants. *Biochem Biophys Res Commun.* 99:1131-1137.
63. Rubio V, Britton HG, Grisolia S. 1983. Activation of carbamoyl phosphate synthetase by cryoprotectants. *Mol Cell Biochem.* 53-54:279-298.
64. Guadalajara A. 1987. Estudio, mediante proteólisis limitada, de la estructura de dominios, unión de ligandos y cambios asociados de conformación de la carbamil fosfato sintetasa de hígado de rata. PhD dissertation. University of Valencia.
65. Powers-Lee SG, Corina K. 1986. Domain structure of rat liver carbamoyl phosphate synthetase I. *J Biol Chem.* 261:15349-15352.
66. Marshall M, Fahien LA. 1988. Proteolysis as a probe of ligand-associated conformational changes in rat carbamyl phosphate synthetase I. *Arch Biochem Biophys.* 262:455-470.
67. Evans DR, Balon MA. 1988. Controlled proteolysis of ammonia-dependent carbamoyl-phosphate synthetase I from Syrian hamster liver. *Biochim Biophys. Acta* 953:185-196.
68. Rubio V, Cervera J, Lusty CJ, Bendala E, Britton HG. 1991. Domain structure of the large subunit of *Escherichia coli* carbamoyl phosphate synthetase. Location of the binding site for the allosteric inhibitor UMP in the COOH-terminal domain. *Biochemistry* 30:1068-1075.

REFERENCES (Introduction and Discussion)

69. Cervera J, Conejero-Lara F, Ruiz-Sanz J, Galisteo ML, Mateo PL, Lusty CJ, 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.
70. Bueso J, Lusty CJ, Rubio V. 1994. Location of the binding site for the allosteric activator IMP in the COOH-terminal domain of *Escherichia coli* carbamyl phosphates synthetase. *Biochem Biophys Res Commun.* 203:1083-1089.
71. Liu X, Guy HI, Evans DR. 1994. Identification of the regulatory domain of the mammalian multifunctional protein CAD by the construction of an *Escherichia coli*-hamster hybrid carbamyl-phosphate synthetase. *J Biol Chem.* 269:27747-27755.
72. Nyunoya H, Lusty CJ. 1983. The *carB* gene of *Escherichia coli*: a duplicated gene coding for the large subunit of carbamoyl-phosphate synthetase. *Proc Natl Acad Sci USA.* 80:4629-4633.
73. Piette J, Nyunoya H, Lusty CJ, Cunin R, Weyens G, Crabeel M, Charlier D, Glansdorff N, Pierard A. 1984. DNA sequence of the *carA* gene and the control region of *carAB*: tandem promoters, respectively controlled by arginine and the pyrimidines, regulate the synthesis of carbamoyl-phosphate synthetase in *Escherichia coli* K-12. *Proc Natl Acad Sci USA.* 81:4134-4138.
74. Nyunoya H, Broglie KE, Widgren EE, Lusty CJ. 1985. Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. *J Biol Chem.* 260:9346-9356.
75. 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.
76. 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.
77. Thoden JB, Holden HM, Wesenberg G, Raushel FM, Rayment I. 1997. Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. *Biochemistry* 36:6305-6316.
78. Thoden JB, Raushel FM, Benning MM, Rayment I, Holden HM. 1999. The structure of carbamoyl phosphate synthetase determined to 2.1 Å resolution. *Acta Crystallogr D Biol Crystallogr.* 55:8-24.
79. Marina A, Bravo J, Fita I, Rubio V. 1995. Crystallization, characterization, and preliminary crystallographic studies of mitochondrial carbamoyl phosphate synthetase I of *Rana catesbeiana*. *Proteins* 22:193-196.

REFERENCES (Introduction and Discussion)

80. Yamaguchi H, Kato H, Hata Y, Nishioka T, Kimura A, Oda J, Katsube Y. 1993. Three-dimensional structure of the glutathione synthetase from *Escherichia coli* B at 2.0 Å resolution. *J Mol Biol.* 229:1083-1100.
81. Fan C, Moews PC, Walsh CT, Knox JR. 1994. Vancomycin resistance: structure of D-alanine: D-alanine ligase at 2.3 Å resolution. *Science* 266:439-443.
82. Waldrop GL, Rayment I, Holden HM. 1994. Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry* 33: 10249-10256.
83. Climent I, Rubio V. 1986. ATPase activity of biotin carboxylase provides evidence for initial activation of HCO₃⁻ by ATP in the carboxylation of biotin. *Arch Biochem Biophys.* 251: 465-470.
84. Rubio V. 1986. Enzymatic HCO₃⁻ fixation: a common mechanism for all enzymes involved? *Biosci Rep.* 6:335-347.
85. Holden HM, Thoden JB, Raushel FM. 1998. Carbamoyl phosphate synthetase: a tunnel runs through it. *Curr Opin Struct Biol.* 8:679-685.
86. Holden HM, Thoden JB, Raushel FM. 1999. Carbamoyl phosphate synthetase: an amazing biochemical odyssey from substrate to product. *Cell Mol Life Sci.* 56:507-522.
87. Yefimenko I, Fresquet V, Marco-Marín C, Rubio V, Cervera J. 2005. Understanding carbamoyl phosphate synthetase deficiency: impact of clinical mutations on enzyme functionality. *J Mol Biol.* 349:127-141.
88. Thoden JB, Wesenberg G, Raushel FM, Holden HM. 1999. Carbamoyl phosphate synthetase: closure of the B-domain as a result of nucleotide binding. *Biochemistry* 38:2347-2357.
89. Cervera J, Bendala E, Britton HG, Bueso J, Nassif Z, Lusty CJ, Rubio V. 1996. Photoaffinity labeling with UMP of lysine 992 of carbamoyl phosphate synthetase from *Escherichia coli* allows identification of the binding site for the pyrimidine inhibitor. *Biochemistry* 35:7247-7255.
90. Bueso J, Cervera J, Fresquet V, Marina A, Lusty CJ, Rubio V. 1999. Photoaffinity labeling with the activator IMP and site-directed mutagenesis of histidine 995 of carbamoyl phosphate synthetase from *Escherichia coli* demonstrate that the binding site for IMP overlaps with that for the inhibitor UMP. *Biochemistry* 38:3910-3917.
91. Thoden JB, Raushel FM, Wesenberg G, Holden HM. 1999. The binding of inosine monophosphate to *Escherichia coli* carbamoyl phosphate synthetase. *J Biol Chem.* 274: 22502-22507.
92. Thoden JB, Huang X, Kim J, Raushel FM, Holden HM. 2004. Long-range allosteric transitions in carbamoyl phosphate synthetase. *Protein Sci.* 13:2398-2405.
93. Xie Y, Ihsanawati K, Kishishita S, Murayama K, Takemoto C, Shirozu M. 2007. RIKEN Structural Genomics/Proteomics Initiative. Crystal

REFERENCES (Introduction and Discussion)

- structure of MGS domain of carbamoyl-phosphate synthetase from *Homo sapiens*. Protein DataBank file 2YVQ, <http://www.rcsb.org/pdb>.
94. Murzin AG. 1999. Structure classification-based assessment of CASP3 predictions for the fold recognition targets. *Proteins. Suppl* 3:88-103.
 95. Britton HG, Garcia-España A, Goya P, Rozas I, Rubio V. 1990. A structure-reactivity study of the binding of acetylglutamate to carbamoyl phosphate synthetase I. *Eur J Biochem.* 188:47-53.
 96. Simmer JP, Kelly RE, Rinker AG Jr, Scully JL, Evans DR. 1990. Mammalian carbamyl phosphate synthetase (CPS). DNA sequence and evolution of the CPS domain of the Syrian hamster multifunctional protein CAD. *J Biol Chem.* 265:10395-10402.
 97. Martínez AI, 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.
 98. Stojanovski D, Johnston AJ, Streimann I, Hoogenraad NJ, Ryan MT. 2003. Import of nuclear-encoded proteins into mitochondria. *Exp Physiol.* 88:57-64.
 99. Raymond Y, Shore GC. 1981. Processing of the precursor for the mitochondrial enzyme, carbamyl phosphate synthetase. Inhibition by rho-aminobenzamidine leads to very rapid degradation (clearing) of the precursor. *J Biol Chem.* 256:2087-2090.
 100. Mori M, Miura S, Morita T, Takiguchi M, Tatibana M. 1982. Synthesis, intracellular transport and processing of mitochondrial urea cycle enzymes. *Adv Enzyme Regul.* 21:121-132.
 101. Freeman JM, Nicholson JF, Masland WS, Rowland LP, Carter S, Curnen EC. 1964. Ammonia intoxication due to a defect in urea synthesis. *J Pediat.* 65:1039-1040.
 102. Gelehrter TD, Snodgrass PJ. 1974. Lethal neonatal deficiency of carbamyl phosphate synthetase. *N Engl J Med.* 290:430-433.
 103. Shi VE. 1976. Hereditary urea cycle disorders. In Grisolia S, Báguena R, Mayor F, editors. *The Urea Cycle*. John Wiley and Sons, New York, pp. 367-414.
 104. Hoogenraad NJ, Mitchell JD, Don NA, Sutherland TM, Mc Leay AC. 1980. Detection of carbamyl phosphate synthetase 1 deficiency using duodenal biopsy samples. *Arch Dis Child.* 55:292-295.
 105. Brusilow SW, Horwich AL. 2001. Urea cycle enzymes. In Scriver CR, Beaudet AL, Sly WS, Valle D, editors; Child B, Kinzler KW, Vogelstein B, associated editors. *The Metabolic and Molecular Bases of Inherited Disease*, 8e. New York: McGraw-Hill. Vol 2, pp 1909-1963.
 106. McReynolds JW, Crowley B, Mahoney MJ, Rosenberg LE. 1981. Autosomal recessive inheritance of human mitochondrial carbamyl phosphate synthetase deficiency. *Am J Hum Genet.* 33:345-353.

REFERENCES (Introduction and Discussion)

107. Smith I. 1981. The treatment of inborn errors of the urea cycle. *Nature* 291:378-380.
108. Anonymous. 2014. Glycerol phenylbutyrate (Ravicti) for urea cycle disorders. *Med Lett Drugs Ther.* 56:77-78.
109. Bachmann C, Krähenbühl S, Colombo JP, Schubiger G, Jaggi KH, Tönz O. 1981. N-acetylglutamate synthetase deficiency: a disorder of ammonia detoxication. *N Engl J Med.* 304:543.
110. Guffon N, Schiff M, Cheillan D, Wermuth B, Häberle J, Vianey-Saban C. 2005. Neonatal hyperammonemia: the N-carbamoyl-L-glutamic acid test. *J Pediatr.* 147:260-262.
111. 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.
112. Piceni Sereni L, Bachmann C, Pfister U, Buscaglia M, Nicolini U. 1988. Prenatal diagnosis of carbamoyl-phosphate synthetase deficiency by fetal liver biopsy. *Prenat Diagn.* 8:307-309.
113. Neill MA, Aschner J, Barr F, Summar ML. 2009. Quantitative RT-PCR comparison of the urea and nitric oxide cycle gene transcripts in adult human tissues. *Mol Genet Metab.* 97:121-127.
114. Summar ML. 1998. Molecular genetic research into carbamoyl-phosphate synthase I: molecular defects and linkage markers. *J Inherit Metab Dis.* 21 Suppl 1:30-39.
115. Rapp B, Häberle J, Linnebank M, Wermuth B, Marquardt T, Harms E, Koch HG. 2001. Genetic analysis of carbamoylphosphate synthetase I and ornithine transcarbamylase deficiency using fibroblasts. *Eur J Pediatr.* 160:283-287.
116. Kretz R, Hu L, Wettstein V, Leiteritz D, Häberle J. 2012. Phytohemagglutinin stimulation of lymphocytes improves mutation analysis of carbamoylphosphate synthetase 1. *Mol Genet Metab.* 106:375-378.
117. Summar ML, Hall LD, Eeds AM, Hutcheson HB, Kuo AN, Willis AS, Rubio V, Arvin MK, Schofield JP, Dawson EP. 2003. Characterization of genomic structure and polymorphisms in the human carbamyl phosphate synthetase I gene. *Gene* 311:51-57.
118. Häberle J, Schmidt E, Pauli S, Rapp B, Christensen E, Wermuth B, Koch HG. 2003. Gene structure of human carbamylphosphate synthetase 1 and novel mutations in patients with neonatal onset. *Hum Mutat.* 21:444.
119. Funghini S, Donati MA, Pasquini E, Zammarchi E, Morrone A. 2003. Structural organization of the human carbamyl phosphate synthetase I

REFERENCES (Introduction and Discussion)

- gene (CPS1) and identification of two novel genetic lesions. *Hum Mutat.* 22:340-341.
120. 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.
121. Summar ML, Dobbelaere D, Brusilow S, Lee B. 2008. Diagnosis, symptoms, frequency and mortality of 260 patients with urea cycle disorders from a 21-year, multicentre study of acute hyperammonaemic episodes. *Acta Paediatr.* 97:1420-1425
122. Tuchman M, Lee B, Lichter-Konecki U, Summar ML, Yudkoff M, Cederbaum SD, Kerr DS, Diaz GA, Seashore MR, Lee HS, McCarter RJ, Krischer JP, Batshaw ML; Urea Cycle Disorders Consortium of the Rare Diseases Clinical Research Network. 2008. Cross-sectional multicenter study of patients with urea cycle disorders in the United States. *Mol Genet Metab.* 94:397-402.
123. Funghini S, Thusberg J, Spada M, Gasperini S, Parini R, Ventura L, Meli C, De Cosmo L, Sibilio M, Mooney SD, Guerrini R, Donati MA, Morrone A. 2012. Carbamoyl phosphate synthetase 1 deficiency in Italy: clinical and genetic findings in a heterogeneous cohort. *Gene* 493:228-234.
124. Häberle J. 2013. Clinical and biochemical aspects of primary and secondary hyperammonemic disorders. *Arch Biochem Biophys.* 536:101-118.
125. van Karnebeek CD, Sly WS, Ross CJ, Salvarinova R, Yaplito-Lee J, Santra S, Shyr C, Horvath GA, Eydoux P, Lehman AM, Bernard V, Newlove T, Ukpeh H, Chakrapani A, Preece MA, Ball S, Pitt J, Vallance HD, Coulter-Mackie M, Nguyen H, Zhang LH, Bhavsar AP, Sinclair G, Waheed A, Wasserman WW, Stockler-Ipsiroglu S. 2014. Mitochondrial carbonic anhydrase VA deficiency resulting from CA5A alterations presents with hyperammonemia in early childhood. *Am J Hum Genet.* 94:453-461.
126. Coude FX, Sweetman L, Nyhan WL. 1979. Inhibition by propionyl-coenzyme A of N-acetylglutamate synthetase in rat liver mitochondria. A possible explanation for hyperammonemia in propionic and methylmalonic acidemia. *J Clin Invest.* 64:1544-1551.
127. Häberle J. 2011. Clinical practice: the management of hyperammonemia. *Eur J Pediatr.* 170:21-34.
128. Horiuchi M, Kobayashi K, Tomomura M, Kuwajima M, Imamura Y, Koizumi T, Nikaido H, Hayakawa J, Saheki T. 1992. Carnitine administration to juvenile visceral steatosis mice corrects the suppressed expression of urea cycle enzymes by normalizing their transcription. *J Biol Chem.* 267:5032-5035

REFERENCES (Introduction and Discussion)

129. Nakagawa T, Lomb DJ, Haigis MC, Guarente L. 2009. SIRT 5 deacetylates carbamoyl phosphate synthetase 1 and regulates the urea cycle. *Cell* 137:560-570.
130. Ogura M, Nakamura K, Tanaka D, Zhuang X, Fujita Y, Obara A, Hamasaki M, Hosokawa N, Inagaki N. 2010. Overexpression of SIRT5 confirms the involvement in deacetylation and activation of carbamoyl phosphate synthetase I. *Biochem Biophys Res Commun.* 393:73-78.
131. Corvi MM, Soltys CL, Berthiaume LG. 2001. Regulation of mitochondrial carbamoyl-phosphate synthetase 1 activity by active site fatty acylation. *J Biol Chem.* 276:45704-45712.
132. Nakagawa T, Guarente L. 2009. Urea cycle regulation by mitochondrial sirtuin, SIRT5. *Aging (Albany NY)* 1:578-581.
133. Tan M, Peng C, Anderson KA, Chhoy P, Xie Z, Dai L, Park J, Chen Y, Huang H, Zhang Y, Ro J, Wagner GR, Green MF, Madsen AS, Schmiesing J, Peterson BS, Xu G, Ilkayeva OR, Muehlbauer MJ, Braulke T, Mühlhausen C, Backos DS, Olsen CA, McGuire PJ, Pletcher SD, Lombard DB, Hirschey MD, Zhao Y. 2014. Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. *Cell Metab.* 19:605-617.
134. Summar ML, Scott N, Cummings E, Hutcheson H, Dawling S, Christman B. 1999. Analysis of 200 patients undergoing bone marrow transplant shows allelic disequilibrium between drug related toxicity and a common exonic polymorphism in the CPSI gene and correlates with disruption of urea cycle intermediates. *Am J Hum Genet.* 65:Suppl:A25.
135. Pearson DL, Dawling S, Walsh WF, Haines JL, Christman BW, Bazyk A, Scott N, Summar ML. 2001. Neonatal pulmonary hypertension-urea-cycle intermediates, nitric oxide production, and carbamoyl-phosphate synthetase function. *N Engl J Med.* 344:1832-1838.
136. Canter JA, Summar ML, Smith HB, Rice GD, Hall LD, Ritchie MD, Motsinger AA, Christian KG, Drinkwater DC Jr, Scholl FG, Dyer KL, Kavanaugh-McHugh AL, Barr FE. 2007. Genetic variation in the mitochondrial enzyme carbamyl-phosphate synthetase I predisposes children to increased pulmonary artery pressure following surgical repair of congenital heart defects: a validated genetic association study. *Mitochondrion* 7:204-210.
137. Barr FE, Tirona RG, Taylor MB, Rice G, Arnold J, Cunningham G, Smith HA, Campbell A, Canter JA, Christian KG, Drinkwater DC, Scholl F, Kavanaugh-McHugh A, Summar ML. 2007. Pharmacokinetics and safety of intravenously administered citrulline in children undergoing congenital heart surgery: potential therapy for postoperative pulmonary hypertension. *J Thorac Cardiovasc Surg.* 134:319-326.

REFERENCES (Introduction and Discussion)

138. Fike CD, Summar M, Aschner JL. 2014. L-Citrulline provides a novel strategy for treating chronic pulmonary hypertension in newborn infants. *Acta Paediatr.* 103:1019-1026.
139. Summar ML, Hall L, Christman B, Barr F, Smith H, Kallianpur A, Brown N, Yadav M, Willis A, Eeds A, Cermak E, Summar S, Wilson A, Arvin M, Putnam A, Wills M, Cunningham G. 2004. Environmentally determined genetic expression: clinical correlates with molecular variants of carbamyl phosphate synthetase I. *Mol Genet Metab.* 81 Suppl 1: S12-19.
140. Ahuja V, Powers-Lee SG. 2008. Human carbamoyl-phosphate synthetase: insight into N-acetylglutamate interaction and the functional effects of a common single nucleotide polymorphism. *J Inherit Metab Dis.* 31:481-491.
141. Yagi M, Nakamura T, Okizuka Y, Oyazato Y, Kawasaki Y, Tsuneishi S, Sakaeda T, Matsuo M, Okumura K, Okamura N. 2010. Effect of CPS14217C>A genotype on valproic-acid-induced hyperammonemia. *Pediatr Int.* 52:744-748.
142. Janicki PK, Bezinover D, Postula M, Thompson RS, Acharya J, Acharya V, McNew C, Bowman JD, Kurkowska-Jastrzebska I, Mirowska-Guzel D. 2013. Increased occurrence of valproic acid-induced hyperammonemia in carriers of T1405N polymorphism in carbamoyl phosphate synthetase 1 gene. *ISRN Neurol.* 2013:261497.
143. Inoue K, Suzuki E, Takahashi T, Yamamoto Y, Yazawa R, Takahashi Y, Imai K, Miyakawa K, Inoue Y, Tsuji D, Hayashi H, Itoh K. 2014. 4217C>A polymorphism in carbamoyl-phosphate synthase 1 gene may not associate with hyperammonemia development during valproic acid-based therapy. *Epilepsy Res.* 108:1046-1051.
144. Moonen RM, Paulussen AD, Souren NY, Kessels AG, Rubio-Gozalbo ME, Villamor E. 2007. Carbamoyl phosphate synthetase polymorphisms as a risk factor for necrotizing enterocolitis. *Pediatr Res.* 62:188-190.
145. Moonen RM, Reyes I, Cavallaro G, González-Luis G, Bakker JA, Villamor E. 2010. The T1405N carbamoyl phosphate synthetase polymorphism does not affect plasma arginine concentrations in preterm infants. *PLoS One.* 5:e10792.
146. Paré G, Chasman DI, Parker AN, Zee RR, Mälarstig A, Seedorf U, Collins R, Watkins H, Hamsten A, Miletich JP, Ridker PM. 2009. Novel associations of CPS1, MUT, NOX4, and DPEP1 with plasma homocysteine in a healthy population: a genome-wide evaluation of 13 974 participants in the Women's Genome Health Study. *Circ Cardiovasc Genet.* 2:142-150.
147. Lange LA, Croteau-Chonka DC, Marvelle AF, Qin L, Gaulton KJ, Kuzawa CW, McDade TW, Wang Y, Li Y, Levy S, Borja JB, Lange EM,

REFERENCES (Introduction and Discussion)

- Adair LS, Mohlke KL. 2010. Genome-wide association study of homocysteine levels in Filipinos provides evidence for CPS1 in women and a stronger MTHFR effect in young adults. *Hum Mol Genet.* 19:2050-2058.
148. Kleber ME et al. 2013. Genome-wide association study identifies 3 genomic loci significantly associated with serum levels of homoarginine: the AtheroRemo Consortium. *Circ Cardiovasc Genet.* 6:505-513.
149. Danik JS, Paré G, Chasman DI, Zee RY, Kwiatkowski DJ, Parker A, Miletich JP, Ridker PM. 2009. Novel loci, including those related to Crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17686 women: the Women's Genome Health Study. *Circ Cardiovasc Genet.* 2:134-141.
150. Köttgen A, Pattaro C, Böger CA, Fuchsberger C, Olden M, Glazer NL, et al. and Fox CS. 2010. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 42:376–384.
151. Moore PE, Ryckman KK, Williams SM, Patel N, Summar ML, Sheller JR. 2009. Genetic variants of GSNOR and ADRB2 influence response to albuterol in African-American children with severe asthma. *Pediatr Pulmonol.* 44:649-654.
152. Miran SG, Chang SH, Raushel FM. 1991. Role of the four conserved histidine residues in the amidotransferase domain of carbamoyl phosphate synthetase. *Biochemistry* 30:7901-7907.
153. Guillou F, Liao M, Garcia-España A, Lusty CJ. 1992. Mutational analysis of carbamyl phosphate synthetase. Substitution of Glu841 leads to loss of functional coupling between the two catalytic domains of the synthetase subunit. *Biochemistry* 31:1656-1664.
154. Kothe M, Powers-Lee SG. 2004. Nucleotide recognition in the ATP-grasp protein carbamoyl phosphate synthetase. *Protein Sci.* 13:466-475.
155. Aoshima T, Kajita M, Sekido Y, Kikuchi S, Yasuda I, Saheki T, Watanabe K, Shimokata K, Niwa T. 2001. Novel mutations (H337R and 238-362del) in the CPS1 gene cause carbamoyl phosphate synthetase I deficiency. *Hum Hered.* 52:99-101.
156. Eroglu B, Powers-Lee SG. 2002. Mutational analysis of ATP-grasp residues in the two ATP sites of *Saccharomyces cerevisiae* carbamoyl phosphate synthetase. *Arch Biochem Biophys.* 407:1-9.
157. Eroglu B, Powers-Lee SG. 2002. Unmasking a functional allosteric domain in an allosterically nonresponsive carbamoyl-phosphate synthetase. *J Biol Chem.* 277:45466-45472.
158. Guy HI, Evans DR. 1994. Cloning and expression of the mammalian multifunctional protein CAD in *Escherichia coli*. Characterization of the recombinant protein and a deletion mutant lacking the major interdomain linker. *J Biol Chem.* 269:23808-23816.

REFERENCES (Introduction and Discussion)

159. Guy HI, Evans DR. 1995. Substructure of the amidotransferase domain of mammalian carbamyl phosphate synthetase. *J Biol Chem.* 270:2190-2197.
160. Hewagama A, Guy HI, Vickrey JF, Evans DR. 1999. Functional linkage between the glutaminase and synthetase domains of carbamoyl-phosphate synthetase. Role of serine 44 in carbamoyl-phosphate synthetase-aspartate carbamoyl transferase-dihydroorotase (CAD). *J Biol Chem.* 274:28240-28245.
161. Niu B. 1993. Expression of rat CPSI full-length cDNA gene in non-liver culture cells. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao.* 15:313-319.
162. Park H, Kim IH, Kim IY, Kim KH, Kim HJ. 2000. Expression of carbamoyl phosphate synthetase I and ornithine transcarbamoylase genes in Chinese hamster ovary *dhfr*-cells decreases accumulation of ammonium ion in culture media. *J Biotechnol.* 81:129-140.
163. Saeed-Kothe A, Powers-Lee SG. 2003. Gain of glutaminase function in mutants of the ammonia-specific frog carbamoyl phosphate synthetase. *J Biol Chem.* 278:26722-26726.
164. Hart EJ, Powers-Lee SG. 2009. Role of Cys-1327 and Cys-1337 in redox sensitivity and allosteric monitoring in human carbamoyl phosphate synthetase. *J Biol Chem.* 284:5977-5985.
165. Eeds AM, Mortlock D, Wade-Martins R, Summar ML. 2007. Assessing the functional characteristics of synonymous and nonsynonymous mutation candidates by use of large DNA constructs. *Am J Hum Genet.* 80:740-750.
166. Eeds AM, Hall LD, Yadav M, Willis A, Summar S, Putnam A, Barr F, Summar ML. 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.
167. Blissard, GW, Rohrmann GF. 1990. Baculovirus diversity and molecular biology. *Annu Rev Entomol.* 35:127-155
168. Hofmann, C, Sandig, V, Jennings, G, Rudolph, M, Schlag, P, Strauss, M. 1995. Efficient gene transfer into human hepatocytes by baculovirus vectors. *Proc Natl Acad Sci USA.* 92: 10099-10103.
169. Rohrmann, GF. 2013. Baculovirus molecular biology: Third Edition [Internet]. Bethesda (MD): National Center for Biotechnology Information (US). <http://www.ncbi.nlm.nih.gov/books/NBK114593/>
170. van Oers MM, Pijlman GP, Vlak JM. 2015. Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. *J Gen Virol.* 96:6-23.

REFERENCES (Introduction and Discussion)

171. Hill-Perkins MS, Possee RD. 1990. A baculovirus expression vector derived from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus. *Gen Virol.* 71:971-976.
172. Han BD, Livingstone LR, Pasek DA, Yablonski MJ, Jones ME. 1995. Human uridine monophosphate synthase: baculovirus expression, immunoaffinity column purification and characterization of the acetylated amino terminus. *Biochemistry.* 34:10835-10843.
173. Luckow VA, Lee SC, Barry GF, Olins PO. 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*. *J Virol.* 67:4566-4579.
174. http://kirschner.med.harvard.edu/files/protocols/Invitrogen_bactobacexpression.pdf
175. Craig NL. 1991. Tn7: a target site-specific transposon. *Mol Microbiol.* 5:2569-2573.
176. Li JA, Happ B, Schetter C, Oellig C, Hauser C, Kuroda K, Knebel-Mörsdorf D, Klenk HD, Doerfler W. 1990. The expression of the *Autographa californica* nuclear polyhedrosis virus genome in insect cells. *Vet Microbiol.* 23:73-78.
177. Ciccarone, VC, Polayes, D, Luckow, VA. 1997. Generation of recombinant baculovirus DNA in *E. coli* using baculovirus shuttle vector. In U. Reischt (ed.). *Molecular Diagnosis of Infectious Diseases.* Humana Press Inc., Totowa, NJ. Vol. 13, pp 213-235.
178. Barry GF. 1988. A broad-host-range shuttle system for gene insertion into the chromosomes of gram-negative bacteria. *Gene* 71:75-84.
179. Beji A, Izard D, Gavini F, Leclerc H, Leseine-Delstanche M, Krembel J. 1987. A rapid chemical procedure for isolation and purification of chromosomal DNA from gram-negative bacilli. *Anal Biochem.* 162:18-23.
180. Hopkins R, Esposito D. 2009. A rapid method for titrating baculovirus stocks using the Sf-9 Easy Titer cell line. *Biotechniques* 47:785-788.
181. Brown M, Faulkner P. 1978. Plaque assay of nuclear polyhedrosis viruses in cell culture. *Appl Environ Microbiol.* 36:31-35.
182. Fuchs LY, Woods MS, Weaver RF. 1983. Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *J Virol.* 48:641-646.
183. Kitamura S, Imai T, Grace TD. 1973. Adaptation of two mosquito cell lines to medium free of calf serum. *J Med Entomol.* 10:488-489.
184. [http://tools.lifetechnologies.com/content/sfs/manuals/CellfectinII Reagent_protocol.pdf](http://tools.lifetechnologies.com/content/sfs/manuals/CellfectinII_Reagent_protocol.pdf).
185. <http://tools.lifetechnologies.com/content/sfs/manuals/3408SF900II.pdf>.

REFERENCES (Introduction and Discussion)

186. http://viralzone.expasy.org/all_by_protein/537.html
187. Volkman LE, Goldsmith PA. 1983. *In vitro* survey of *Autographa californica* nuclear polyhedrosis virus interaction with nontarget vertebrate host cells. *Appl Environ Microbiol.* 45:1085-1093.
188. Peracchi A. 2001. Enzyme catalysis: removing chemically 'essential' residues by site-directed mutagenesis. *Trends Biochem Sci.* 26:497-503.
189. Hutchison CA 3rd, Phillips S, Edgell MH, Gillam S, Jahnke P, Smith M. 1978. Mutagenesis at a specific position in a DNA sequence. *J Biol Chem.* 253:6551-6560.
190. Carter P. 1986. Site-directed mutagenesis. *Biochem J.* 237:1-7.
191. Taylor JW, Schmidt W, Cosstick R, Okruszek A, Eckstein F. 1985. The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA. *Nucleic Acids Res.* 13:8749-8764.
192. Wells JA, Vasser M, Powers DB. 1985. Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene* 34:315-323.
193. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. 1986. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 51:263-273.
194. Shimada A. 1996. PCR-based site-directed mutagenesis. *Methods Mol Biol.* 57:157-165.
195. Eckert KA, Kunkel TA. 1991. DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Appl.* 1:17-24.
196. Terpe K. 2013. Overview of thermostable DNA polymerases for classical PCR applications: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol.* 97:10243-10254.
197. <http://www.chem.agilent.com/Library/usermanuals/Public/200521.pdf>
198. Mierzejewska K, Siwek W, Czapinska H, Kaus-Drobek M, Radlinska M, Skowronek K, Bujnicki JM, Dadlez M, Bochtler M. 2014. Structural basis of the methylation specificity of *R.DpnI*. *Nucleic Acids Res.* 42:8745-8754.
199. Lundberg KS, Shoemaker DD, Adams MW, Short JM, Sorge JA, Mathur EJ. 1991. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* 108:1-6.
200. Hogrefe HH, Hansen CJ, Scott BR, Nielson KB. 2002. Archaeal dUTPase enhances PCR amplifications with archaeal DNA polymerases by preventing dUTP incorporation. *Proc Natl Acad Sci USA.* 99:596-601.
201. Cheng S, Fockler C, Barnes WM, Higuchi R. 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc Natl Acad Sci USA.* 91:5695-5699.

REFERENCES (Introduction and Discussion)

202. Spector, L, Jones, ME, Lipman, F. 1957. Carbamyl Phosphate. In Methods in Enzymology, Colowick SP, Kaplan NO, editors. Academic Press, New York and London, Vol. 111, pp: 653-655.
203. Allen, CM Jr, Jones, ME. 1964. Decomposition of carbamylphosphate in aqueous solutions. *Biochemistry* 3: 1238-1247.
204. Barcelona-Andrés B, Marina A, Rubio V. 2002. Gene structure, organization, expression, and potential regulatory mechanisms of arginine catabolism in *Enterococcus faecalis*. *J Bacteriol.* 184:6289-6300.
205. Nakamura, M, Jones, ME. 1970. Ornithine carbamyltransferase (*Streptococcus faecalis*). In Methods in Enzymology, Tabor H, Tabor CW, eds. Academic Press, New York and London, Vol. 17, Part A, pp 286-294.
206. Archibald RM. 1944. Determination of citrulline and allantoin and demonstration of citrulline in blood plasma. *J Biol Chem.* 156:121-142.
207. Bergmeyer, HU. 1983. NAD(P)H dependent reactions. In Methods of Enzymatic Analysis Third ed., Bergmeyer, J, Grassl, M, eds. Verlag Chemie GmbH, Weinheim. Vol I, pp 357-359.
208. Bergmeyer, HU. 1983. Adenosine 5'-Triphosphate. In Methods of Enzymatic Analysis Third ed., Bergmeyer, J, Grassl, M, eds. Verlag Chemie GmbH, Weinheim. Vol VII, pp 340-364.
209. Klaus V, Vermeulen T, Minassian B, Israelian N, Engel K, Lund AM, Roebrock K, Christensen E, Häberle J. 2009. Highly variable clinical phenotype of carbamylphosphate synthetase 1 deficiency in one family: an effect of allelic variation in gene expression? *Clin Genet.* 76:263-269.
210. Post LE, Post DJ, Raushel FM. 1990. Dissection of the functional domains of *Escherichia coli* carbamoyl phosphate synthetase by site-directed mutagenesis. *J Biol Chem.* 265:7742-7747.
211. Miles BW, Mareya SM, Post LE, Post DJ, Chang SH, Raushel FM. 1993. Differential roles for three conserved histidine residues within the large subunit of carbamoyl phosphate synthetase. *Biochemistry* 32:232-240.
212. Stapleton MA, Javid-Majd F, Harmon MF, Hanks BA, Grahmann JL, Mullins LS, Raushel FM. 1996. Role of conserved residues within the carboxy phosphate domain of carbamoyl phosphate synthetase. *Biochemistry* 35:14352-14361.
213. Javid-Majd F, Stapleton MA, Harmon MF, Hanks BA, Mullins LS, Raushel FM. 1996. Comparison of the functional differences for the homologous residues within the carboxy phosphate and carbamate domains of carbamoyl phosphate synthetase. *Biochemistry* 35:14362-14369.

REFERENCES (Introduction and Discussion)

214. Blanchard CZ, Lee YM, Frantom PA, Waldrop GL. 1999. Mutations at four active site residues of biotin carboxylase abolish substrate-induced synergism by biotin. *Biochemistry*. 38:3393-3400.
215. Levert KL, Lloyd RB, Waldrop GL. 2000. Do cysteine 230 and lysine 238 of biotin carboxylase play a role in the activation of biotin? *Biochemistry*. 39:4122-4128.
216. Marshall M, Fahien LA. 1985. Proximate sulfhydryl groups in the acetylglutamate complex of rat carbamylphosphate synthetase I: their reaction with the affinity reagent 5'-p-fluorosulfonylbenzoyladenine. *Arch Biochem Biophys*. 241:200-214.
217. Geschwill K, Lumper L. 1989. Identification of cysteine residues in carbamoyl-phosphate synthase I with reactivity enhanced by N-acetyl-L-glutamate. *Biochem J*. 260:573-576.
218. Díaz-Moreno I, García-Heredia JM, González-Arzola K, Díaz-Quintana A, De la Rosa MÁ. 2013. Recent methodological advances in the analysis of protein tyrosine nitration. *Chemphyschem*. 14:3095-3102.