U.PORTO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

GENETIC, BIOCHEMICAL AND BIOPHYSICAL STUDIES OF THE HEME BIOSYNTHETIC PATHWAY

LUÍS FILIPE LOPES DA CUNHA

Dissertação de doutoramento em Ciências Biomédicas

2008

LUÍS FILIPE LOPES DA CUNHA

GENETIC, BIOCHEMICAL AND BIOPHYSICAL STUDIES OF THE HEME BIOSYNTHETIC PATHWAY

Dissertação de Candidatura ao grau de Doutor
em Ciências Biomédicas, submetida ao Instituto
De Ciências Biomédicas de Abel Salazar

ACKNOWLEDGEMENTS:

I would like to thank Dr. Robert Desnick and Dr. Graça Porto, for being my mentors. I would also like to thank Dr. David Bishop for his daily supervision in the lab, valuable discussions, and help with publishing. Dr. Miklos Kuti for teaching me NMR, and for reviewing this manuscript, Dr. Mihaly Mezei for teaching me computational biology; Dr. Kenneth Astrin for his cheerful support in the lab; Alicia and Peter Warburton for their friendship since the first day I arrived in New York City; I wish also to thank those who organized the GABBA program 1997, for giving me the opportunity to participate in the University of Porto's graduate program. To my family, specially my parents, Abel Dias Da Cunha and Maria Almerinda Da Costa Lopes, all of those who supported me during this endeavor, and to my wife, Jane Marie Hitch Cunha, for her support during the final leg of this journey.

CONTENTS: Page

FIGURE INDEX

TABLE INDEX

APPENDIX INDEX

ABBREVIATION LIST

δ-aminolevulinic acid synthase: ALAS; aminolevulinic acid: ALA; δ-aminolevulinic acid dehydratase ALAD; hydroxymethylbilane HMB; uroporphyrinogen decarboxylase: URO-decarboxylase; hydroxymethylbilane synthase: HMB-synthase; uroporphyrinogen III synthase: URO-synthase; coproporphyrinogen III: COPRO'gen III; ferrochelatase: FECH; Acute Intermittent Porphyria: AIP; Congenital Erythropoietic Porphyria: CEP; coproporphyrinogen I: COPRO'gen I; Familial Porphyria Cutanea Tarda: fPCT; Porphyria Cutanea Tarda: PCT; hepatoerythropoietic porphyria: HEP; porphobilinogen: PBG; nuclear magnetic resonance: NMR; DNA – deoxyribonucleic acid; RNA – ribonucleic acid; mRNA: messenger ribonucleic acid; PCR – polymerase chain reaction; Restriction Fragment Length Polymorphism: RFLP; RT-PCR: reverse transcriptase PCR; bp: base pair; nonsense mediated mRNA decay: NMD; single nucleotide polymorphisms: SNPs; uroporphyrinogen I: URO'gen I; uroporphyrin I: URO I; coproporphyrin I: COPRO I; uroporphyrinogen III: URO'gen III; matrix-assisted laser desorption ionization/time-of-flight: MALDI-TOF; isopropanol: IPA; acid/water/IPA: FWI; Trifluoroacetic acid: TFA; Unit: U;

erythroid HMB-synthase or E-HMBS: E-HMB-synthase; housekeeping HMB-synthase or H-HMBS: H-HMB-synthase; Small Ubiquitin-related Modifier: SUMO; immobilized metal affinity chromatography: IMAC; gel filtration: GF; Da – Dalton; SDS- sodium dodecyl sulphate; TCA: trichloroacetic acid; TCEP: tris (2-carboxyethyl) phosphine; DTT: dithiothreitol; Ulp1: ubiquitin-like protein; IPTG: isopropyl-beta-D-thiogalactopyranoside; SDS-PAGE: sodium dodecyl sulphate-polyacrylamide gel; ubiquitin-like: UBL; p-aminobenzoate - PABA; 3D – tri-dimensional; Nuclear Overhauser Effect: NOE; TOCSY: Total correlated spectroscopy; NMF-bilane: N_D-methyl-1-formylbilane; tris (2-carboxyethyl) phosphine; PMSF, phenylmethylsulphonyl fluoride; HSQC: heteronuclear single quantum correlation; TOCSY, total correlated spectroscopy; PDB: protein databank; molecular dynamics: MD; steered molecular dynamics: SMD; Metropolis Monte Carlo: MMC; K: Kelvin; Biological Magnetic Resonance Data Bank: BMRB; root mean square deviation: r.m.s.d.; kT/e: k = the Boltzman constant, T = temperature and e = electron volt; Ansgtrom: Å; inhibition constant - K_i; lognormal – ln; maximum velocity: V_{max}; Michaelis-Menten constant: K_M; normal mode analyses: NMA; Gibbs free energy: ΔG; Alanine: Ala or A; Arginine: Arg or R; Asparagine: Asn or N; Aspartic acid: Asp or D; Cysteine: Cys or C; Glutamine: Gln or Q; Glutamic acid: Glu or E; Glycine: Gly or G; Histidine: His or H; Isoleucine; Ile or I; Leucine: leu or L; Lysine: Lys or K; Methionine: Met or M; Phenylalanine: Phe or F; Prloline: Pro or P; Serine: Ser or S; Threonine: Thr or T; Tryptophan: Trp or W; Tyrosine: Tyr or Y; Valine: Val or V.

SUMMARY

The term “Porphyria” refers to a heterogeneous group of diseases, each due to the deficient activity of a specific heme biosynthetic enzyme, characterized by the accumulation in the body of porphyrins, or porphyrin precursors. There are several possible classifications of the porphyrias. The main place of accumulation (erythron or liver) distinguishes between erythropoietic (congenital erythropoietic porphyria and erythropoietic protoporphyria) and hepatic porphyrias (ALAD deficient porphyria, AIP, hereditary coproporphyria and variegate porphyria). The acute porphyrias are characterized by neurological manifestations, and cutaneous porphyrias have skin manifestations.

In the first study, the molecular basis of porphyria cutanea tarda (PCT) in two Portuguese families was determined. It is well known that the biochemical diagnosis of the porphyrias by enzyme or porphyrin analysis has been problematic, and that the definitive diagnosis of a particular porphyria requires the demonstration of the responsible genetic mutation. In an early component of this work, we analyzed two, presumably unrelated, Portuguese families diagnosed with PCT, and identified E314E as the disease-causing mutation in the *UROD* gene of both families. This mutation was also present in several other family members. This was a splicing mutation that caused the deletion of exon 9. The resultant mutant *UROD* mRNA encoded a normal polypeptide sequence through exon 8 (291 residues), then the misspliced mRNA was out-of-frame and encoded seven abnormal amino acids from exon 10, followed by a premature stop codon at residue 298. These results permit pre-symptomatic molecular diagnosis and counseling of these families to enable family members to avoid disease-precipitating factors.

The second and third studies focused on the structural and functional relationships of hydroxymethylbilane synthase (HMB-synthase), uroporphyrinogen III synthase (URO-synthase) and uroporphyrinogen decarboxylase (URO-decarboxylase), the third, fourth and fifth enzymes in the heme biosynthetic pathway. For over four decades, investigators have suggested that URO-synthase and HMB-synthase interact in a cytosolic complex in order to efficiently convert porphobilinogen (PBG) to uroporphyrinogen III (URO’gen III), which is required for porphyrin biosynthesis. A HMB-synthase/URO-synthase complex would prevent the non-enzymatic conversion of hydroxymethylbilane (HMB) to the URO’gen I isomer, a non-physiologic and toxic compound. An enzymatic complex remains to be demonstrated; although much progress has been made in understanding the molecular basis of the enzymes involved in the pathway, the reaction mechanism of URO-synthase remains an unsolved mystery. For studies of these two pending questions of heme biosynthesis, which require these enzymes to be available pure and in large

quantities, we describe in chapter two their high-level recombinant expression, and efficient, high-yield purification. The recombinant human enzymes were subcloned in specific prokaryotic expression vectors, optimally selected for each enzyme, and expressed in *E. coli*. The cDNAs of HMB-synthase (housekeeping and erythroid specific enzyme) were fused to a 10-histidine sequence and a Factor Xa recognition site. URO-synthase was fused with the Ubiquitin-related Modifier (SUMO) peptide, Smt3, between an N-terminal His•Tag™ and URO-synthase, and the UROD cDNA was fused in an expression construct to a modified thioredoxin polypeptide systems. The recombinant human enzymes were overexpressed in *E. coli*, purified to homogeneity, and characterized. The respective recombinant enzymes showed similar properties to those reported for the human enzymes purified from erythrocytes or liver. Conditions for long term stability of each enzyme were also determined.

In chapter three, we performed structural studies using NMR and computer simulation techniques. First, we investigated the long-held hypothesis that URO-synthase and HMB-synthase interact in a complex, using NMR spectroscopy. For this purpose, we assigned the backbone amide resonances of URO-synthase. Chemical shift perturbation mapping, in which ^{15}N-labeled URO-synthase was titrated with unlabelled HMB-synthase and/or URO-decarboxylase, showed that the 2D HSQC (heteronuclear single quantum coherence) was not specifically perturbed by either enzyme, therefore not supporting the formation of an enzymatic complex. Since the enzyme concentrations used were at least two orders of magnitude higher than those found in the liver, it is unlikely that these enzymes interact directly *in vivo*. A second pending question in heme biosynthesis relates to the mechanism by which URO-synthase catalyses the formation of uroporphyrinogen III (URO III), the common precursor of all tetrapyrroles, from hydroxymethylbilane (HMB), with inversion of ring D. For decades, much effort has been put into solving this biochemical puzzle. Attempts were made to synthesize intermediate state inhibitors, as well as to trap the biological transition state. Although the later has proven elusive, possibly due to a very short lived species that remains bound to the enzyme, the synthesis of several inhibitors have led to suggest a spiro-pyrrolenine intermediate. In an ongoing effort to demonstrate the mechanism of the enzyme catalyzed conversion of HMB to URO III, we described in this work the mapping of the active site of URO-synthase, the first step in determining a mode of action. For this, we employed both NMR and computational studies to characterize the substrate binding site. NMR experiments involved the titration of isotope-labeled URO-synthase with two inhibitors of substrate binding, and revealed important residues involved in substrate binding. These experiments localized the active site to the large cleft between the enzyme’s two globular domains. This result was congruent with computer docking studies, which provided a model of the enzyme-substrate complex. The Docking studies dock the ligand to the rigid protein, and predict the best conformers, based on the lowest free energy of the complex. Although the x-ray crystal structure of URO-synthase was previously available, the putative active site cleft was too large to tightly fit the substrate molecule. Therefore we solved the 3D solution structure of URO-synthase by NMR. Both structures are very similar, but the active site cleft between domains is smaller in the solution structure. The smaller cleft in the NMR structure better explains the predicted ligand contacts with both domains, judging from the NMR perturbation of residues in both sides of the cleft. Large domain movements between open/closed conformations may be important for the reaction mechanism. Molecular dynamics simulations, a more elegant, but computationally expensive method that simulates the time-dependent behavior of the flexible protein-ligand complex, confirmed the accuracy of the rigid-body docking results. Furthermore, preliminary mutagenesis of specific active-site residues showed reduced activity of some of these mutants, adding further evidence of the active site localization.

In summary, these studies have provided further understanding of the molecular genetics, cellular biochemistry and structural biology of three critical cytosolic enzymes in the heme biosynthetic pathway, and provide the basis for future studies of the URO-synthase reaction mechanism by molecular dynamics and quantum mechanics simulations.

The work relevant to this thesis has been published in the scientific journal *Proteins: Structure, Function and Bioinformatics* (Cunha, L., Kuti, M., Bishop, D. F., Mezei, M., Zeng, L., Zhou, M.-M., Desnick, R. J., 2008, Human Uroporphyrinogen III Synthase: NMR-based Mapping of the Active Site, Proteins: Structure, Function and Bioinformatics, vol 71:855-873).

RESUME

Le terme de porphyrie ne se relie pas à une maladie spécifique, mais implique plutôt un groupe hétérogène de maladies héréditaires, chacun étant lié à l’activité déficiente d’une enzyme spécifique dans la biosynthèse de l'hème, qui entraîne l'accumulation dans le corps des porphyrines et de leurs précurseurs. Il y a plusieurs classifications possibles des porphyries. L’organe principal de l'accumulation (érythron ou foie) permet de distinguer les formes de porphyries érythropoiétiques (CEP, protoporphyries érythropoiétiques) et hépatiques (porphyrie déficiente d'ALA, AIP, le coproporphyrie héréditaire et porphyrie variagata). Les porphyries aiguës sont caractérisées par des manifestations neurologiques, et les porphyries cutanées ont des manifestations dermatologiques.

Dans la première étude, la base moléculaire du porphyrie cutanée tardive (PCT) dans deux familles portugaises a été déterminée. Puisque le diagnostic biochimique a été problématique, un diagnostic définitif d'une porphyrie particulière nécéssite l’identification de la mutation génétique responsable. Dans une première partie de ce travail, nous avons analysé deux, vraisemblablement indépendante, familles portugaises diagnostiquées avec de la PCT. Une mutation E314E a été identifiée comme cause de la maladie, dans le gène *UROD* chez ces deux familles. Cette mutation d'épissage entraîne la perte de l’exon 9. L’ARNm résultant du mutant UROD code une séquence polypeptidique normale de l’exon 8 (291 residus), puis le décallage du cadre de code sept acides aminés anormaux de l’exon 10 suivi d’un codon stop au résidu 298. Ces résultats permettent un diagnostic moléculaire pré-symptomatiques et l’analyse de ces familles, notamment la consultation de ces membres de famille, devrait prévenir les facteurs entraînant cette maladie.

La deuxième et troisième études sont concentrée sur le rapport structural et de functional des trois enzymes biosynthétiques, «*hydroxymethylbilane synthase* (*HMB-synthase*)», «*uroporphyrinogen III synthase (URO-synthase*)», et «*uroporphyrinogene decarboxylase*», les troisième, quatrièmes, et cinquième enzymes respectivement dans la voie biosynthétique du groupe heme. Depuis plus de quarante ans, il y a été suggéré que «*HMB-synthase*» et *«URO-synthase*», impliquées dans la biosynthèse de l’hème, forment un complexe cystosolique actif pour transformer efficacement la «*porphobilinogène (PBG)*», nécessaire à la biosynthèse de la porphyrine, en l’isomère *uroporphyrinogène III (URO’gen III*). Ce complexe *« HMB-synthase/URO-synthase»* empêcherait ainsi la conversion non enzymatique de «l’hydroxyméthylbilane (HMB)» en l’isomère «uroporphyrinogène I (URO’gen I), qui est une molécule toxique physiologiquement.

Afin de caractériser l'interaction moléculaire entre «HMB-synthase», «URO-synthase», et «URO-decarboxylase», et de déterminer la structure tri-dimensionnelle de la protéine «URO-synthase» par résonance magnétique nucléaire (RMN), les ADNc des enzymes humaines ont été clonés dans un vecteur d'expression spécifique des procaryotes. Le gène «HMB-synthase» a été fusionné avec une séquence de 10 histidine et un site de reconnaissance du facteur Xa. Le gène «URO-synthase» a été fusionné avec le peptide Ubiquitin-connexe demodificateur «(SUMO)» - Smt3, avec en N-terminal la séquence His.Tag et en C-terminal le gène «URO-synthase». Le gène UROD fusionné a été exprimé dans un système inducteur grâce au polypeptide modifié de «*thioredoxin*». Les enzymes humaines recombinaées ont été surexprimées dans *E. coli*, et purifiées pour leur caractérisation. Ces enzymes ont montré des propriétés semblables à celles rapportées pour les enzymes humaines purifiées des érythrocytes ou des cellules dufoie. Des conditions pour leur stabilité à long terme ont été également déterminées.

Dans le chapitre trois,nous avons employé RMN et les techniques de simulation informatiques pour étudier deux aspects de structure-fonction non résolus à propos de la biosynthèse de heme: Premier, nous avons examiné l'hypothèse longtemps-tenu que «URO-synthase» et l' «HMB-synthase» réagissent réciproquement dans un complexe, utilisant la spectroscopie de RMN. A cet effet, nous avons assigné les résonances de amide de la chaîne peptidic principale d'URO-synthase. Les techniques RMN dans lequel «^{15}N-URO-synthase» a été titrated avec l' «URO-decarboxylase» de et/ou d' «HMB-synthase», ont montré que le 2D HSQC (heteronuclear de cohérence quantique seule) n'a pas été perturbé en particulier par les deux l'enzymes, ne soutenant pas donc la formation d'un complexe enzymatique. Puisque les concentrations d'enzyme utilisées étaient au moins deux ordres de magnitude plus élevés que ces trouvé dans le foie, c'est peu probable que ces enzymes réagissent réciproquement de maniere direct in vivo.

Dans la troisième étude, nous avons investigué l'interaction putative protéine-protéine entre l' «URO-synthase» et l' «HMB-synthase» et/ou l' «URO-decarboxylase», par des techniques de spectroscopie en RMN. À ce sujet, nous avons assigné les résonances de URO-synthase. Pour la mappage de la perturbation chimique, l' «URO-synthase» a été marqué avec ^{15}N et titrée avec «HMB-synthase» et/ou «URO-decarboxylase» non marqué. Les résultats obtenus montrent que le spectrum 2D HSQC (cohérence monoquantique hétéronucléaire) «URO-synthase» se sont altérés pas en présence de «HMB-synthase», ou «URO-decarboxylase».

Une seconde question dans la biosynthèse du groupe heme est le mécanisme par lequel URO-synthase convertit la formation de «uroporphyrinogen III (URO III)», le précurseur commun de tout tetrapyrroles, apartir de «hydroxymethylbilane (HMB)», avec l'inversion d'anneau D. Pour les décennies, beaucoup d'effort a été mis pour résoudre cette énigme biochimique. Des tentatives ont été faites pour synthétiser des inhibiteurs des états intermédiaires, de même que piéger les étatsde transition biologique. Bien que pas demonstré, probablement en raison de la présence des espèces a court durée de vie liés à l'enzyme, la synthèse de plusieurs inhibiteurs à suggère l'existence de un « spiro-pyrrolenine » intermédiaire. Dans un effort continu pour démontrer le mécanisme stereospecifique d'enzyme pour catalyser la formation de URO III, nous avons employé RMN et les études computationnelles, de manière à faire la carte du site actif et caractérise la liesont de substrat.

L'évaluation expérimentale de résidus perturbés par le titration de deux inhibiteurs de liesont au substrat a révélé ques résidus importants sont impliqués , localisées dans la grande fente entre les deux domaines globuleux de l'enzyme. La localisation de ces résidus était en harmonie avec des modèles informatiques, qui ont prédit la liesont à la même région et ont fourni une visualisation du complexe d'enzyme-substrat. Bien que la structure d'URO-synthase était précédemment disponible, la fente de site putative actif était trop grande pour s'ajuster solidement au substrat. Donc nous avons résolu la structure 3D de par RMN. Les deux structures sont très similaires, mais comme prévu, la fente de site active entre les domaines est plus petite dans la structure de solution. La structure RMN de la plus petite fent explique de plus satisfaisanteles contacts de la molecule ligand avec les domaines et des grands mouvements entre une conformation d'ouverture/fermé pourraient être impliqués dans le mécanisme de réaction. Technique de dynamique moléculaire a été employée pour simuler le comportement dépendant du temps du complexe, et a confirmé les résultats d'amarrage. En outre, la mutagénèse préliminaire des résidus d'actif-emplacement a montré l'activité réduite des résidus principaux, de ce fait confirmant la localisation active d'emplacement.

En résumé, nous études ont permis une meilleur compréhension de la génétique moléculaire, de la biochimie cellulaire et des relations entre structure et fonction des trois enzymes cytosoliques critiques dans la voie de biosynthèse de l'hème, et ils fournissent la base pour l'explication future du mécanisme de réaction d'URO-synthase par les simulations moleculaires de mécanique quantique.

Les travaux relatifs à cette thèse a été publiée dans la revue scientifique *Proteins: Structure, Function and Bioinformatics* (Cunha, L., Kuti, M., Bishop, D. F., Mezei, M., Zeng, L., Zhou, M.-M., Desnick, R. J., 2008, Human Uroporphyrinogen III Synthase: NMR-based Mapping of the Active Site, Proteins: Structure, Function and Bioinformatics, vol 71:855-873).

SUMÁRIO

O termo “porfiria” refere-se a um grupo heterogeneo de doenças, cada uma resultante da actividade enzimática deficiente de uma enzima específica da via biossintética do heme, e é caracterizada por uma acumulação de porfirinas ou precursores das porfirinas, no corpo. Existem várias classificações possiveis das porfírias. A classificação segundo o local de acumulação distingue entre as porfirias eritropoieticas (porfíria congenita eritropoiética e protoporfíria eritropoiética) e as porfírias hepáticas (porfíria deficiente em ácido aminolevulinico, porfiria aguda intermitente, coproporfíria hereditária, e *variagate porphyria*). Noutra forma de classificação, as porfírias agudas têm manifestações neurológicas, e as porfírias cutâneas têm manifestações na pele.

No primeiro estudo, foi determinada a base molecular responsável pela porfíria cutânea tarda (PCT) em duas famílias portuguesas. É bem sabido que o diagonóstico de porfírias por métodos bioquimícos é problemático, e que um diagonóstico definitivo requer a demonstração do defeito genético responsável. Numa fase inicial deste trabalho analizamos duas familias, portuguesas diagonosticadas com PCT e não relacionadas entre si, e identificamos a mutação E314E como sendo responsável pela doença nestas familias. Esta mutação foi detectada em vários elementos destas famílias. Trata-se uma mutação que afecta o *splicing* normal do gene *UROD*, causando a delecção do exon 9. O ARNm resultante codifíca uma proteína com uma sequência polipeptídica normal até ao exon 8 (291 residuos) e depois o ARNm continua «*out-of-frame*» e codifica 7 amino ácidos anormais do exon 10, e finalmente termina com um codon *stop* preaturo, ao resíduo 298. Estes resultados permitem o diagonóstico molecular pre-simptomático e o aconselhamento destas famílias, permitindo aos individuos portadores desta mutação evitar factores que precipitem a doença.

No segundo e terceiro estudos abordamos aspectos da correlação entre a estrutura e a função de três das enzimas biossintéticas do heme, a «*hydroxymethylbilane synthase* (HMB-synthase)», «*uroporphyrinogen III synthase* (URO-synthase)», e «*uroporphyrinogen decarboxylase* (URO-decarboxylase)», a terceira, quarta, e quinta enzima na cadeia de síntese do heme. Há mais de quatro décadas que vários investigadores sugerem que URO-synthase e HMB interagem num complexo citosólico, para converter porfobilinogénio (PBG) em uroporphyrinogénio III (URO III) eficazmente, conforme é requerido para a biossíntese do heme. Um complexo entre «HMB-synthase» e «URO-synthase» preveniria a conversão não enzimática de «*hydroxymethylbilane* (HMB)» em URO I. Embora bastante progresso tenha sito feito em relação à compreensão da base molecular das enzimas envolvidas nesta via metabólica, um puzzle continua por resolver relativamente ao mecanismo de acção da enzima «*uroporphyrinogen III synthase*». Para

estudos destas duas questões, que requerem quantidades grandes de enzimas puras, no capítulo 2 descrevemos a expressão das enzimas recombinantes em grande quantidade, e um método eficaz de purificação usando vectores de expressão procarióticos, produzidas em *E. coli*. Os ADNc de «HMB-synthase» foram fundidos com uma sequência de 10 histidinas e uma sequência de reconhecimento de Factor Xa. O URO-synthase foi fundido com um péptido modificador relacionada com a ubiquitina (SUMO), o Smt3, entre a sequência de histidinas N-terminais e URO-synthase, e o ADNc de «URO-decarboxylase» foi fundido com um polipéptido modificado da tiorredoxina. As enzimas recombinantes foram expressas em *E. coli*, e purificadas e caracterizadas.

No capítulo 3, descrevemos o uso tecnicas de espectroscopia de ressonancia magnetica nuclear (RMN), e computacionais, para investigar duas questões importantes relativas a sintese do heme: Em primeiro lugar, investigamos a interacção hipotética entre «HMB-synthase» e «URO-synthase». Pare este efeito, as ressonancias dos grupos NH da cadeia peptidica foram identificadas no espectro de coerência quântica heteronuclear unica *(CQNU)*. O estudo da perturbacao das variações quimicas («*chemical shifts*»), em que «URO-synthase» foi marcada com ^{15}N e titulada com «*HMB-synthase*» e/ou «*URO-decarboxylase*» não marcadas, demonstrou que o espectro de «*URO-synthase*» nao foi alterado significativamente na presença de «*HMB-synthase*» ou «*URO-decarboxylase*». Uma vez que as concentrações enzimáticas usadas sao muito maiores do que aquelas encontradas no fígado, é muito pouco provavel que estas enzimas interactuem *in vivo.*

Uma segunda questão aínda por resolver relativamente à sintese do heme tem a ver com o mecanismo através do qual a enzima «*URO-synthase*» converte o «HMB» em URO III. Durante décadas, esforços foram concentrados em resolver este *puzzle*. Tentativas foram feitas em sintetizar o composto intermédio da reacção, assim como em capturar esse estado transitório. Embora o ultimo caso tem sido díficil de conseguir, provavelmente devido a esta espécie química ter uma duração muito curta, or se manter ligado a enzima, a sintese de vários inibidores levou a sugerir um composto intermediario denominado «*spiro-pyrrolenine*». Neste trabalho descrevemos a localização do centro activo da enzima usando técnicas de RMN e computacionais, o primeiro passo para compreender o mecanismo enzimático,.

Experiências de RMN, em que estudamos a perturbação de aminoácidos da enzima titulada com dois inibidores, localizaram o centro-activo na fenda entre os dois domínios globulares da enzima. Esta localização está em acordo com o modelo estabelecido por «docking» que prevê a conformação energéticamente mais favorável do complexo enzima-inibidor. Apesar da estrutura tri-dimensional cristalográfica já ser conhecida anteriormente, o centro-activo hipotético (a fenda entre os dois domínios) é demasiado grande para permitir um ligar justo do inibidor à enzima. Por esse motivo, nós determinamos a estrutura da enzima em solução por técnicas de RMN. Esta estrutura assemelha-se à estrutura determinada por cristalografia de raios-x, no entanto com algumas diferenças. O espaço entre os dois domínios da enzima é mais pequeno na estrutura de RMN. Esta estrutura explica melhor os resultados obtidos por pelas experiencias de RMN usando titulações de inibidores, e o modelo computacional. Grandes movimentos envolvendos dos dois domínios globulares, transitanto entre uma conformaçáo aberta (estrutura cristalina), e uma conformação fechada (estrutua de NMR em solução), pode ser importante para o mecanismo enzimático. Simulações computacionais de dinâmica molecular, uma técnica mais sofisticada de estudar complexos moleculares, confirmaram o modelo obtido através de «docking». Além disso, mutações de residuos especificos localizados na fenda entre os dois dominios resultaram em reduao da actividade especifica da enzima, confirmando que esta fenda é de facto o centro activo.

Em resumo, estes estudos proporcionam um melhor conhecimento da genética molecular, bioquímica e biologia estrutural de três enzimas citosólicas essenciais à síntese do heme.

Os resultados obtidos com o trabalho desta tese foram publicados na revista científica *Proteins: Structure, Function and Bioinformatics* (Cunha, L., Kuti, M., Bishop, D. F., Mezei, M., Zeng, L., Zhou, M.-M., Desnick, R. J., 2008, Human Uroporphyrinogen III Synthase: NMR-based Mapping of the Active Site, Proteins: Structure, Function and Bioinformatics, vol 71:855-873).

INTRODUCTION

Molecular Biology of the Human Heme Biosynthesis

Heme (ferroprotoporphyrin IX) is an important molecule of life, not only because it's the prosthetic group in various proteins, including hemoglobin, cytochromes, catalase, peroxydase, tryptophan pyrrolase, and sulfite reductase hemoproteins, but also, a number of genetic diseases, are associated with its synthesis (Table 1). The hepatic and erythropoietic porphyrias of man constitute a family of inherited disorders each resulting from the deficient activity of a specific enzyme in the heme biosynthetic pathway, which involves eight enzymatic steps for the biosynthesis of heme from glycine and succinyl-coA (Fig. I.1).

The first reaction in the heme biosynthetic pathway occurs in the mitochondria, where δ-aminolevulic acid synthase (ALAS) catalysis the formation of δ-aminolevulinic acid (ALA) from the condensation of a molecule of glycine and a molecule of succinyl-CoA [1]. ALA is then transported into the cytosol, where the next four enzymatic reactions take place. δ-aminolevulinic acid dehydratase (ALAD) dimerizes two molecules of ALA to form porphobilinogen (PBG) [2, 3]. The enzyme binding sites of the two ALA substrates have been designated the A and P sites. The A site gives rise to the acetic side chain, while the P site gives rise to the propionic side chain of PBG [4]. Four molecules of PBG are polymerized, head-to-tail by HMB-synthase [5]. The resulting linear tetrapyrrole hydroxymethylbilane (HMB) is then cyclized to URO III, with inversion of ring D, by URO-synthase [6], also known as co-synthase, re-recognizing the concerted action of both enzymes to form the physiologic, URO III tetrapyrrole. URO III is the precursor of all intermediates of the heme biosynthetic pathway. The fifth enzyme in the pathway, uroporphyrinogen decarboxylase (URO-decarboxylase) catalyzes the decarboxylation of the acetic side chains of URO-synthase III to form the methyl groups of coproporphyrinogen III (COPRO'gen III) [7]. This compound then enters the mitochondria, where the last three steps of the pathway take place. Coproporphyrinogen oxidase catalyzes the oxidative decarboxylation of the two propionate side chains of ring A and B and their conversion to the vinyl groups of protoporphyrinogen IX [8-9]. In the next step, catalyzed by the seventh enzyme in the pathway, protoporphyrinogen oxidase, protoporphyrinogen IX is oxidized to protoporphyrin IX [8-9]. The oxidase reaction requires molecular oxygen and results in the loss of 6 protons and 6 electrons, yielding a completely conjugated ring system, which is responsible for the characteristic red color of heme. In the final step of heme biosynthesis, ferrochelatase catalyzes the insertion of ferrous iron into protoporphyrin IX to form heme [8-9].

Each of the heme biosynthetic enzymes is encoded by a single nuclear gene, with the exception of the first enzyme, ALA-S, which has separate genes for the housekeeping and erythroid-specific enzymes [10]. The full-length human cDNAs for all the nine enzymes have been isolated and sequenced and the genomic organization and/or the entire genomic sequences for all the nine genes have been determined [8,9].

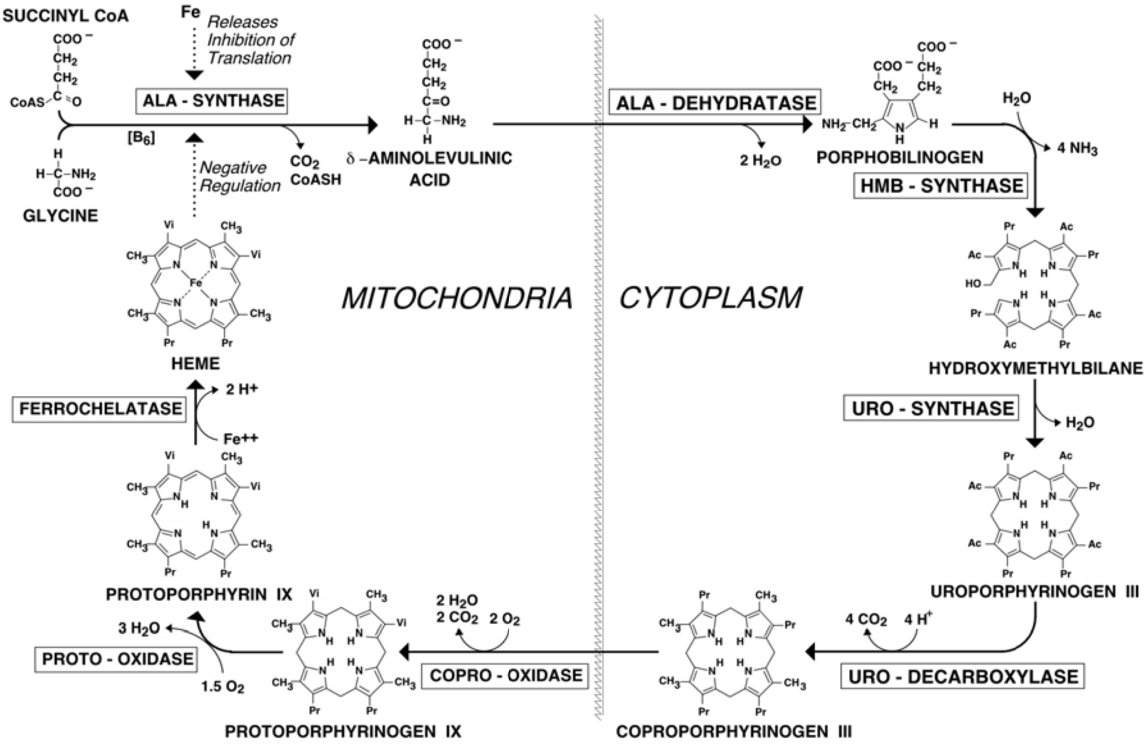


Figure I.1. The heme biosynthetic pathway. The pathway consists of eight enzymes, four localized in the cytoplasm and four in the mitochondria.

Tissue specific regulation:

In humans, about 85% of the heme is synthesized in erythroid cells to provide heme for hemoglobin. Most of the other 15% are produced in the liver and used mainly by cytochrome P-450 enzymes, as well as other hemoproteins. The regulation of heme biosynthesis in the liver is different from that in erythroid cells [10-14]. In the liver, the pathway is under negative feedback control, where the concentration of the first enzyme (the housekeeping form of ALAS 1) is regulated by the concentration of free heme. Heme represses the transcription and translation of liver ALAS 1 mRNA and can also interfere with the transport of the enzyme from the cytosol to the mitochondria. ALAS is inducible by many of the same chemicals that induce cytochrome P-450 enzymes. Because most of the heme synthesized in the liver is used for the synthesis of cytochrome P-450 enzymes, the induction of hepatic ALAS and cytochrome P-450 occur in a coordinated fashion. In erythroid cells, *ALAS* has a separate gene, and *ALAD*, *HMBS* and *UROS* have unique erythroid specific promoters. The erythroid *ALAS* transcription is not negatively regulated by heme and translation is controlled by an iron-responsive element in the 5'-untranslated region of the mRNA [10-14]. *HMB-synthase* has a specific erythroid promoter, in addition to its housekeeping promoter, which generates two isozymes. *ALAD* and *UROS* have housekeeping and erythroid-specific promoters which generate transcripts that encode an identical polypeptide.

The Porphyrias

The inherited porphyrias are a diverse group of inborn errors of metabolism, each due to the deficient activity of a specific heme biosynthetic enzyme (see table I.1). Based on the site of excess porphyrin precursor production, they are classified as hepatic (excess production in the liver) or erythropoietic (excess production in the erythron) [15]. The four hepatic porphyrias (ALAD deficient porphyria, AIP, hereditary coproporphyria and variegate porphyria) have life-threatening, acute neurological attacks characterized by elevated plasma and urinary concentrations of the porphyrin precursors, ALA and PBG. The erythropoietic porphyrias include CEP and erythropoietic protoporphyria. Patients with erythropoietic porphyrias have elevated bone marrow and erythrocyte porphyrins, and their clinical manifestations usually include transfusion-dependent to mild anemia and/or cutaneous photosensitivity causing mild to severe dermatological involvement. For medical management and genetic counseling it is important to establish which porphyria is present in a given patient. Because the symptoms of the various porphyrias are not specific, and misdiagnosis is common in patients with symptoms due to other diseases, when a porphyria is suspected, proper laboratory testing is important. Biochemical tests for porphyria are hard to interpret and abnormal porphyrin results may be due to other disorders. A definitive diagnosis of a particular porphyria requires the identification of the responsible genetic mutation.

Relevant to this work are three of the cytosolic heme biosynthetic enzymes, HMB-synthase, URO-synthase and URO-decarboxylase and their associated porphyrias. During the past decades, biochemical and molecular studies of these disorders have provided insights into the properties and reaction mechanisms of these enzymes, the genomic organization and chromosomal location of their genes, and the regulation and control of heme biosynthesis. Most recently, the 3D structures of *E. coli* ALAD and HMB-synthase and human URO-decarboxylase, URO-synthase, and ferrochelatase (FECH) have been solved [16-20] and their reaction mechanisms (except of URO-synthase) have been characterized at the molecular level. In addition, the nature and frequency of the mutations that cause perturbations in the heme biosynthesis and exacerbate the clinical expression of these diseases have been identified [21-28]. However, these advances have only paved the way to further research to gain fundamental understanding of these enzymes including the three-dimensional (3D) structural basis of the unique reaction mechanisms of URO-synthase, and the metabolic and cellular basis for the bi-compartmental location in the cytosol and mitochondria.

	Compound	**Enzyme**	**Disease**
	Glycine and Succinyl CoA		
		δ-Aminolevulinic acid Synthase	*(Sideroblastic anemia)*
	δ-Aminolevulinic acid		
		δ-Aminolevulinic Acid Dehydratase	ALAD porphyria
	Porphobilinogen		
		Hydroxymethylbilane Synthase	Acute intermittent porphyria
	Hydroxymethylbilane		
Non-enzymatic		Uroporphyrinogen Synthase	Congenital Erythropoietic porphyria
Uroporphyrinogen I	Uroporphyrinogen III		
		Uroporphyrinogen Decarboxylase	Porphyria Cutanea Tarda; Hepatoerythropoietic Porphyria
Coproporphyrinogen I	Coproporphyrinogen III		
		Coproporphyrinogen Oxidase	Hereditary Coproporphyria
	Protoporphyrinogen IX		
		Protoporphyrinogen Oxidase	Variegate Porphyria
	Protoporphyrin IX		
		Ferrochelatase	Erythropoietic Protoporphyria
	Heme		

Table I.1. The heme biosynthetic pathway and associated porphyrias.

Acute Intermittent Porphyria (AIP): AIP, the most common hepatic porphyria (~1 in 10,000 Caucasians), is a dominantly inherited acrogenic condition which results from the half-normal activity of HMB-synthase [23]. There are two major disease subtypes. In the classical AIP subtype, patients have half-normal HMB-synthase activities in all tissues, whereas patients with the erythroid variant have the classic disease manifestations, but normal levels of erythrocyte HMB-synthase activity due to mutations that are housekeeping promoter - or exon 1 splicing - specific [26]. Clinical expression is highly variable in both subtypes and it is estimated that about 80 percent of AIP heterozygotes remain clinically asymptomatic throughout life. However, life-threatening acute neurologic attacks are precipitated by certain dietary or hormonal changes and, particularly, by various drugs that lead to induction of hepatic ALAS1. Therefore, it is important to identify at-risk heterozygotes in affected families so that they may be counseled to avoid porphyrinogenic precipitants. Because the biochemical diagnosis of asymptomatic AIP classical patients may be problematic due to a significant overlap of the high heterozygote and low normal ranges of erythrocyte HMB-synthase activity [26, 27] (Moreover, asymptomatic heterozygotes with the erythroid variant have normal erythrocyte HMB-synthase levels) a reliable diagnosis of classical or variant AIP heterozygotes requires mutation analysis and because . To date, 289 mutations causing AIP have been identified (The Human Gene Mutation Database, http://www.hgmd.cf.ac.uk/ac/index.php), most being “private,” (i.e., occurring in only one or a few unrelated families).

The acute neurologic attacks in all four hepatic porphyrias are precipitated by porphyrinogenic drugs, or by dietary and hormonal changes that induce hepatic ALAS1 activity, resulting in markedly increased levels of ALA, PBG, and other porphyrins [30]. That the ALA and PBG levels remain increased during the attack (as monitored by their high levels in plasma and urine) and return to normal or near normal after the attack in the hepatic porphyrias, suggests a common pathogenesis. However, the mechanism(s) that precipitate ALAS1 induction during an acute attack have not been systematically determined.

<u>Treatment of AIP</u>: The major approach to treat the life-threatening acute attacks of the hepatic porphyrias is prevention. By identifying at-risk heterozygotes for these dominant diseases, they can be counseled to avoid porphyrinogenic drugs, as well as dietary and other changes known to trigger acute attacks [15]. Therefore, efforts have focused on identifying the specific mutations in each family, and screening at-risk family members. Currently, acute neurologic attacks are treated by the intravenous administration of heme or a stable heme conjugate (e.g., heme arginate) which provides exogenous heme for the negative feedback inhibition of ALAS1 and decreased production of ALA and PBG [28-33].

Congenital Erythropoietic Porphyria (CEP): CEP results from the markedly deficient, but not absent, URO-synthase activity [34]. In this autosomal recessive disease, very low URO-synthase activities (≤1% of normal) are compatible with life, although these patients have severe symptoms and are transfusion-dependent. The deficient URO-synthase activity leads to the accumulation of HMB, most of which is converted non-enzymatically to URO'gen I (Fig. I.2). URO'gen I can undergo decarboxylation by URO-decarboxylase to form coproporphyrinogen I (COPRO'gen I), but cannot be metabolized further, except to their non-physiologic and pathogenic oxidized porphyrins, URO I and COPRO I [35]. They damage erythrocytes, cause cutaneous photosensitivity, and are deposited in tissues and bones. Although the disease is an erythropoietic porphyria, the porphyrin levels and symptoms increase at puberty, analogous to the onset of symptoms in the hepatic porphyrias, suggesting that there may be a hepatic or housekeeping component to the disease pathophysiology.

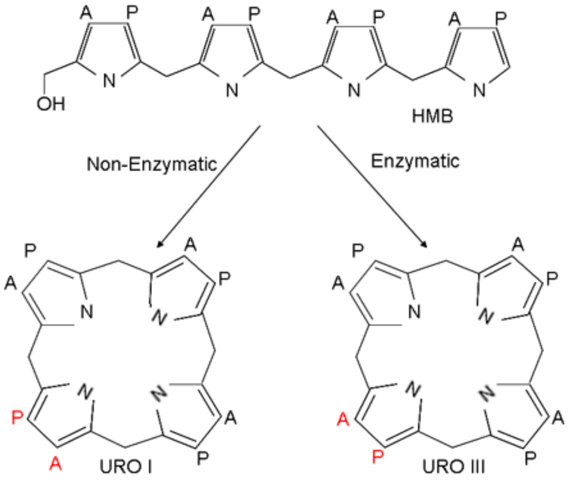


Figure I.2. Conversion of the linear tetrapyrrole HMB: non-enzymatically to URO I, or by uroporphyrinogen synthase to URO III.

The age at onset and clinical severity are highly variable, ranging from non-immune hydrops fetalis due to severe hemolytic anemia *in utero* to milder, later-onset forms, which have only cutaneous lesions in adult life [36]. Factors responsible for the variability include: 1) the amount of residual URO-synthase activity, 2) the degree of hemolysis and consequent stimulation of erythropoiesis, and 3) exposure to ultraviolet light. Life expectancy may be markedly diminished in severely affected patients due to hematologic complications and infections. In most patients, cutaneous photosensitivity usually begins in early infancy with exposure to sunlight and other UV sources and is manifested by blistering of the epidermis on the hands, face, and other sun-exposed areas. The bullae and vesicles are prone to rupture and secondary infection can lead to significant cutaneous scarring and deformities, including loss of digits, eyelids, nostrils, and ears. Corneal scarring can lead to blindness. Other manifestations include erythrodontia, hypertrichosis of the face and extremities [37] and bone loss [38]. Anemia due to hemolysis can be severe if the marrow can not compensate, and severely affected patients are transfusion-dependent. Secondary splenomegaly may cause leukopenia and thrombocytopenia (see table I.2).

To date, 38 *URO-synthase* mutations that cause CEP have been reported. Three mutations account for over 40% of the mutant alleles, while most other lesions were "private". Many of the missense mutations have been expressed and genotype/phenotype correlations have been described.

Classical Symptoms	Phenotypic Heterogeneity
Autosomal recessive inheritance	Severe:
Porphyrinuria	Non-immune hydrops fetalis
Hemolytic anemia	Transfusion-dependent
Secondary Hypertrichosis	Marked skin involvement
Erythrodontia	Moderate:
Hypertrichosis	Moderate anemia
Skin photosensitivity:	Skin involvement
Bullous lesions → Scaring →	Mild:
Infection &	Later onset
deformities	Only skin involvement

Table I.2. Clinical Features of CEP

<u>Treatment of CEP</u>: Current strategies to treat patients with CEP include: skin protection, transfusions with iron removal, marrow suppression to decrease porphyrin production, and BMT [39]. Protection of the skin from sunlight and minor trauma is essential. Sunscreens and β-carotene are sometimes beneficial [40]. Bacterial infections that complicate cutaneous blisters require timely treatment to prevent scarring and mutilation. Severe infections such as cellulitis and bacteremia may require intravenous antibiotics. In severely affected, transfusion-dependent patients, frequent blood transfusions are required. In patients with milder anemia, chronic transfusions (every 2 to 4 weeks) are employed to suppress erythropoiesis, which will decrease porphyrin

production, and result in reduced porphyrin levels and photosensitivity [39]. Such therapy requires deferoxamine treatment to reduce the resulting iron overload. Hydroxyurea has been used to reduce bone marrow porphyrin synthesis [41], and splenectomy may substantially reduce transfusion requirements in some patients [40]. Heme therapy has not been extensively studied [42], but seems unlikely to provide long-term benefit.

Familial Porphyria Cutanea Tarda (fPCT): PCT results from decreased uroporphyrinogen decarboxylase activity in the liver, and is characterized by blistering skin lesions on sun-exposed areas. PCT can be classified into three subtypes: type I (sporadic), type II (familial, autosomal dominant), and type III (familial, rare). The disease can also be induced by exposure to certain environmental chemicals [43]. In Type I PCT, URO-decarboxylase activity is normal in erythrocytes and other non-hepatic tissues. In the Type II disease, an autosomal dominant disorder, the enzymatic activity of URO-decarboxylase is reduced in all cells at to about 50%. Homozygosity for type II PCT results in hepatoerythropoietic porphyria (HEP), which has markedly deficient URO-decarboxylase decarboxylase activity (to about 3-27%). Type III disease is a rare autosomal dominant, characterized by normal erythrocyte URO-decarboxylase activity and deficient activity only in the liver [43]. To date, 84 PCT causing gene mutations have been described.

In PCT, porphyrins accumulate in large amounts in the liver, and are increased in the plasma and urine (predominately uroporphyrin and heptacarboxyl porphyrin). The pattern of porphyrins is complex since URO-decarboxylase deficiency results in the accumulation of porphyrinogens with different numbers of carboxyl groups of both the I and III isomer types, as well as isocoproporphyrinogen, which undergo non-enzymatic oxidation to the corresponding porphyrins [44]. Accumulation of isocoproporphyrins is specific for URO-decarboxylase deficiency.

All three PCT subtypes include cutaneous manifestations and hepatic abnormalities in the absence of neurologic involvement. Symptoms usually develop in adult life. Multiple factors can contribute to inactivation or inhibition of hepatic URO-decarboxylase in this disease, probably by an iron-dependent oxidative mechanism, including alcohol, hepatitis C infection, estrogen, HIV, smoking and factors that increase hepatic iron content such as mutations of the *HFE* gene [44]. Exposures to certain chlorinated polycyclic aromatic hydrocarbons are responsible for occasional outbreaks of PCT.

Cutaneous involvement: Lesions develop on sun-exposed areas of skin. The most common lesions include superficial erosion, bullae, milia (small white plaques) hypertrichosis and pigmentation. Less common lesions may include cutaneous thickening, scarring, and calcification (pseudoscleroderma), scarring alopecia, onycholysis of the nails and photodamage to the eyes (particularly to the conjunctivae and sclerae) [44].

The phototoxic reaction appears to involve the generation of reactive oxygen species, which may be followed by activation of the complement system and lysosomal damage [45-49].

Hepatic involvement: PCT is usually associated with abnormal liver function tests, (especially transaminases and γ-glutamyltranspeptidase [44]) and abnormalities such as siderosis, fatty infiltration, focal necrosis and inflammation of the fibrotic portal.

Treatment of PCT: Symptoms may disappear, following the abstention of precipitating factors, such as alcohol and estrogens. However, phlebotomy remains the standard therapy for PCT, reducing the excess hepatic iron by removing 450 ml of blood at intervals of 1-2 weeks until the serum ferritin reaches the lower limit of normal. Chloroquine, which complexes with uroporphyrin and promotes its release from the liver, is also an effective treatment for PCT [50,51].

The Metabolon Hypothesis and the Bi-Compartmental Subcellular Distribution of the Heme Biosynthetic Enzymes

Four of the heme biosynthetic enzymes function in the mitochondria and four are cytosolic. This bi-compartmental subcellular organization of the heme biosynthetic enzymes presumably evolved to maintain the highly reduced porphyrinogens in the relatively non-oxidative environment of the cytosol (to protect them from oxidation to unmetabolizable porphyrins) The formation of URO’gen III requires that the linear tetrapyrrole, HMB, be immediately converted to the cyclic tetrapyrrole, URO’gen III by URO-synthase [52]. The rapidity of cyclization by URO-synthase and the co-isolation of HMB-synthase and URO-synthase have suggested that these enzymes work in a complex [53-55]. HMB is unstable and in the absence of URO-synthase, rapidly cyclizes non-enzymatically to URO’gen I with a half-life, at pH 8.2 and 37 °C, of 4 min [55-56]. This non-physiologic and toxic URO’gen I isomer can be converted enzymatically to equally toxic COPRO’gen I by the sixth enzyme of the pathway, uroporphyrinogen decarboxylase (URO-decarboxylase), but COPRO’gen I is not a substrate of coproporphyrinogen III oxidase and thus accumulates if formed *in vivo.* Under normal conditions *in vivo*, no URO I or COPRO I is detected.

The first physical evidence of an interaction between any of the cytosolic enzymes was the retention of human URO-synthase on an enzymatically active human erythroid HMB-synthase-Sepharose 4B affinity column [53]. While intriguing, the significance of the putative interaction was unclear since the HMB-synthase was in >50-fold excess over URO-synthase and only one-half of the applied URO-synthase was bound to the affinity column. Using HMB-synthase and URO-synthase purified from wheat germ, Higuchi and

Bogorad (1975) demonstrated an increase in the apparent molecular weight of HMB-synthase in the presence of URO-synthase by sucrose density gradient centrifugation only when porphobilinogen was present [54]. This result can be interpreted that either PBG caused a conformational change that induced interaction or that the HMB-synthase interacted with URO-synthase via its stable enzyme-substrate intermediates [57, 57a]. A marked “rate enhancement” of URO-synthase was observed when HMB-synthase was mixed with URO-synthase in the coupled enzyme assay [58]. In this experiment, when the substrate hydroxymethylbilane (HMB) was presented to URO-synthase as it was generated by HMB-synthase from porphobilinogen (PBG), the apparent affinity constant for HMB binding was 10-fold lower than that for synthetic HMB added directly to URO-synthase (in the absence of HMB-synthase enzyme). This result indicated a functional effect of the HMB-synthase-URO-synthase interaction, but we lacked a demonstration of their physical association.

Studies were needed to determine if URO-synthase and HMB-synthase interact in a complex or “metabolon,” and to determine the precise sub-cellular localization of the four cytosolic enzymes. Furthermore, it would be “organizationally attractive” and efficient if all the four cytosolic enzymes were localized in a multi-enzyme complex for the conversion of ALA to COPRO’gen III. Moreover, the delivery of COPRO’gen III to the mitochondria would be extremely effective if the cytosolic complex was situated on the mitochondrial outer membrane.

Previously, small quantities of human HMB-synthase and URO-synthase were purified from outdated human erythrocytes [55, 57, 58]. Subsequently, milligram quantities of recombinant *E. coli* and *B. subtilis* HMB-synthase and human URO-synthase were purified to homogeneity for physical and kinetic characterization [56, 59, 60]. In order to investigate whether the cytosolic heme biosynthetic enzymes exist in a complex, we established a method for high-level purification and efficient purification of the three enzymes. In order to study the putative cytosolic enzyme complex by nuclear magnetic resonance (NMR) methods (chemical shift perturbation), we determined the NMR backbone resonance assignments of URO-synthase, and titrated [^{15}N, ^{1}H]-labeled enzyme with the putative binding partners.

The active-site of URO-synthase

The 3D structure of human URO-synthase containing N-terminal 22 residue His$_{10}$ tag extension was recently obtained at 1.85 Å resolution, but efforts to investigate the reaction mechanism were unsuccessful [16] due to he inability to crystallize the enzyme with enzyme inhibitors. Furthermore, this structure has large cleft opening, the putative active site, which is too large to tightly bind the substrate, HMB.

Several mechanisms for the conversion of HMB into URO’III have been proposed [61-62], but these are only concerned with the chemical process which the substrate might undergo, and do not take into account the enzyme’s mode of action (these mechanisms are discussed in chapter 3 and 4). To map the active site of URO-synthase, in an effort to understand the reaction mechanism, we determined the NMR 3D solution structure of URO-synthase, and performed NMR chemical shift perturbation experiments in which labeled URO-synthase was titrated with two inhibitors of substrate binding. The NMR experiments provide information regarding the location of the active site and protein-residues likely to be involved in inter-molecular contacts, however does not provide insights into the orientation of the ligand in the active-site. Therefore, we investigated the mode of substrate-binding using *in silico* docking and molecular simulations methods. Interestingly, the computational results were in good agreement with the NMR experimental data.

BIBLIOGRAPHY:

1. Gibson, K.D., W.G. Laver, and A. Neuberger, 1958, Initial stages in the biosynthesis of porphyrins. 2. The formation of delta-aminolevulinic acid from glycine and succinyl-coenzyme A by particles from chicken erythrocytes. Biochem J, 70(1): p. 71-81.
2. Gibson, K.D., A. Neuberger, and J.J. Scott, 1954, The enzymic conversion of delta-aminolaevulic acid to porphobilinogen. Biochem J, 58(4): p. 11-12.
3. Gibson, K.D., A. Neuberger, and J.J. Scott, 1955, The purification and properties of delta-aminolaevulic acid dehydrase. Biochem J, 61(4): p. 618-29.
4. Jaffe E.K., Abrams W.R., Kaempfen H.X., Harris K.A. Jr., 1992, 5-Chlorolevulinate modification of porphobilinogen synthase identifies a potential role for the catalytic zinc. Biochemistry, 31(7): p. 2113-23.
5. Higuchi, M. and L. Bogorad, 1975, The purification and properties of uroporphyrinogen I synthases and uroporphyrinogen III cosynthase. Interactions between the enzymes. Ann N Y Acad Sci,, 244: p. 401-18.
6. Jordan, P.M. and J.S. Seehra, 1979, The biosynthesis of uroporphyrinogen III: order of assembly of the four porphobilinogen molecules in the formation of the tetrapyrrole ring. FEBS Lett., 104(2): p. 364-6.
7. Sano, S. and S. Granick, 1961, Mitochondrial coproporphyrinogen oxidase and protoporphyrin formation. J Biol Chem., 236: p. 1173-80.
8. Karl E. Anderson, S.S., David F. Bishop, Robert J. Desnick, 2001, Disorders of Heme Biosynthesis: X-Linked Sideroblastic Anemia and the Porphyrias. eighth ed. The Metabolic and Molecular Basis of Inherited Diseases, ed. A.L.B. Charles S. Scriver, Willian S. Sly, David Valle,. Vol. II., New York: McGraw-Hill. 71.
9. Desnick, R.J., Astrin, K. H. , Anderson, K., 2002, Inherited Porphyrias, in Principles and Practice of Medical Genetics, J.M.C. David Rimoin, Reed E. Pyeritz, Bruce K. Korf, Editor., Churchill Livingston: London. p. 2586-2623.
10. Sadlon, T.J., Dell'Oso T., Surinya, K.H., May, B.K., 1999, Regulation of erythroid 5-aminolevulinate synthase expression during erythropoiesis. Int J Biochem Cell Biol, 31(10): p. 1153-67.
11. Aizencang, G., Solis, C., Bishop, D.F., Warner, C., Desnick, R.J., 2000, Human uroporphyrinogen-III synthase: genomic organization, alternative promoters, and erythroid-specific expression. Genomics, 70(2): p. 223-31.
12. Bishop, T.R., Miller, M.W., Beall, J., Zon, L.I., Dierks, P., 1996, Genetic regulation of delta-aminolevulinate dehydratase during erythropoiesis. Nucleic Acids Res,24(13): p. 2511-8.

13. Kaya, A.H., Plewinska, M., Wong, D.M., Desnick ,R.J., Wetmur, J.G., 1994, Human delta-aminolevulinate dehydratase (ALAD) gene: structure and alternative splicing of the erythroid and housekeeping mRNAs. Genomics, 19(2): p. 242-8.
14. Cox, T.C., Bawden, M.J., Martin, A., May, B.K., 1991, Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron-responsive element in the mRNA. Embo J,10(7): p. 1891-902.
15. Anderson, K.E., Sassa, S., Bishop, D.F. & Desnick, R.J. , 2001, Disorders of haem biosynthesis: X-linked sideroblatic anemia and the porphyrias, Disorders of haem biosynthesis: X-linked sideroblatic anemia and the porphyrias. Eighth ed. The Metabolic and Molecular Bases of Inherited Disease, ed. A.L.B. C.S. Scriver, W.S. Sly & D. Valle. Vol. II. 2001, New York: McGraw-Hill. 2961-3062.
16. Mathews, M.A., Schubert, H.L., Whitby, F.G., Alexander, K.J., Schadick, K., Bergonia, H.A., Phillips, J.D., Hill, C.P., 2001, Crystal structure of human uroporphyrinogen III synthase. Embo J, 20(21): p. 5832-9.
17. Al-Karadaghi S., Hansson M., Nikonov S., Jonsson B., Hederstedt L.., 1997, Crystal structure of ferrochelatase: the terminal enzyme in heme biosynthesis. Structure,5(11): p. 1501-10.
18. Erskine P.T., Senior N., Awan S., Lambert R., Lewis G., Tickle I.J., Sarwar M., Spencer P., Thomas .P, Warren M.J., Shoolingin-Jordan P.M., Wood .SP., Cooper J.B., 1997,.X-ray structure of 5-aminolaevulinate dehydratase, a hybrid aldolase. Nat Struct Biol, 4(12): p. 1025-31.
19. Louie, G.V,, Brownlie PD, Lambert, R., Cooper, J.B., Blundell, T.L., Wood, S.P., Malashkevich, V.N., Hadener, A., Warren, M.J., Shoolingin-Jordan, P.M., 1996, The three-dimensional structure of Escherichia coli porphobilinogen deaminase at 1.76-A resolution. Proteins, 25(1): p. 48-78.
20. Whitby F.G., Phillips J.D., Kushner J.P., Hill C.P., 1998, Crystal structure of human uroporphyrinogen decarboxylase. Embo J, 17(9): p. 2463-71.
21. Desnick, R.J., Astrin, K. and Anderson, K.E., 2002, Inherited Porphyrias. 4th ed. ed. Emery and Rimoin's Principles and Practice of Medical Genetics, ed. D.L. Rimoin, Connor, J.M. and Pyeritz, R.E., New York: Churchill-Livingstone. 2586-2623.
22. Desnick, R.J., Inborn Errors of Heme Biosynthesis, The Porphyrias. 21st ed. ed. Rudolph's Pediatrics, ed. C.D. Rudolph, Rudolph, A.M., Hostetter, M., Lister, G., and Siegel, N.,. 2002, New York: McGraw Hill Medical Publishing Co. 682-690.
23. Mustajoki, P., 1981, Normal erythrocyte uroporphyrinogen I synthase in a kindred with acute intermittent porphyria. Ann Intern Med, 95(2): p. 162-6.

24. Kappas, A., Sassa, S., Galbraith, R. A. and Nordmann, Y, 1995, The porphyrias. The Metabolic and Molecular Bases of Inherited Disease, ed. A.L.B. C.S. Scriver, W.S. Sly and D. Valle., New York: McGraw-Hill. 2103-2160.
25. Chen, C.H., Astrin, K.H., Lee, G., Anderson, K.E., Desnick, R.J., 1994, Acute intermittent porphyria: identification and expression of exonic mutations in the hydroxymethylbilane synthase gene. An initiation codon missense mutation in the housekeeping transcript causes "variant acute intermittent porphyria" with normal expression of the erythroid-specific enzyme. J Clin Invest, 94(5): p. 1927-37.
26. Astrin, K.H. and R.J. Desnick, 1994, Molecular basis of acute intermittent porphyria: mutations and polymorphisms in the human hydroxymethylbilane synthase gene. Hum Mutat, 4(4): p. 243-52.
27. Moore, M.R., McColl, K. E. L., Rimington, C. and Goldberg, A, 1987, Disorders of porphyrin metabolism, New York. 1-374.
28. Mustajoki, P. and Y. Nordmann, 1993, Early administration of heme arginate for acute porphyric attacks. Arch Intern Med, 153(17): p. 2004-8.
29. Dover, S.B., Moore, M.R., Fitzsimmons, E.J., Graham, A., McColl, K.E., 1993, Tin protoporphyrin prolongs the biochemical remission produced by heme arginate in acute hepatic porphyria. Gastroenterology, 105(2): p. 500-6.
30. Lamon J.M., Frykholm B.C., Hess R.A., Tschudy D.P., 1979, Hematin therapy for acute porphyria. Medicine (Baltimore), 58(3): p. 252-69.
31. Bissell, D.M., 1986, Treatment of acute hepatic porphyria with hematin. J Hepatol, 1988. 6(1): p. 1-7.
32. Jones, R.L., Hematin-derived anticoagulant. Generation in vitro and in vivo. J Exp Med,. 163(3): p. 724-39.
33. Khanderia, U., 1986, Circulatory collapse associated with hemin therapy for acute intermittent porphyria. Clin Pharm, 5(8): p. 690-2.
34. Romeo, G. and E.Y. Levin, 1969, Uroporphyrinogen 3 cosynthetase in human congenital erythropoietic porphyria. Proc Natl Acad Sci U S A, 63(3): p. 856-63.
35. Desnick, R.J. and K.H. Astrin, 2002, Congenital erythropoietic porphyria: advances in pathogenesis and treatment. Br J Haematol, 117(4): p. 779-95.
36. Poh-Fitzpatrick, M.B., 1985, Porphyrin-sensitized cutaneous photosensitivity: pathogenesis and treatment. Clin Dermatol, 3(2): p. 41-82.
37. Desnick R.J., Glass I.A., Xu W., Solis C., Astrin K.H., 1998, Molecular genetics of congenital erythropoietic porphyria. Semin Liver Dis, 18(1): p. 77-84.
38. Poh-Fitzpatrick, M.B., 1986, The erythropoietic porphyrias. Dermatol Clin, 1986. 4(2): p. 291-6.

39. Piomelli S., Poh-Fitzpatrick M.B., Seaman C., Skolnick L.M., Berdon W.E., Complete suppression of the symptoms of congenital erythropoietic porphyria by long-term treatment with high-level transfusions. N Engl J Med, 314(16): p. 1029-31.
40. Mathews-Roth, M.M., 1993, Carotenoids in erythropoietic protoporphyria and other photosensitivity diseases. Ann N Y Acad Sci, 691: p. 127-38.
41. Guarini, L., S. Piomelli, and M.B. Poh-Fitzpatrick, 1994., Hydroxyurea in congenital erythropoietic porphyria. N Engl J Med, 330(15): p. 1091-2.
42. Watson, C.J., Bossenmaier, I., Cardinal, R., Petryka, Z.J., 1974, Repression by hematin of porphyrin biosynthesis in erythrocyte precursors in congenital erythropoietic porphyria. Proc Natl Acad Sci U S A, 71(2): p. 278-82.
43. Granick, S. and D. Mauzerall, 1958, Enzymes of porphyrin synthesis in red blood cells. Ann N Y Acad Sci, 75(1): p. 115-21.
44. Elder, G.H., 1998, Porphyria cutanea tarda. Seminar in Liver Disease, 18(1): p. 67-75.
45. Lim, H.W., M.B. Poh-Fitzpatrick, and I. Gigli, 1984, Activation of the complement system in patients with porphyrias after irradiation in vivo. J Clin Invest, 74(6): p. 1961-5.
46. Meurer M., Schulte C., Weiler A., Goerz G., 1985, Photodynamic action of uroporphyrin on the complement system in porphyria cutanea tarda. Arch Dermatol Res, 277(4): p. 293-8.
47. Torinuki, W., T. Miura, and H. Tagami, 1985, Activation of complement by 405-nm light in serum from porphyria cutanea tarda. Arch Dermatol Res, 277(3): p. 174-8.
48. Sandberg S., Romslo I., Hovding G., Bjorndal T., 1982, Porphyrin-induced photodamage as related to the subcellular localization of the porphyrins. Acta Derm Venereol Suppl (Stockh), 100: p. 75-80.
49. Nomura N., Lim H.W., Levin J.L., Sassa S., 1998, Effect of soluble complement receptor type 1 on porphyrin-induced phototoxicity in guinea pigs. J Photochem Photobiol B, 42(1): p. 28-31.
50. Tsega E., Besrat A, Damtew B., Seyoum E., Landells J.W., 1981, Chloroquine in the treatment of porphyria cutanea tarda. Trans R Soc Trop Med Hyg, 75(3): p. 401-4.
51. Kordac V., Jirsa M, Kotal P., Kalab M., Cervinka J., Kotyk A., Martasek P., Agents affecting porphyrin formation and secretion: implications for porphyria cutanea treatment. Semin Hematol, 1989. 26(1): p. 16-23.
52. Battersby A.R., Fookes C.J., Matcham G.W., McDonald E., 1980, Biosynthesis of the pigments of life: formation of the macrocycle. Nature, 285(5759): p. 17-21.

53. Frydman, R.B. and G. Feinstein, 1974, Studies on porphobilinogen deaminase and uroporphyrinogen 3 cosynthase from human erythrocytes. Biochim Biophys Acta, 350(2): p. 358-73.

54. Higuchi, M. & Bogorad, L., 1975, The purification and properties of uroporphyrinogen I synthases and uroporphyrinogen III cosynthase. Interactions between the enzymes. Ann. N. Y. Acad. Sci. 244, 401-18.

54. Rossetti M.V., Juknat de Geralnik A.A., Kotler M., Fumagalli S., Batlle A.M., 1980, Occurrence of multiple molecular forms of porphobilinogenase in diverse organisms: the minimum quaternary structure of porphobilinogenase is a protomer of one deaminase and one isomerase domain. Int J Biochem, 12(5-6): p. 761-7.

55. Anderson, P.M. and R.J. Desnick, 1980, Purification and properties of uroporphyrinogen I synthase from human erythrocytes. Identification of stable enzyme-substrate intermediates. J Biol Chem, 255(5): p. 1993-9.

56. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1987, Coupled-enzyme and direct assays for uroporphyrinogen III synthase activity in human erythrocytes and cultured lymphoblasts. Enzymatic diagnosis of heterozygotes and homozygotes with congenital erythropoietic porphyria. Anal Biochem, 166(1): p. 120-33.

57. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1987, Purification and properties of uroporphyrinogen III synthase from human erythrocytes. J Biol Chem, 262(3): p. 1268-73.

58. Alwan, A.F., B.I. Mgbeje, and P.M. Jordan, 1989, Purification and properties of uroporphyrinogen III synthase (co-synthase) from an overproducing recombinant strain of Escherichia coli K-12. Biochem J, 264(2): p. 397-402.

59. Stamford, N.P., Capretta A., and Battersby A.R., 1995, Expression, purification and characterisation of the product from the *Bacillus subtilis* hemD gene, uroporphyrinogen III synthase. Eur J Biochem, 231(1): p. 236-41.

60. Omata Y., Sakamoto H., Higashimoto Y., Hayashi S., And Masato Noguchi, 2004, Purification and Characterization of Human Uroporphyrinogen III Synthase Expressed in *Escherichia coli.* J Biochem, 136, 211–220

61. Mathewson J.H., Alsoph C., Biosynthesis of pyrrole Pigments: A Mechanism of Porphobilinogen Polymerization. *J. Am. Chem. Soc.* 1961, *83*, 135-137.

62. Battersby A.R., Baker M.G., Broadbent H. A., Fookes C.J.R., Leeper F.J., 1987, Biosynthesis of Porphyrins and Related Macrocycles. Part 29. Synthesis and Chemistry of 2,2-Disubstituted 2H-Pyrroles (Pyrrolenines). J. chem. *soc.* Perkin Trans. I p. 2027-2048

CHAPTER 1: A SPLICING MUTATION IN THE UROPORPHYRINOGEN DECARBOXYLASE GENE IS RESPONSIBLE FOR PCT IN TWO PRESUMABLY UNRELATED PORTUGUESE FAMILIES

SUMMARY

Familial porphyria cutanea tarda (fPCT) in two presumably unrelated Portuguese families was caused by a splicing mutation in the uroporphyrinogen decarboxylase gene (*UROD*). This mutation (E314E) led to skipping of exon 9 and was previously identified in an Argentinean family. RT-PCR using primers spanning exons 5 to 10 showed both a normal allele and the mutated allele 67 bp shorter, consistent with the deletion of exon 9. The resultant mutant *UROD* mRNA encoded a normal polypeptide sequence through exon 8 (291 residues), then was out of frame and encoded seven abnormal amino acids from exon 10, and finally truncated prematurely at residue 298. Of interest, the mutant protein was 69 amino acids shorter than wild type enzyme, indicating that the C-terminal residues are important for enzymatic activity or stability. This study permits pre-symptomatic molecular diagnosis and counseling of these families to enable family members to avoid disease precipitating factors.

INTRODUCTION

Familial porphyria cutanea tarda (fPCT), an autosomal dominant inborn error of metabolism, results from the half normal activity of uroporphyrinogen decarboxylase (UROD, EC 4.1.1.37) in all cells [1]. This enzyme is the fifth in the heme biosynthetic pathway and sequentially decarboxylates the four acetate side chains of uroporphyrinogen III and its isomer, uroporphyrinogen I [2]. The 40.8 kDa monomeric enzyme is encoded by a single gene localized in chromosomal region 1p34 [3]; It contains 10 exons spread over 3 kb [4]. The complete cDNA sequence has been reported [4,5]. In contrast to the alternative genes or promoters for erythroid regulation of the first four enzymes of the heme biosynthetic pathway, the *UROD* gene has only a single promoter, containing both housekeeping and erythroid control elements.

The enzyme from human erythrocytes has been purified and its physical and kinetic properties have been determined [6,7]. Recently, the recombinant human enzyme was over-expressed, purified, and its crystal structure determined [8,9]. The protein is comprised of a single domain containing a $(\beta/\alpha)_8$-barrel with the active site having a deep cleft formed by loops at the C-terminal ends of the barrel strands. fPCT is an autosomal dominant disorder characterized by incomplete penetrance, as 90% of carriers remain asymptomatic [10]. The half-normal activity of UROD caused by a genetic mutation is not sufficient to cause disease and other factors are required [11]. The major risk factors for the expression of PCT are chronic hepatitis C, alcohol abuse, iron overload, estrogen use, and co-inheritance of haemochromatosis gene (HFE) mutations [12-15]. Of interest, iron does not appear to directly affect UROD activity. The effect of iron on hepatic UROD probably results from the formation of reactive oxygen species and cytochrome P450 induction leading to an oxidation product that is derived from ALA, or possibly uroporphyrinogen and that irreversibly inhibits UROD [10, 16]. Other subtypes of PCT are sporadic porphyria cutanea tarda, in which the decreased UROD activity is restricted to the liver and there is no family history, and hepatoerythropoietic porphyria (HEP), the rare homozygous form of fPCT, characterized by 75-90% reduction in UROD activity in all cells [10]. In all these subtypes, porphyrins accumulate in the liver and are increased in the plasma and urine, leading to skin deposition causing photosensitivity. The main clinical manifestations are skin fragility, bullous lesions, and hypertrichosis on sun-exposed areas [17]. Phlebotomy and low-dose chloroquine are specific forms of treatment in PCT that are virtually always effective.

To date, 84 mutations causing fPCT have been described. Most of these are missense mutations, but ten splice-site mutations - all resulting in exon deletions, nine small deletions, five small insertions, four gross deletions, 1 indel and one regulatory mutation

(Human Gene Mutation Database http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html). With the exception of A80G, R142Q, G281E, IVS6^{+1}, E314E, G10insA, each identified in several families, the other mutations were private, each found only in one family, emphasizing the molecular heterogeneity of the mutations causing this disease. In this communication, we identify E314E, a mutation previously described in an Argentinean family of Italian ancestry [18], as the fPCT causing mutation in the two Portuguese families.

MATERIAL AND METHODS

Patients - Patients from two unrelated f-PCT families diagnosed at the Hospital de Santo Antonio, Porto, Portugal, were studied (Fig. 1.1). A summary of their clinical data at diagnosis is given in Table 1.1.

Fresh peripheral blood samples were collected from each f-PCT patient, with informed consent. Cultured lymphoid cell lines were established and maintained as described elsewhere [19]. DNA was isolated from peripheral blood by the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN) and poly (A)+ mRNA was isolated from cultured lymphoblasts by the RNeasy Mini Kit (Qiagen, Valencia, CA).

DNA Analysis – The entire URO-D gene was amplified from genomic DNA (1 μg) in a single reaction using the long-range PCR (LR-PCR) primers (Table 1.2) and the GeneAmp XL PCR Kit (Perkin-Elmer), according to the manufacturer's instructions. For the first 16 cycles, annealing and extension were for 3 min each at 67 °C, and for the last 12 cycles the times were increased 15 s each cycle, with a final extension cycle of 10 min at 72 °C. A portion of the PCR product was analyzed by agarose-gel electrophoresis to determine that the long-range reaction was successful. PCR amplicons encompassing one or more exons, and the respective intron/exon boundaries, were sequenced with the primers indicated in Table 1.2, with 50 ng of the LR-PCR product used as a template. Sequencing was accomplished in the Mount Sinai DNA core facility by the Sanger enzymatic method using Applied Biosystems (Foster City, CA) Prism BigDye® fluorescent terminators, version 1.1 chemistry. Unreacted dye terminators were removed by gel filtration on Qiagen (Valencia, CA) DyeEx 96 plates. Analysis was conducted on an ABI Prism® 3700 capillary DNA sequencer using Sequence Analysis software version 3.7.

Restriction Fragment Length Polymorphism (RFLP) – for RFLP analysis, 1μg of PCR-amplified DNA (intron 7-9), was digested with *Dde* I endonuclease according to the manufacture's instructions, and analyzed in a 1% agarose gel.

RT-PCR – RT-PCR analysis was used to characterize the splicing mutation. Poly(A)+ mRNA (1 µg) was used for cDNA synthesis using the PromegaR Reverse Transcription System (Madison, WI), according to the manufacturer's instructions. PCR was performed with 10 μl of the reverse transcription reaction. In order to analyze the E314E G>A splicing mutation, the primers were located in exon 5 and 10 (Table 1.2).

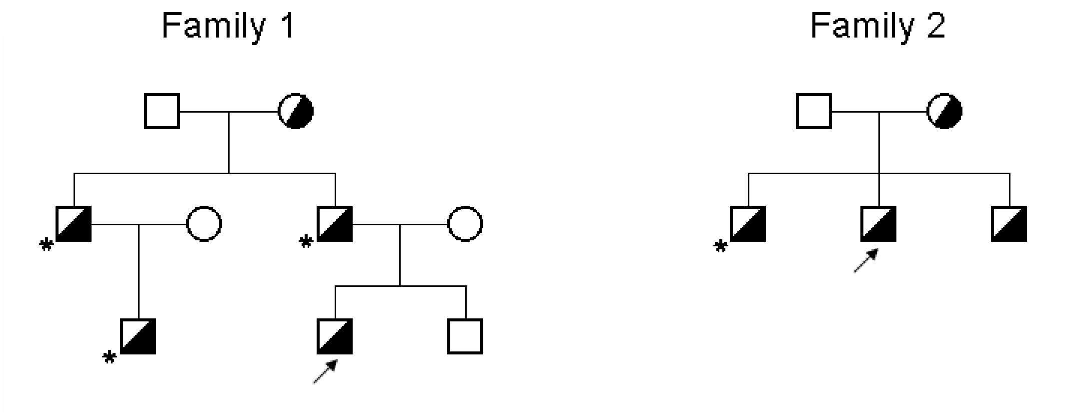


Figure 1.1 - Pedigrees of the Portuguese f-PCT families diagnosed in this study. The proband of each family is indicated with an arrow. The asterisk indicates the family members who participated in this study.

Family number	Generation/ subjects' number	Age at diagnosis	Cutaneous manifestations (1)	Urinary uroporphyrin (μg/24h)	Aspartate-amino-transferase (U/L)	Alanine-amino-transferase (U/L)	Transferrin Saturation (%)	Serum Ferritin (ng/ml)
1	II/1	54	Yes	3855	40	86	36	523
1	II/2	46	Yes	1555	33	42	45	397
1	III/1	27	Yes	2301	27	61	48	245
1	III/2	23	Yes	3726	33	77	55	161
2	II/1	45	Yes	742	35	51	33	242
2	II/2	33	Yes	6254	25	96	32	401

Table 1.1 - Summary of clinical data from patients included in this study. The family numbers and subject's references are reported to Fig. 1.1. (1) - include skin fragility, bullous lesions, skin pigmentation and hypertrichosis.

Primer Type	UROD Region	Amplicon size (bp)	Tm (^{o} C)	Sense F – Forward R - Reverse	Sequence
Long-range PCR:	5’-UTR	3699	67.6 64.9	F R	5'-TATGGACCTGGCTGGATAAGACTGTTGGT-3' 5'-GGGACAATCTTTCACAAACAAAACTACAC-3'
Sequencing	5’-UTR Intron 1	868	47.6 48.8	F R	5'-TTACAGCGGAGTTTTCTTTT-3' 5'-GTGGGGGTAGGTGTCAGTAT-3'
	Intron 1 Intron 4	834	48.1 50.0	F R	5'-GACTAAGTGGAGGGAAGAGG-3' 5'-AGGAGGAAGGAAAAGGAGAA-3'
	Intron 4 Intron 6	681	46.0 56.0	F R	5'-GTTTTATTCTCCTTTTCCTT-3' 5'-TTCTTTCATCCCTGGCAT -3'
	Intron 4 Intron 8	1381	43.0 50.9	F R	5'-TTTATTCTCCTTTTCCTTCC -3' 5'-AACACCCCCACCTCAACCTC -3'
	Intron 6 Intron 9	845	48.0 47.5	F R	5'-TTAGGGTCAGGCAGTATCAG -3' 5'-TGAACAACAGCAACAAAAAG -3'
	Intron 8 3’-UTR	772	57.0 46.2	F R	5'-AGGTTGAGGTGGGGGTGTTG -3' 5'- GAGGCTAAGAAGAACTGAGG-3'
RT-PCR	Exon 5 Exon 10	668	67.4 53.9	F R	5’-CCGACAACGACTGGCTGGAC-3’ 5’-AGCAGACGTGAGTGTTTATG-3’

Table 1.2 - Primers for long-range PCR amplification, sequencing, and RT-PCR of the *UROD* gene.

RESULTS

Mutation characterization - The entire UROD sequence of each proband was amplified by long-range PCR, and each exon and adjacent intronic or flanking boundaries were amplified and sequenced. Sequence analysis revealed a G to A transition at cDNA nucleotide 960, the last nucleotide of exon 9 (Fig. 1.2). This mutation did not alter the coded amino acid (glycine), but it was located in the consensus splice donor site of exon 9.

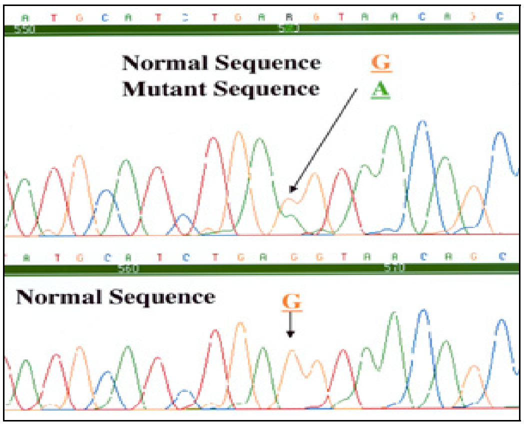


Figure 1.2 – *UROD* gene sequencing. **Above**: Electropherogram of proband 1 showing the G to A transition at nucleotide 960. The proband was heterozygous for this mutation and the wild type allele. **Below**: Electropherogram showing the partial sequence region neighboring nucleotide 960 of a normal individual.

RFLP analysis was used to detect the presence of this mutation in 3 other affected members of family 1 and one affected member of family 2 (Fig 1.3). PCR was performed with primers located in introns 6 and 9, and the amplification product was digested with *Dde* I endonuclease. The nt960 G→A mutation abolished a consensus sequence for this restriction enzyme and when compared to a normal control, the probands showed an extra digestion band of 411 bp, corresponding to the deletion of the *Dde* I recognition sequence.

Exon skipping demonstrated by RT-PCR - RT-PCR experiments conducted on the RNA extracted from cultured lymphoblasts showed that this mutation altered the normal splicing of the *UROD* pre-mRNA. After PCR amplification with primers located in exons 5 and 10, a shorter band of about 600 nucleotides was visible in conjunction with the normal 660 bp fragment (Fig. 1.4). Sequence analysis of the shorter band showed the deletion of 67 bp corresponding to the length of exon 9. A premature stop codon within exon 10 was then created. The resulting protein had 291 normal amino acids followed by other seven different amino acids coded by the out of frame sequence of exon 10. The subsequent out-of-frame codon (299) was a termination codon which resulted in a 298 residue truncated polypeptide instead of the normal 367 long peptide.

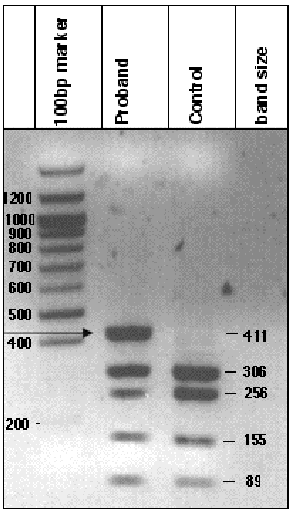


Figure 1.3 – RFLP confirmation of the G to A transition in nucleotide 960 of the Proband 1's *UROD* cDNA. A fragment of DNA from introns 7-9 was digested with *Dde* I endonuclease, as described in methods. The G→A mutation abolished a *Dde* I recognition site. The Proband was heterozygous for the nt960 G→A mutation, hence in addition to the normal number of restriction bands seen in the control sequence, there is an extra band, band of 411 nt, corresponding to non-cutting between the 256 and 155-bp fragments of the mutated allele. Analysis of proband 2's cDNA identified the same E314E. Both proband's family members were screened for this mutation by RFLP analysis (data not shown), resulting in the diagnosis of several other affected family members.

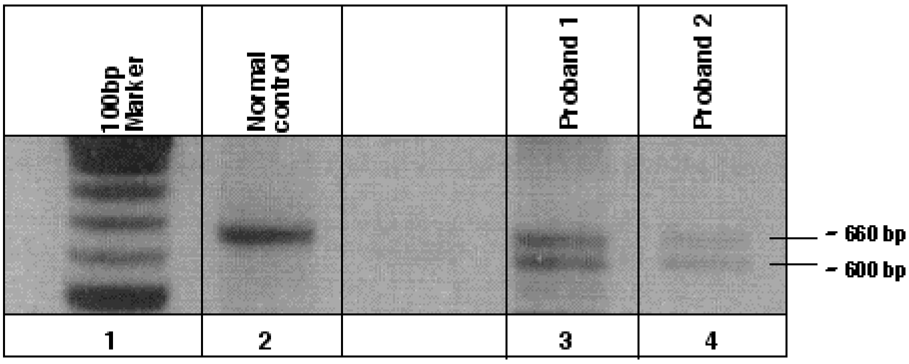


Figure 1.4 - RT-PCR confirmation of the abnormal splicing caused by the E314E mutation. The forward and reverse primers were located in exon 5 and 10, respectively. Lane 1: 100 bp marker; Lane 2: normal control; Lane 3: family 1 proband; lane 4: family 2 proband.

DISCUSSION

A previously described mutation, E314E [18], was identified as the disease-causing mutation in two presumably unrelated Portuguese families with f-PCT. This mutation altered the normal splicing of the *UROD* mRNA, resulting in an aberrant species lacking exon 9. RFLP analysis confirmed the presence of the E314E in all affected family members studied.

This codon was degenerate and correctly encoded a glutamic acid residue, but the 5' donor splice site changed from GAGgtaaca to GAAgtaaca. RT-PCR experiments demonstrated that this mutation resulted in abnormal splicing of the *UROD* mRNA and the joining of exons 8 and 10. The deletion of exon 9 removed 22 amino acids from the translated peptide. Exon 10 was, consequently, out of frame, and encoded an abnormal sequence of 7 amino acids before a stop codon. This represents a major disruption of the protein structure, since several structural elements that define the protein core, as well as the active site, are abolished. Therefore we presume the protein is catalytically inactive and/or structurally unstable. In eukaryotes, an mRNA undergoes several processing steps before it can be translated into the corresponding protein. Numerous mechanisms mediated by the eukaryotic gene expression machinery ensure that only properly processed mRNAs that are capable of encoding functional polypeptides are translated. One such mechanism, the nonsense mediated mRNA decay (NMD) pathway, is a post-transcriptional process that rapidly degrades mRNAs with premature translation termination codons (aberrant mRNAs that terminate translation upstream of the last splice junction) [20]. Preliminary quantitative-PCR studies, using mRNA extracted from cultured lymphoblasts of both probands, have indicated a reduction in the levels of UROD mRNA (data not shown).

To date, 84 mutations have been identified in the *UROD* gene, as the genetic defect responsible for fPCT. Ten of these mutations are splicing mutations, including the G→A, E314E mutation described in this work, previously only identified in an Argentinean family of Italian ancestry [18].

These studies identified the causative mutation, E314E, in these Portuguese f-PCT families which now permit the molecular diagnosis of this disease. Importantly, at risk asymptomatic family members can be screened to identify those carrying the mutation for appropriate counseling to avoid the known precipitating factors. It is intriguing that a high number of affected individuals are present in both families, considering fPCT has a low penetrance value of ~10%. At least one of the families is heterozygous for the H63D haemochromatosis mutation. However a recent study suggested that the prevalence of the H63D mutation in PCT patients is not higher than that in controls and does not play a significant role in PCT [15]. Although the importance of C282Y HFE mutation in iron storage and PCT is supported by several studies, this mutation was not present in these Portuguese families.

BIBLIOGRAPHY:

1. Kushner, J.P., A.J. Barbuto, and G.R. Lee, 1976, An inherited enzymatic defect in porphyria cutanea tarda: decreased uroporphyrinogen decarboxylase activity. J Clin Invest, 58(5): p. 1089-97.
2. Lash, T.D., 1991, Action of uroporphyrinogen decarboxylase on uroporphyrinogen-III: a reassessment of the clockwise decarboxylation hypothesis. Biochem J, 278 (Pt 3): p. 901-3.
3. Dubart A., Mattei M.G., Raich N., Beaupain D., Romeo P.H., Mattei J.F., Goossens M., 1986, Assignment of human uroporphyrinogen decarboxylase (URO-D) to the p34 band of chromosome 1. Hum Genet, 73(3): p. 277-9.
4. Moran-Jimenez M.J., Ged C., Romana M., Enriquez De Salamanca R., Taieb A., Topi G., D'Alessandro L. de Verneuil H., 1996, Uroporphyrinogen decarboxylase: complete human gene sequence and molecular study of three families with hepatoerythropoietic porphyria. Am J Hum Genet, 58(4): p. 712-21.
5. Mendez M., Sorkin L., Rossetti M.V., Astrin K.H., del C Batlle A.M., Parera V.E., Aizencang G., Desnick RJ., 1998, Familial porphyria cutanea tarda: characterization of seven novel uroporphyrinogen decarboxylase mutations and frequency of common hemochromatosis alleles. Am J Hum Genet, 63(5): p. 1363-75.
6. Sassa S., de Verneuil H., Anderson K.E. Kappas A., 1983, Purification and properties of human erythrocyte uroporphyrinogen decarboxylase: immunological demonstration of the enzyme defect in porphyria cutanea tarda. Trans Assoc Am Physicians, 96: p. 65-75.
7. Elder, G.H., Tovey J.A., Sheppard D.M., 1983, Purification of uroporphyrinogen decarboxylase from human erythrocytes. Immunochemical evidence for a single protein with decarboxylase activity in human erythrocytes and liver. Biochem J, 215(1): p. 45-55.
8. Laterriere M., d'Estaintot B.L., Dautant A., Precigoux G., Hombrados I., De Verneuil H., 1998, Expression, purification, crystallization and preliminary X-ray diffraction analysis of human uroporphyrinogen decarboxylase. Acta Crystallogr D Biol Crystallogr, 54 (Pt 3): p. 476-8.
9. Whitby F.G., Phillips J.D., Kushner J.P., Hill C.P., 1998, Crystal structure of human uroporphyrinogen decarboxylase. Embo J, 17(9): p. 2463-71.
10. Granick, S. and D. Mauzerall, 1958, Enzymes of porphyrin synthesis in red blood cells. Ann N Y Acad Sci, 75(1): p. 115-21.

11. Elder, G.H., 1998, Porphyria cutanea tarda. Seminar in Liver Disease, 18(1): p. 67-75.

12. Bonkovsky, H.L., R.W. Lambrecht, and Y. Shan, 2003, Iron as a co-morbid factor in nonhemochromatotic liver disease. Alcohol, 30(2): p. 137-44.

13. Elder, G.H., 1999, Alcohol intake and porphyria cutanea tarda. Clin Dermatol, 17(4): p. 431-6.

14. Egger N.G., Goeger D.E., Payne D.A., Miskovsky E.P., Weinman S.A., Anderson K.E., 2002, Porphyria cutanea tarda: multiplicity of risk factors including HFE mutations, hepatitis C, and inherited uroporphyrinogen decarboxylase deficiency. Dig Dis Sci, 47(2): p. 419-26.

15. Nagy Z., Koszo .F, Par A., Emri G., Horkay I., Horanyi M., Karadi O., Rumi G. Jr., Morvay M., Varga V., Dobozy A., Mozsik G., 2004, Hemochromatosis (HFE) gene mutations and hepatitis C virus infection as risk factors for porphyria cutanea tarda in Hungarian patients. Liver Int, 24(1): p. 16-20.

16. Mukerji, S.K. and N.R. Pimstone, 1990, Free radical mechanism of oxidation of uroporphyrinogen in the presence of ferrous iron. Arch Biochem Biophys, 281(2): p. 177-84.

17. Mascaro, J.M., Porphyria cutanea tarda: clinical manifestations, 1991, Curr Probl Dermatol, 20: p. 79-90.

18. Mendez M., Sorkin L., Rossetti M.V., Astrin K., del Carmen Batlle A.M., Parera V, Aizsencang G., Desnick R. J. 1998, Familial porphyria cutanea tarda: characterization of seven novel uroporphyrinogen decarboxylase mutations and frequency of common hemochromatosis alleles. Am. J. Hum Genetics, Nov;63(5):1363-75.

19. Anderson, M.A. and J.F. Gusella, Use of cyclosporin A, 1984, in establishing Epstein-Barr virus-transformed human lymphoblastoid cell lines. In Vitro, 20(11): p. 856-8.

20. Singh, G. and J. Lykke-Andersen, 2003, New insights into the formation of active nonsense-mediated decay complexes. Trends Biochem Sci, 28(9): p. 464-6.

CHAPTER 2: PURIFICATION AND CHARACTERIZATION OF RECOMBINANT HUMAN HYDROXYMETHYLBILANE SYNTHASE, UROPORPHYRINOGEN III SYNTHASE AND UROPORPHYRINOGEN DECARBOXYLASE

SUMMARY

Hydroxymethylbilane synthase (HMB-synthase), uroporphyrinogen III synthase (URO-synthase) and uroporphyrinogen decarboxylase (URO-decarboxylase) catalyze the third, fourth, and fifth cytosolic reactions in the heme biosynthetic pathway, which converts porphobilinogen (PBG) to coproporphyrinogen III. The cDNA of each enzyme was inserted in a prokaryotic expression system, and the recombinant human enzymes expressed in *E. coli.* The recombinant human enzymes were purified to homogeneity by metal affinity, ion exchange, and gel filtration chromatography. Milligram amounts of each enzyme were obtained and their properties were determined. The recombinant enzymes showed similar properties to those reported for the human enzymes purified from erythrocytes or liver. Contrary to the previous evidence that some of these enzymes are highly unstable, conditions for long term stability, as required for subsequent structural studies by nuclear magnetic resonance, were determined.

INTRODUCTION

Human Heme Biosynthesis

Heme biosynthesis involves eight enzymatic steps for the synthesis of heme from glycine and succinyl-CoA. The first and last three reactions occur in the mitochondria, whereas the other four enzymatic steps occur in the cytosol. Each enzyme is encoded by a single nuclear gene, with the exception of the first enzyme, 5-aminolevulinate synthase (ALAS), which has separate genes for its housekeeping (ALAS1) and erythroid-specific (ALAS2) isozymes. The full-length human cDNAs encoding all nine enzymes have been isolated and sequenced [1-13], and their genomic organization and/or entire genomic sequences have been determined [12-20]. Notably, these genes are dispersed on seven human chromosomes [11-35], in contrast to their adjacent or operon-like genomic organization in many bacteria and other species.

Despite the large understanding of the pathway, genomic organization and gene regulation, molecular basis of disease, and biochemical mechanisms of most of its components, two outstanding questions remain. One concerns the bi-compartmental organization of the pathway in two cellular compartments (the mitochondria and the cytosol) and the long held hypothesis that the cytosolic enzymes are organized in a complex. The second concerns the fascinating and puzzling mechanism by which URO-synthase stereo-specifically converts HMB to URO III, with inversion of ring D, avoiding the formation of the non-physiologic isomer URO I. Previous attempts to probe for an enzymatic complex with traditional methods (e.g. two-hybrid system, co-immunoprecipitations) were not successful in demonstrating a complex [36]. However, these methods might not have the required sensitivity to detect transient, weak interaction (in the milimolar range). Nuclear magnetic resonance (NMR) is an ideal technology to study protein-protein interactions. It can detect very weak protein-ligand interactions, being the ligand either another protein or small-molecule inhibitors for the study of dynamic processes involved in substrate binding and catalysis. A major disadvantage of NMR is that, first it requires access to such equipment; secondly it requires large amounts of isotope-labeled sample (milimolar concentrations), which needs to be stable at about 30^{o} C for large periods of time. Isotope labeling of protein samples is expensive and therefore one needs a high-level expression system and a high yield purification protocol in place. In this chapter we describe the recombinant expression and purification of three of the four human cytosolic heme biosynthetic enzymes - HMB-synthase, URO-synthase, and URO-decarboxylase.

Hydroxymethylbilane Synthase (also known as URO’gen-I-synthase or PBG-deaminase, EC 4.3.1.8) is the third enzyme in the pathway, which catalyzes the step-wise, head-to-tail condensation of four PBG molecules by a series of deaminations to form the linear tetrapyrrole, hydroxymethylbilane (HMB) [37]. The enzyme has been purified from a variety of sources, including human erythrocytes [37]. The human enzyme has five forms, the free enzyme, and the mono-, di-, tri- and tetra-pyrrole substrate-enzyme intermediates [37]. The enzyme also makes a unique, covalently bound, dipyrromethane cofactor from PBG which binds the pyrrole intermediates at the active site [38-41]. A single ~10 kb gene, mapped to 11q24.1->q24.2 [20, 29], encodes erythroid-specific and housekeeping isozymes that are monomers of 40 and 42 kDa, respectively [3, 42, 43]. The gene has 15 exons and a variety of single nucleotide polymorphisms (SNPs) [14]. Tissue specificity results from alternative splicing of exons 1 and 2 [42]. The erythroid-specific promoter contains elements which bind the GATA-1, NF-E1, and NF-E2 erythroid-specific transcription factors [44]. The crystal structure of *E. coli* HMB-synthase (which has ~48% amino acid identity with the human enzyme) has been solved [45]. However, both human isozymes have 48 additional residues before the homologous 34 C-terminal amino acids. These 48 residues may form another domain involved in interactions with URO-synthase or in binding to a component of a complex involving the other cytosolic heme biosynthetic enzymes. To date, the human recombinant erythroid and housekeeping enzymes have not been expressed and purified in large amounts, nor have their physicokinetic properties been characterized.

Uroporphyrinogen-III-Synthase (URO-synthase, also known as URO'gen co-synthase, EC 4.2.1.75) is the fourth enzyme in the pathway catalyzes the rapid cyclization and rearrangement of HMB (by inversion of the pyrrole D ring) to form the asymmetric and physiologic uroporphyrinogen III (URO'gen III) isomer (Scheme. 2.1). In the absence of URO-synthase, HMB non-enzymatically cyclizes to form the toxic uroporphyrinogen I (URO'gen I) isomer. This non-physiologic compound can be metabolized to coproporphyrinogen (COPRO'gen) I, but further metabolism cannot proceed since the next enzyme is stereo-specific for the III isomer. URO'gen I and COPRO'gen I are oxidized to their respective porphyrins, uroporphyrin I (URO I) and coproporphyrin I (COPRO I). URO-synthase has been purified in small amounts from several sources including human erythrocytes [50]. The human enzyme is a monomer with a molecular weight of about 29,500. The full-length cDNA encoding human URO-synthase has been isolated, and the gene has been mapped to the chromosome region 10q25.2→q26.3 [30]. The ~34 kb human URO-synthase gene has 10 exons and two alternative promoters that generate housekeeping or erythroid-specific transcripts with unique 5'-untranslated sequences (exons 1 and 2A) followed by nine shared coding exons (2B to 10) [51]. Their transcripts encode the same 265 amino acid polypeptide that is active as a monomer [47]. Expression arrays revealed that the housekeeping transcript was present in all tissues, while the erythroid transcript was only in erythropoietic tissues. The housekeeping promoter lacked TATA and SP1 sites, consistent with its observed low level expression in most cells, while the erythroid promoter contained GATA1 and NF-E2 sites for erythroid specificity. The proximal GATA1 site was shown to be functional and near a putative CP2 site.

Recently, the recombinant human enzyme was used to determine its 3D crystal structure [48]. However, this recombinant enzyme contained an N-terminal 21 residue His_{10} tag extension. The 3D structure was obtained at 1.85 Å resolution, but efforts to investigate the reaction mechanism were unsuccessful. Thus, large amounts of purified, active, native human URO-synthase are needed for studies of the enzyme's properties, its reaction mechanism, and of the putative interaction with HMB-synthase).

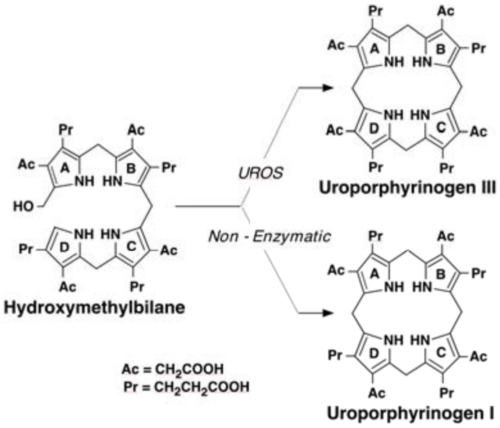


Figure 2.1 - HMB to URO'gen I or III.

Uroporphyrinogen-Decarboxylase (URO-decarboxylase, EC 4.1.1.37), the fifth enzyme in the pathway sequentially decarboxylates the four acetate side chains of URO'gen III and its isomer, URO'gen I [49-53]. The enzyme from human erythrocytes has been purified and its physical and kinetic properties have been determined [48]. The 40.8 kDa protein is encoded by a single gene localized in the chromosomal region 1p34 [24]; the human URO-decarboxylase cDNA was cloned and sequenced [4]. It contains 10 exons spread over 3 kb [16]. In contrast to the unique expression mechanisms for erythroid regulation of the first four enzymes of the heme biosynthetic pathway, the URO-decarboxylase gene has only a single promoter. Analysis of the URO-decarboxylase promoter revealed a TATA-like sequence at -21 and one box at -60. The gene contain two initiation-of-transcription sites, however, both sites are apparently used with the same frequencies in all tissues, including erythroid cells, and the gene is transcribed as a single mRNA [54]. The 3D structure of the human enzyme has been recently determined [55].

MATERIAL AND METHODS

Materials – Fluorescamine, porphobilinogen, leupeptin, lysozyme, and imidazole were obtained from Sigma (St Louis, MO). SuperBroth media was from Qbiogene, Inc. (Carlsbad, CA) and Luria Broth from GIBCO/Invitrogen (Carlsbad, CA) and tris (2-carboxyethyl) phosphine, hydrochloride (TCEP-HCl) was from BioVectra™dcl (Charlottetown, PEI, Canada). Factor Xa was obtained from Novagen (Madison, WI). The Bradford protein assay reagent and bovine serum albumin were purchased from Bio-Rad (Hercules, CA). All other chemicals were from Fisher Scientific (Pittsburgh, PA) except as noted

Protein Determination - Protein concentrations were routinely carried out using the Bradford method according to the manufacturer's instructions. In order to facilitate specific activity comparisons for enzymes previously assayed by different methods, protein concentrations were also determined by the modified Lowry (*DC*™ Protein Assay, BioRad, Richmond, CA), fluorescamine [56], and UV absorbance [57] methods, all referenced to HMB-synthase and URO-synthase solutions whose protein concentrations were determined by amino acid composition analyses performed at the Keck Molecular Analysis Laboratory, Yale University (New Haven, CT) is shown (Tables 2.2, 2.4, 2.6). When reporting kinetic values for the pure enzymes, the data were corrected to the amino acid composition values. The molecular mass of each purified protein was determined by gel filtration chromatography and by matrix-assisted laser desorption ionization/time-of-flight (MALDI-TOF) mass spectroscopy using a Voyager DE-STR instrument (Applied Biosystems, Foster City, CA). A Superdex 70 column was calibrated with the following molecular weight standards: bovine serum albumin (M_r 67,000), ovalbumin (M_r 44,000), carbonic anhydrase (M_r 29,000) and cytochrome c (M_r 14,200) (5 mg each in a volume of 1ml). For MALDI-MS analysis, a modified thin-layer method of sample preparation was used [58]. A saturated 4HCCA matrix was prepared in a 1:2 water/acetonitrile solvent mixture and further diluted 4-fold with isopropanol (IPA). Diluted matrix solution (20 μL) was applied over the gold surface of the sample plate (50 × 50 mm). The organic solvents were allowed to dry in ambient air until only traces of moisture were left on the plate. The matrix was then gently wiped with a tissue, leaving behind a faint layer of 4HCCA visible only as a yellowish reflection when the plate was examined at an angle, “the ultrathin layer".

<u>MALDI-MS Sample Preparation</u>: The 4HCCA matrix was prepared as a saturated solution in a 3:1:2 (v/v/v) mixture of formic acid/water/IPA (FWI). Each sample was diluted 120-fold.

An aliquot (0.5 μL) of the protein-matrix solution was spotted onto the ultrathin layer. A white opaque polycrystalline film of matrix forms at the spot within seconds. As soon as the crystalline film appeared homogeneous, the excess liquid was removed by vacuum aspiration. The spots were then washed for a few seconds with 0.1% aqueous Trifluoroacetic acid (TFA) to remove salts. The 0.1% TFA was subsequently removed by vacuum suction. Myoglobin (100 fmol/μl) was used as an internal calibrant. The samples and the calibrant were analyzed in a MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometer (Voyager DE-STR, Applied Biosystems, Foster City, CA) operating in linear delayed extraction mode. This instrument uses a nitrogen laser that delivers pulses of ultraviolet light (wavelength = 337 nm) at 3 Hz to the matrix spots, each pulse yielding a full mass spectrum. Five hundred individual scans were averaged prior to acquisition into a single spectrum. The spectra were smoothed and further analyzed using the software M-over-Z (http://prowl.rockefeller.edu).

Heme Biosynthetic Enzyme Assays –

Assay of Hydroxymethylbilane Synthase (HMB-synthase) - The HMB-synthase assay was performed with the following modifications of the previously described method [37]. The HMB-synthase reaction mixture contained 500 μl of 0.1 M Tris•HCl, pH 8.0 and 0.1 mM DTT, 200 μl of 0.5 mM PBG, and 50 μl of enzyme preparation in the same buffer. Following a 30 min incubation at 37° C, the reaction was terminated by the addition of 250 μl of 50% TCA (wt/vol) and exposed to ~310 nm UV light for 30 min at room temperature to oxidize the porphyrinogens. The fluorescence was quantitated in an Optical Technology Devices Ratio-2 System Fluorometer (Elmsford, NY) with excitation at 405 nm using a narrow band-pass interference filter and emission at 618 nm obtained with an Optical Technology Devices 595 nm No. 2-63 sharp cut-off filter. The instrument was calibrated with 0 to 200 nmoles/ml of URO I Fluorescent Standard from Frontier Scientific (Logan, UT). One unit of activity is defined as that amount of enzyme required to form one nanomole of URO I per hour at 37° C.

Assay of Uroporphyrinogen III Synthase (URO-synthase) - URO-synthase was assayed as previously described for the coupled-enzyme assay [59], with the following modifications. Erythrocyte lysates containing HMB-synthase were replaced with 5 units (one unit is that amount of enzyme that catalyzes the formation of one nanomole of HMB per hour) of recombinant HMB-synthase. In addition, increased HPLC resolution of the porphyrin isomers was obtained by using a 5 micron particle size BDS Hypersil column (Thermo Electron Corp., Bellefonte, PA) equilibrated with 12% acetonitrile in 1 M ammonium acetate, pH 5.16. Detection was enhanced using a Model 474 scanning fluorometer with excitation at 405 nm and emission at 618 nm using a gain of 100 and an attenuation of 32 (Waters, Milford, MA). One unit (U) of activity is defined as that amount of enzyme required to form one nanomole of URO III per hour at 37° C under the conditions of the assay.

Assay of Uroporphyrinogen decarboxylase (URO-decarboxylase) - URO-decarboxylase was assayed using enzymatically reduced porphyrins as substrates [60]. HMB-synthase (6 U) and 50 μl of 2.4 mM PBG were incubated at 37° C in 0.1 M Tris, pH 7.65 containing 0.1 mM DTT, to produce 50 mM URO'gen I in 30 min. The reaction was stopped by adjusting the buffer pH to 6.8 with 0.15 M KH_2PO_4, pH 4.5. In the dark, 300 μl of this mixture was combined with 200 μl URO-decarboxylase in 0.1M KH_2PO_4 pH 6.8. The reaction was incubated at 37° C for 30 min and stopped with 500 μl of 3 M HCl. UV-mediated oxidation and HPLC separation of the decarboxylation products was as described above, except the starting gradient solutions were 50% (v/v) methanol/NaH_2PO_4, pH 4.5 and 100% methanol. One U of activity is defined as that amount of enzyme required to form one nanomole of coproporphyrin (COPRO) I per hour at 37° C.

Gel Electrophoresis - The expression and purification of the recombinant proteins were monitored by SDS-PAGE electrophoresis according to Laemmli [61]. The *Benchmark* Prestained Protein Ladder (Invitrogen, Carlsbad, CA) was used as molecular weight standard. Non-denaturing gel electrophoresis was performed in a PhastGel system (Amersham Biosciences, Piscataway, NJ) using homogenous 20 gel and native buffer strips.

Thermostability - For stability studies, samples of each purified enzyme, equalized for activity, were incubated at 4, 30, 37, 42 or 45 and 60^{o} C for up to 72 hours. The stability of URO-synthase also and HMB-synthase was evaluated for 14 days at 30^{o} C and at -80^{o} C for 2 years. For the short term determinations (hours), aliquots were removed at timed intervals, placed on ice, and the enzyme activity determined as described above for each enzyme.

Physicokinetic Determinations - The pH optima of purified URO-synthase and HMB-synthase were characterized over the pH range of 6.5-9.0 in 50 mM Bis-Tris Propane, 100 mM NaCl and 2 mM DTT, and that of URO-decarboxylase over the pH range of 5.0-8.0 in 0.2 M KH_2PO_4, 0.1 mM DTT. The thermostabilities of the purified enzymes were determined on aliquots of each enzyme diluted to the desired concentration (permitting direct assay without further dilution) in the same buffer used for NMR (20 mM NaH_2PO_4, pH 7.45, 100 mM NaCl, 1 mM EDTA, and 2 mM DTT) containing 1 mg/ml bovine serum albumin and incubated at 4, 30, 37, 45, and 60° C for up to 14 days, except as indicated in the text. Samples were placed on ice at timed intervals, and the enzyme activities were determined at the end of the experimental time-course. The coupled-assay was used to determine the apparent K_m of recombinant URO-synthase as described by Tsai [50]. Homogeneous recombinant HMB-synthase was added in amounts yielding HMB concentrations ranging from 0.03 to 7.8 mM. For determining the k_m of HMB-synthase, the enzyme was diluted in 100 mM Tris buffer, pH 8.2, with a series of substrate concentrations ranging from 6.6 to 333 mM.

Construction of Expression Vectors -

Human HMB-Synthase Expression Constructs - The human housekeeping HMB-synthase (H-HMB-synthase; nucleotides 28 to 1113; GenBank accession number X04808) and erythroid HMB-synthase (E-HMB-synthase; nucleotides 79 to 1113; GenBank accession number X04808) cDNA coding regions were PCR-amplified to add an *Nde* I restriction site at the ATG start codon using the forward primers, 5'-AGCGGAGCCCATATGTCTGGTAACGGCAAT-3' (housekeeping) or 5'-AGCCCAAAGCA TCTGAGAGTGATTCGCGTG-3' (erythroid) and a *Bgl* II site was added to the 3' end using the reverse primer, 5'-CAGAGATCTTTAATGGGCATCGTTAAGCTGCCGTGC-3'. The PCR product was ligated into the *Nde* I/*Bam*H I-digested pET-16b expression vector (Novagen, Madison, WI) after the T7 promoter and the 5' His•Tag to generate plasmids pH1 (housekeeping) and pH2 (erythroid) that were then transformed into the *E. coli* strain BL21-CodonPlus(DE3)-RIL-X as described above. The Factor Xa-cleaved recombinant enzyme was identical to the wild-type sequence with a predicted molecular mass of 39,326 Da.

Human URO-Synthase Expression Constructs - This vector was engineered to include the yeast Small Ubiquitin-related MOdifier (SUMO) peptide, Smt3p, between the N-terminal His•Tag™ and URO-synthase. This peptide conferred increased solubility and was attached to URO-synthase via a unique cleavage site for the Smt3p-specific Ubiquitin-Like Protein 1 (Ulp1) protease. The yeast SUMO peptide Smt3p (nucleotides 328-621, GenBank accession number U27233), kindly provided by Dr. Troy Burke (Columbia University, NY), was inserted between the *Nhe* I and *Bam*H I sites of the pET28a N-terminal His•Tag™vector (designated pS1). The human URO-synthase cDNA encoding amino acids 1 to 265 (nucleotides 197 to 994; GenBank accession number J03824) was PCR-amplified using a forward primer that added a 5' *Bgl* II cloning site (5'-GGAGGAAGATCTATGAAGGTTCTTTTACTG-3'), and a reverse primer with a 3'-flanking

Sal I site (5'-CGGCGGGTCGACTCACAGCAGCAGCAGCC ATGGGG-3'). The URO-synthase *Bgl* II-*Sal* I insert was subcloned into the *BamH* I-*Sal* I double-digested pS1 vector, designated pSU1, and sequence confirmed in both orientations. The pSU1 plasmid was transformed into *E. coli* strain BL21-CodonPlus(DE3)-RIL-X (Stratagene, La Jolla, CA) for high-level expression. After cleavage by Ulp1, the recombinant URO-synthase was identical to wild-type enzyme plus one amino-terminal serine, for a total predicted molecular mass of 28,712 Da.

Human URO-Decarboxylase Expression Construct - The pThioHisA prokaryotic expression vector (Invitrogen, Carlsbad, CA) was modified by site-directed mutagenesis to replace the internal *Nde* I site with AGCATG, followed by a second mutagenesis step with the sequence 5'-GATGACGATGACAAGGTACCTCTGGTGCCGCGCGGCAGCCATATG CATGAGCTCGA GATCTTC-3' to introduce a 5' *Nde* I cloning site and to replace the enterokinase cleavage site with a thrombin cleavage site. The human URO-decarboxylase cDNA coding region (nucleotides 109 to 1212; GenBank accession number NM_000374) was PCR-amplified to add 5' *Nde* I and 3' *Xho* I cloning sites using the forward primer 5'-CAGCTGACCCATATGGAAGCGAATGGGTTG-3' and reverse primer 5'-CACCTCGAGT CAGTTCTGTCGAAGCAGACGTGAGTGTTT-3', digested with *Nde* I and *Xho* I, cloned into the modified pThioHis vector, and then transformed into *E. coli* strain Top10 (Invitrogen). The thrombin-cleaved recombinant URO-decarboxylase enzyme contained three extra amino-terminal residues (GlySerHis) and had a predicted molecular mass of 41,066 Da.

Yeast Ulp1 Protease Expression Construct - The *Saccharomyces cerevisiae* Ulp1 protease was used for cleavage of the URO-synthase/SUMO fusion peptide. The Ulp1 cDNA region encoding residues 7412 to 8071 of GenBank accession number U36624 was PCR-amplified from the *Saccharomyces cerevisiae* cosmid c9901 (ATCC, Manassas, VA) to yield the forward and reverse sequences 5'-GGGAATTCCATATGCTGGTGCCGCGCG GCAGCCTTGTTCCTGAATTAAATGAAAAAG-3' and 5'-CGCCGCGGATCCTCATTTTAA AGCGTCGGTTAAAAT C-3', respectively. The PCR product was sub-cloned into the *Nde I* and *BamH I* restriction sites of the pET16b vector to generate plasmid pU1 with the protease fused to a Factor Xa-cleavable His•Tag for affinity purification. The plasmid was transformed into the host strain BL21-CodonPlus(DE3)-RIL-X (Stratagene).

Overexpression and Purification of the Recombinant proteins -

Overexpression and Purification of Human Recombinant URO-Synthase - A 10 ml overnight culture in Luria Broth [LB] containing 50 μg/ml kanamycin and the pSU1 vector in BL21-CodonPlus(DE3)-RIL was used to inoculate 1.5 liters of SuperBroth (Qbiogene, Carlsbad, CA), or 2 L of M9 minimal medium (20% (W/V) glucose, 37.5 mM Na_2HPO_4, 17 mM KH_2PO_4, 18.7 mM NH_4Cl, 8.6 mM NaCl, 1.2 mM $MgSO_4$, 0.12 mM $CaCl_2$, and 5% thiamine in deionized water or 99.9% ^2H_2O). At A_{600} = 2 for SuperBroth or 0.7 for M9 minimal medium cultures, expression was induced with 1 mM IPTG for 4 hr at 37° C, or 16 hr at 20° C (SuperBroth or minimal media cultures, respectively) with shaking. All subsequent steps were preformed at 4° C. The cells were harvested at 3,000 x g for 30 min. For lysis, the bacterial pellet was resuspended in 40 ml of cold buffer A (50 mM HEPES, pH 7.8, 0.5 M NaCl and 5 mM TCEP-HCl) containing, 1 μM leupeptin and 0.1 mM PMSF, 330 μg/ml lysozyme (Type L7651, Sigma, St. Louis, MO), 5 μg/ml each DNase and RNase A (Sigma), and stirred gently for 30 min on ice. The suspension was freeze-thawed three times using a dry-ice/ethanol bath. The lysate was cleared by centrifugation at 35,000 x g for 30 min and then directly loaded onto a 1.6 x 7 cm immobilized metal affinity chromatography (IMAC) column containing 15 ml of (nickel) Chelating Sepharose™ Fast Flow resin (GE Healthcare, Piscataway, NJ) equilibrated with buffer A. After washing with 10 column volumes of buffer A containing 50 mM imidazole, the bound URO-synthase was eluted with 5 column volumes of buffer A containing 250 mM imidazole.

Prior to ion-exchange chromatography, the eluted protein was digested overnight with Ulp1 protease (1 μg/mg URO-synthase) and buffer exchanged to buffer B (50 mM HEPES, pH 7.0, 2 mM DTT) by ultrafiltration, using an Amicon Ultra 15 ml, 10,000 molecular weight cut-off, centrifugal ultrafiltration device (Millipore, Bedford, MA) according to the manufacturer's instructions. The cleaved URO-synthase preparation was loaded onto a Mono Q HR 1.0 x10 cm ion exchange column (GE Healthcare) equilibrated with buffer B and eluted with a 240 ml gradient of 0 to 100% buffer B containing 200 mM KCl at a flow-rate of 1.0 ml/min. The peak fractions containing URO-synthase were combined, concentrated to 1 ml and exchanged to buffer C (20 mM sodium phosphate, pH 7.45, 150 mM NaCl and 2 mM DTT), as described above.

The concentrated post-Mono Q URO-synthase sample was fractionated into 1.0 ml aliquots from a HiLoad 1.6 x 60 cm Superdex 75 gel filtration (GF) column (GE Healthcare) equilibrated with buffer C at a flow-rate of 1.0 ml/min. The peak fractions were concentrated by ultrafiltration as described above and the buffer was exchanged to 20 mM sodium phosphate, pH 7.45, 100 mM NaCl, 2 mM DTT, and 0.1 mM EDTA as described above. For long-term storage at -80° C, 10% glycerol was added.

Overexpression and Purification of Human Recombinant HMB-Synthase - HMB-synthase was expressed and purified in three chromatographic steps as described for URO-synthase with the following modifications. Selection was with 50 μg/ml ampicillin. The IMAC eluate was adjusted to 5 mM $CaCl_2$ and the 21-residue His•Tag N-terminal extension was removed by digestion with 0.5 units of Factor Xa protease per mg of HMB-synthase for 2 days at 4° C. Any undigested enzyme was removed by binding the eluate to a 2 ml IMAC column equilibrated with buffer A and eluted with 5 column volumes of 30 mM imidazole in buffer A. Mono Q ion-exchange chromatography resolved the HMB-synthase isozymes with 0 to 3 covalently-attached PBG moieties (stable enzyme-substrate intermediates E, ES_1, ES_2 and ES_3 [37]) which were further purified by gel filtration and stored as described above.

Overexpression and Purification of Human Recombinant URO-Decarboxylase - URO-decarboxylase was expressed in the Top10 host (Invitrogen), lysed from 4 liters of culture, and purified as for HMB-synthase with the following modifications. The IMAC wash buffer contained 30 mM imidazole. The eluate was adjusted to 2.5 mM $CaCl_2$, and the thioredoxin N-terminal extension was removed by digestion with 1 U of thrombin protease (Novagen, Madison, WI) per mg of URO-decarboxylase eluate for 24 hr at 4° C. The Mono-Q elution buffer contained 250 mM KCl.

Overexpression and Purification of Yeast Recombinant Ulp1 Protease - The Ulp1 protease used to cleave the Smt3p-URO-synthase fusion protein was expressed from the pU1 vector in the *E. coli* BL21-CodonPlus(DE3)-RIL host and purified in one-step using IMAC chromatography as described above. The His•Tag was not removed. One microgram of protease was sufficient to cleave 1 mg of the URO-synthase fusion protein when incubated overnight in the IMAC elution buffer. Due to the fact that the Ulp1 protease recognizes a specific structure rather than a peptide sequence, it was highly specific and provided a clean product with no additional cleavages. The enzyme was stored at 1 mg/ml with addition of 10% glycerol, at -135° C, as storage at higher temperatures resulted in loss of proteolytic activity within a few months.

RESULTS

Overexpression and Purification of Recombinant human Housekeeping HMB-synthase and erythroid HMB-Synthase) - The purification of erythroid (E-) and housekeeping (H-) HMB-synthase, as well as their properties were the same. As an example, the purification and properties of H-HMB-synthase are described in detail. Following IPTG-induced synthesis in *E. coli*, the levels of human housekeeping or erythroid HMB-synthase expressed by the pET16b vector were up to 25% of total bacterial proteins as estimated by SDS-PAGE. The enzymes were purified to homogeneity in three chromatographic steps with about 36% yield. Table 2.1 summarizes a typical purification profile recombinant human HMB-synthase from a 6 L culture. The crude extract contained 7.6 x 10^5 units of activity at a concentration of 250 U/mg of protein. Nickel affinity chromatography provided an additional 3.4-fold purification of the fusion protein to 852 U/mg with 84% yield. Of note, the optimal concentration of imidazole for washing and elution of the fusion protein bound to the nickel column were 100 mM and 200 mM, respectively.

Purification	**Vol**	**HMB-synthase Activity**		**Protein**		**Purification**		
(Step)	(ml)	(U/ml)	(Units)	(mg/ml)	(mg)	(U/mg)	(Fold)	(%)
Crude Lysate	245	3125	7.6 x 10^5	12.5	3060	250	1	100
Nickel Affinity	137	4686	6.4 x 10^5	5.5	753	852	3.4	84
Mono Q	150	1878	2.8 x 10^5	1.6	240	1170	4.6	37
Superdex 70	179	1509	2.7 x 10^5	1.18	211	1280	5.12	35

Table 2.1 - Typical Purification of recombinant human housekeeping HMB-synthase from a 6L culture

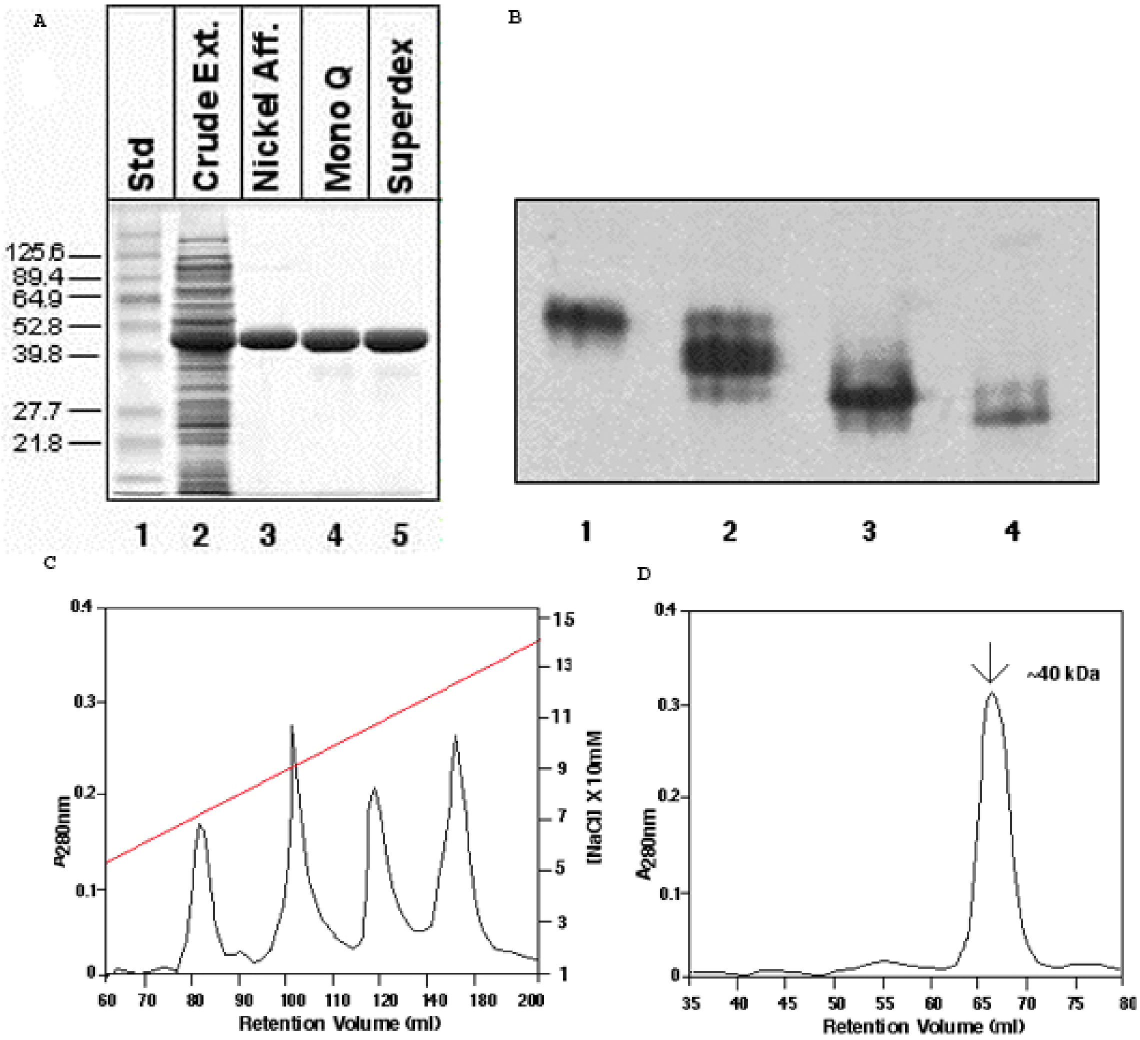


Figure 2.2 – Protein purification of H-HMB-synthase. **A,** SDS-PAGE, documenting protein expression and purification of H-HMB-synthase: Lane 1 – LMW marker; lane 2 – crude lysate; lane 3 – IMAC chromatography; lane 4 – ion exchange chromatography ; lane 5 – gel filtration chromatography**; B,** Native electrophoresis of the MonoQ isolated enzyme-substrate intermediates; **C,** Mono Q ion-exchange chromatogram showing the separation of the enzyme-substrate intermediates; **D,** Gel filtration chromatogram of the purified HMB-synthase (pooled MonoQ fractions), run on a Superdex-70 (16/60) column (Amersham Biosciences).

Most bacterial proteins were excluded after the chelation affinity purification (Fig. 2.2A, lane 3). The histidine tag was cleaved by digestion with Factor Xa, and the enzyme preparation was again chromatographed on the nickel affinity column to separate any undigested enzyme. Further purification and separation of the enzyme-substrate intermediates was achieved with high resolution MonoQ ion exchange chromatography

(Fig. 2.2B, [37]). This step resulted in 1.3 fold purification (Fig. 2.2A, lane 4), with an excellent yield. Subsequent size selection by gel filtration increased the purity by 1.11-fold, and resulted in a homogeneous enzyme protein (Table 2.1 and Figure 2.2). From the original 6 L of culture media, over 200 mg of the homogeneous enzyme was obtained. This represented a 7-fold over the crude bacterial extract with an overall yield of 36%.

Properties of recombinant human HMB-synthase –

The molecular weight was accurately determined by mass spectroscopy (Figure 2.3). The apoenzyme with the bound dipyrrole cofactor, after cleavage of the histidine tag with Factor Xa protease, had a molecular weight of 39978 Da (theoretical weight 39874 Da). Depending on the number of bound PBG molecules, the ES_1, ES_2 and ES_3 fractions had molecular weights of 40290, 40484 and 40775 Da respectively. Note that these fractions were not homogeneous but contained primarily the indicated enzyme-substrate complex as shown in Fig. 2.1B. To determine the physical and biochemical properties of purified H-HMB-synthase, the time course of the H-HMB-synthase activity was determined. The optimal pH was 8.2 and the enzyme had less than 20% of that activity at pH 6.0 (Figure 2.4). Several different protein determination methods were referenced to the enzyme concentration determined by amino-acid analysis to facilitate comparisons with other studies (Table 2.2), thereby avoiding up to two-fold errors of estimation.

Thermostability of housekeeping HMB-synthase - The thermostability of purified HMB-synthase was determined at 4, 30, 37, 42, and 60° C at pH 8.2 (100 mM Tris, 2 mM DTT) and pH 7.4 (20 mM $NaPO_4$, 100 mM NaCl, 2 mM DTT). The enzyme retained full activity after 2 hrs, even when incubated at 60^o C. Further studies revealed that the enzyme remained fully active after 14 days at 30^o C, and 6 months at 4^o C (data not sown). For long term storage, the enzyme was stable at -80° C, in 15% glycerol, for over 2 yrs, in 20 mM $NaPO_4$, 100 mM NaCl, 2 mM DTT.

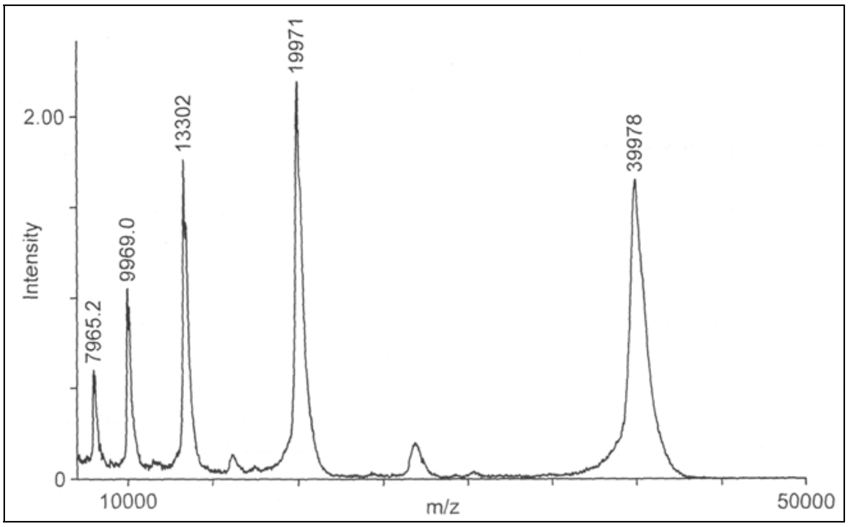


Figure 2.3 – Mass spectrometry profile of the MonoQ fraction containing the recombinant H-HMB-synthase apoenzyme (with the dipyrrole cofactor). The enzyme has a molecular of 39.978 kDa.

Kinetic properties of housekeeping HMB-synthase – For kinetic studies, different concentrations of substrate and HMB-synthase were tested to find the linear range of initial velocity. In steady state, the Km of housekeeping HMB-synthase was 48.4 μM and the V_{max} was 1717 nmoles/hour/mg (Fig 2.5). The properties of recombinant human Housekeeping HMB-synthase are summarized in table 2.2.

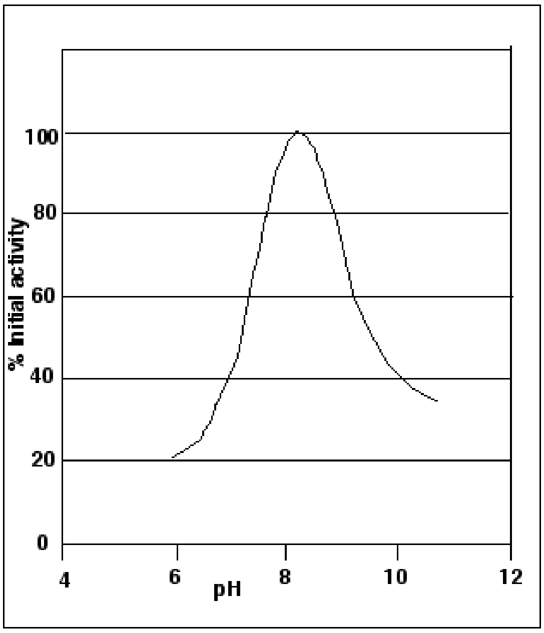


Figure 2.4 – Effect of pH on the activity of purified H-HMB-synthase, indicating the optimal pH value of 8.2

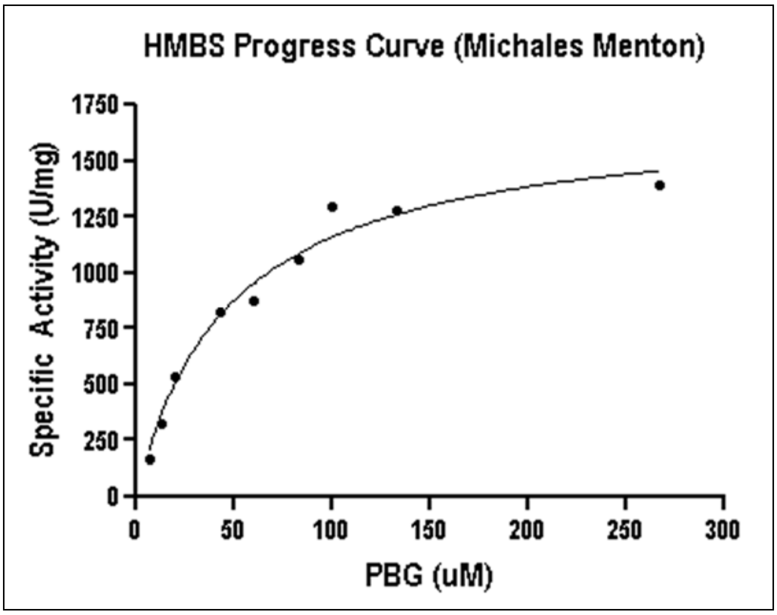


Figure 2.5 - Plot of the specific activity of H-HMB-synthase as a function of the substrate concentration. The Km was 48.4 μM and the V_{max} was 1717 nmoles/hour/mg. This plot was generated by non-linear regression fitting of the kinetic data using the statistics software GraphPad Prism (GraphPad Software, Inc., San Diego, CA).

Property	HMB-Synthase
Mr (Theoretical, daltons)	39,742[a]
Mr (Mass Spec, daltons)[b]	39,874
Protein (Bradford, mg/ml)[c]	1.54 ± 0.09
Protein (DC Lowry, mg/ml)[c]	1.67 ± 0.18
Protein (Fluram, mg/ml)[c]	0.74 ± 0.02
Protein (A_{280}, mg/ml)[c]	0.94 ± 0.06
pH Optimum	8.2
K_m (μM)	48.4
K_{cat} (min^{-1})	1.13
Catalytic Efficiency ($sec^{-1}M^{-1}$)	3.9×10^2
Stability at 30° C, pH 7.4 ($t_{1/2}$d)	>14
Stability at 60° C (pH, $t_{1/2}$min)	8.2, >120

Table 2.2 – Summary of the properties of the housekeeping HMB-synthase. [a.]The calculated mass is for the apoenzyme containing the covalently bound dipyrrole cofactor. [b.]Determined for the purified enzymes by mass spectrometry. [c.]Apparent protein concentrations for a 1 mg/ml solution as determined by amino acid analysis.

Overexpression and Purification of recombinant human URO-synthase - URO-synthase was purified to homogeneity in 3 steps with about 50 % yield using the pET28/Smt3 expression system. Table 2.3 summarizes a typical purification profile for the recombinant human enzyme from a 1.5 L culture of *E. coli* strain BL21Codon Plus transformed with the pET28-Smt3-URO-synthase, a vector that produces a fusion polypeptide of URO-synthase and Smt3. The Smt3 fusion peptide enhanced the solubility of URO-synthase and after removal by the Ulp1 protease the recombinant human URO-synthase enzyme was identical to the wild-type sequence with the addition of an amino-terminal serine.

Purification	**Vol**	**URO-synthase Activity**		**Protein**		**Purification**		
(Step)	(ml)	(U/ml)	(Units)	(mg/ml	(mg)	(U/mg)	(Fold)	(%)
Crude Lysate	35.0	5.854E+06	2.049E+08	17.00	600.0	341,459	1.0	100.0
Nickel Affinity	78.0	2.184E+06	1.703E+08	1.70	132.6	1284586	3.8	83.1
Mono Q	22.5	4.866E+06	1.095E+08	1.20	27.0	4,055,389	11.9	53.4
Superdex 70	23.0	4.298E+06	9.886E+07	1.00	23.0	4,298,092	12.6	48.3

Table 2.3 - Typical Purification of Recombinant Human URO-synthase from 1.5L of culture. The specific activities of the fusion protein in the first two steps were corrected to subtract the contribution of the 11.3 kDa Smt3 peptide. Cleavage of the fusion protein was carried out after the nickel affinity purification step using the Ulp1 protease.

Enzyme purity after each step was evaluated by SDS PAGE (Figure 2.6). Compared to the presence of the *E. coli* proteins, as seen in crude extracts (lane 2), most bacterial proteins were excluded after chelation affinity purification (lane 3). Optimal IMAC conditions were 50 mM for washing step, and 250 mM for the elution step, in a 20 mM sodium phosphate buffer containing 500 mM NaCl, 5 mM TCEP, pH 7.8. From 1.5 litter of cell culture in SuperBroth Medium, 132 mg of partially pure URO-synthase could be recovered from the column. The Smt3 tag was cleaved with 1μg of Ulp1 protease per mg of URO-synthase (16 hours digestion at 4^{o} C), and separated from URO-synthase with a high resolution Mono Q chromatographic step, using similar conditions to the HMB-synthase ion exchange chromatography (lane 4). The enzyme eluted at about 42% of buffer B (84 mM NaCl) and totally separated from the Smt3 peptide, which eluted at a higher salt concentration (76% of buffer B, 154 mM NaCl). Although The Mono Q chromatography resulted in only 53% recovery, the enzyme was nearly homogeneous after this 3.1-fold purification step. Subsequent size selection by gel filtration increased purity by only 1.06-fold, but resulted in a homogeneous enzyme protein (Table 2.3 and Figure 2.6). 23 mg of pure protein could be recovered from 1.5 liters of culture, with 12.6 fold purification and 48% yield.

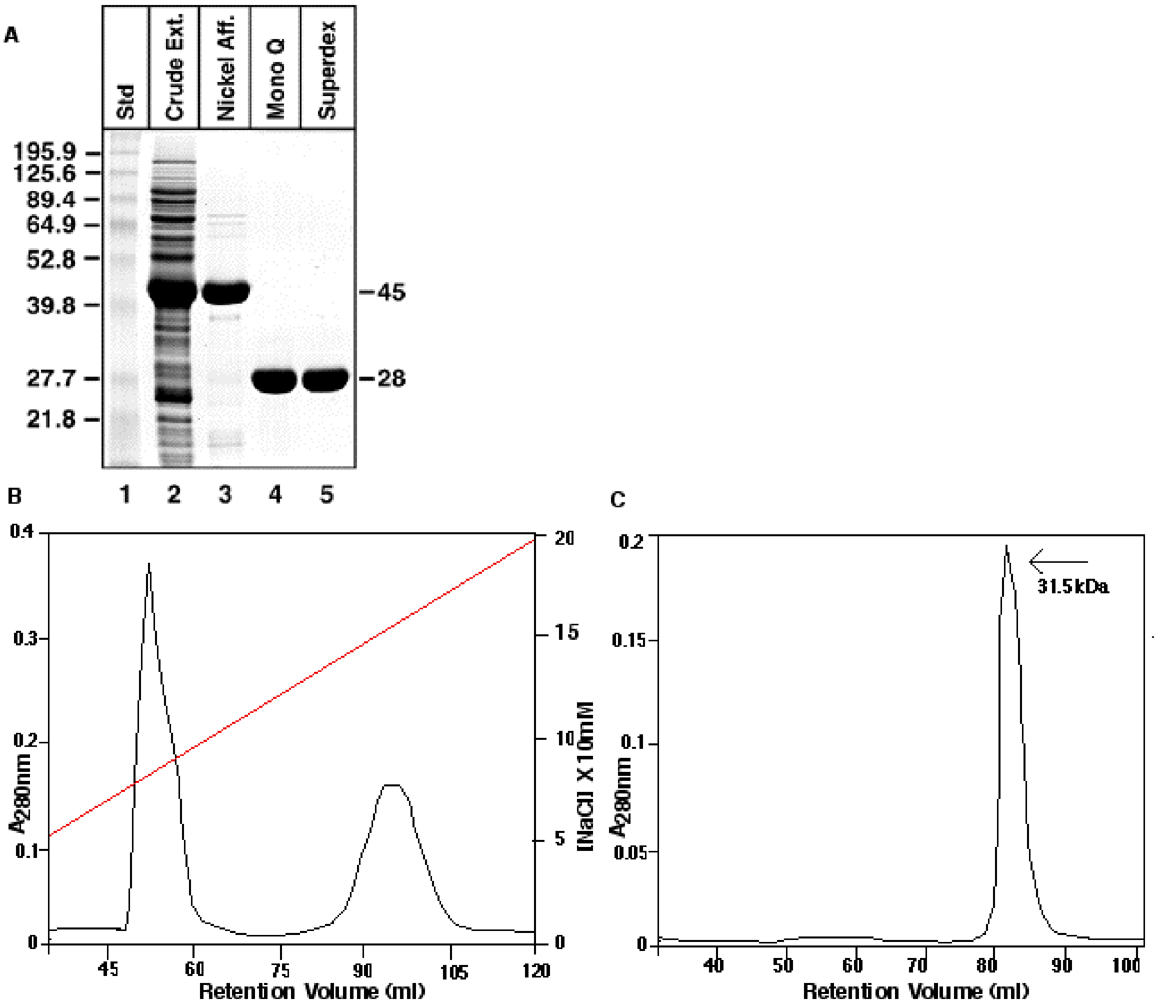


Figure 2.6 – Protein purification of URO-synthase. **A,** SDS-PAGE, documenting protein expression and purification of H-HMB-synthase: Lane 1 – LMW marker; lane 2 – crude lysate; lane 3 – IMAC chromatography; lane 4 – ion exchange; lane 5 - GF**. B** Mono Q ion-exchange chromatogram showing URO-synthase (first peak) after cleavage of the Smt3 tag (second peak). **D,** Gel filtration chromatogram of the purified URO-synthase run on a Superdex-70 (16/60) column (Amersham Biosciences). Note that URO-synthase exists as a monomer in 20 mM phosphate buffer of pH 7.4, containing 150 mM NaCl and 2 mM DTT. (Presence of DTT in the buffer is essential to maintain the monomer conformation, data not shown).

Properties of recombinant human URO-synthase

The molecular weight was accurately determined by mass spectroscopy. The enzyme, after cleavage of the Smt3 tag with Ulp1 protease, had a molecular weight of 28808 kDa (Fig. 2.7). To determine the physical and biochemical properties of purified URO-synthase, the time course of the URO-synthase activity was determined first. The enzymatic assay was modified from the previously described coupled assay [59], in which we replaced the HMB-generating blood lysate with pure, recombinant HMB-synthase. The production of uroporphyrinogen III was in linear proportion to incubation period within 30 minutes, using 5 U of HMB-synthase, 50 μl of 3.4 mM PBG and a certain amount of URO-synthase in the assay (data not shown). The pH effect on the enzyme revealed that the optimal pH of URO-synthase was 8.2 (enzyme coupled assay) and the enzyme showed only residual activity below pH 6.5 (Fig. 2.8). The permit comparison of protein determinations between several methods, these were referenced to a 1 mg/ml protein solution determined by amino-acid composition, as shown in table 2.4.

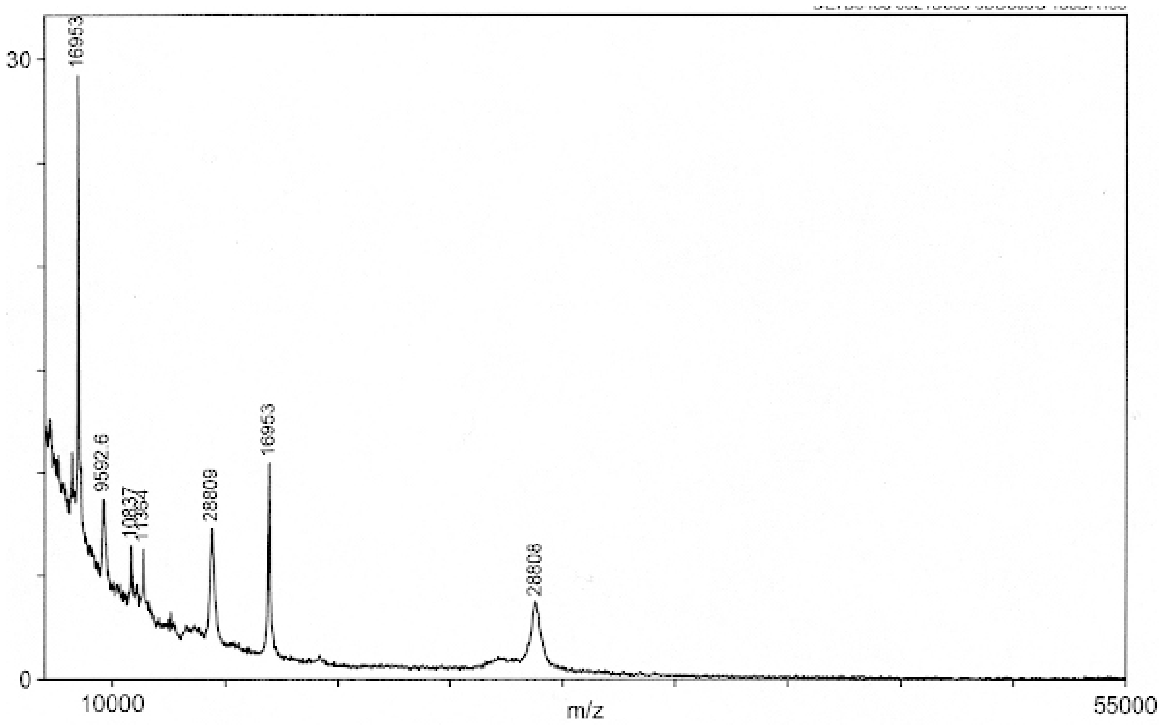


Figure 2.7 – Mass spectrometry profile of URO-synthase. The enzyme has a molecular of 28808 kDa.

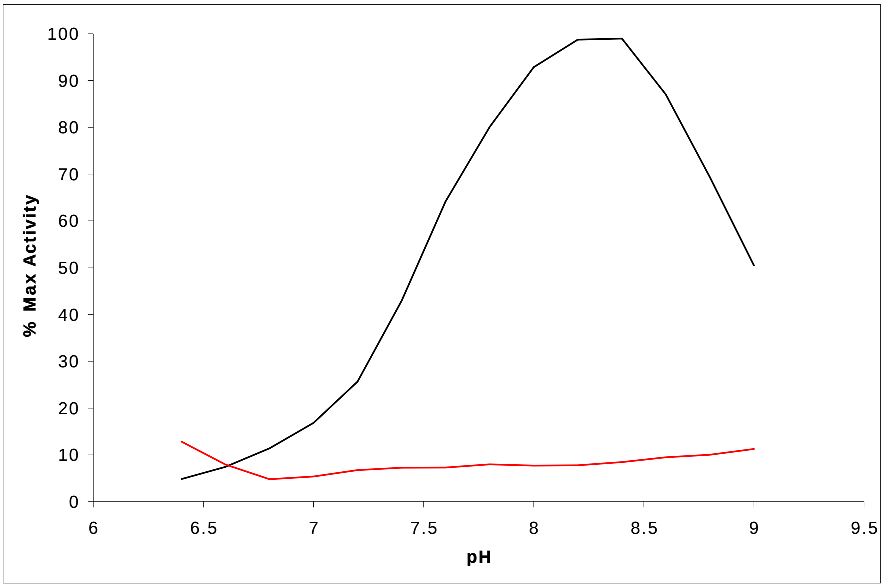


Figure 2.8 – Optimal pH of URO-synthase activity. The amount of HMB-synthase in the assay was adjusted according to its optimal activity, to generate a constant about of HMB-synthase in every assay at each pH (the red line represents the units of HMB-synthase in the assay). URO-synthase has an optimal pH value of 8.2

Thermostability of recombinant human URO-synthase - Thermostability studies revealed that the enzyme was rapidly inactivated at 60° C with a half life of about 2 min. URO-synthase proved to be less stable at pH 8.2 than pH 7.4. At pH 8.2 the half life was 11 hours at 37^{o} C and 5 hours at 45° C. However, the enzyme is quite thermostable at pH 7.4, with a half life of 45 hours at 45° C. At 37^{o} C the enzyme retained full activity for the assay period of 72 hours, as it did for 14 days at 30^{o} C (data not sown). At 4^{o} C the enzyme retained 100% of the initial activity after a period of 6 months (data not shown). For long term storage, we evaluated the stability of the enzyme at -80° C. URO-synthase was stable in for a period longer than 2 years (15% glycerol was added to prevent precipitation of the enzyme upon freezing).

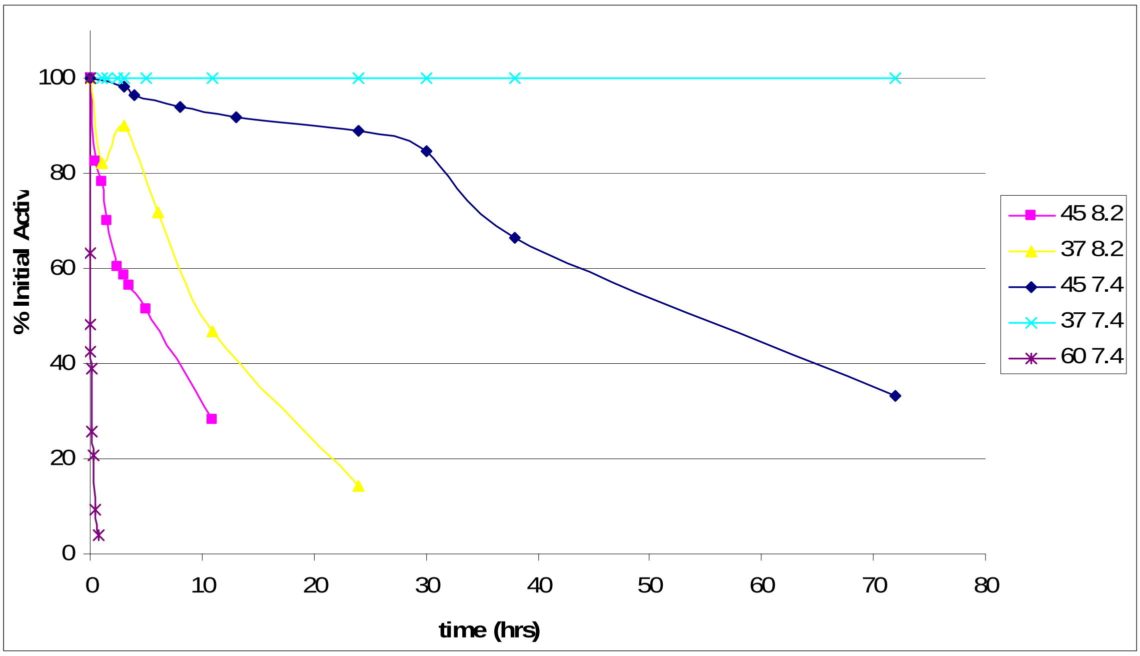


Figure 2.9 – Thermostability profile of recombinant human URO-synthase at pH 7.4 (20 mM $NaPO_4$, 100 mM NaCl, 2 mM DTT), and 8.2 (50 mM Tris, 100 mM NaCl, 2 mM DTT). The half-life at pH 8.2 was 11 hours at 37^{o} C, 5 hours at 45^{o} C. At pH 7.4 the half-life was about 45 hours at 45^{o} C and 2 min at 60^{o} C. At this pH, the enzyme retained 100% activity for a minimal period of 72 hours at 37^{o} C, 15 days at 30^{o} C and 6 months at 4^{o} C (data not shown).

Kinetic properties of recombinant human URO-synthase – The coupled-assay was modified to determine the apparent K_m of recombinant URO-synthase as described by Tsai, 1987 [59], with the following modification: Pure recombinant HMB-synthase was diluted in 0.1 M Hepes, pH 8.2 to obtain a series of concentrations ranging from 0.5-5.0 U/ml HMB-synthase. An aliquot of each dilution was pre-warmed at 37° C for 5 min, and then 50 μl of 3.4 mM PBG hydrate was added at zero time. The mixture was pre-incubated for 13 min to obtain a steady-state HMB levels, at which time, 10 μl of purified URO-synthase were added. The reaction mixture was incubated for an additional 15 min and was terminated by the addition of 440 μl of 10% TCA (w/v) containing 0.5% iodine (w/v). In steady-state kinetics, the km of URO-synthase was 0.14 μM (Figure 2.10) and the V_{max} was 4.5 x 10^6 nmol/hr/mg.

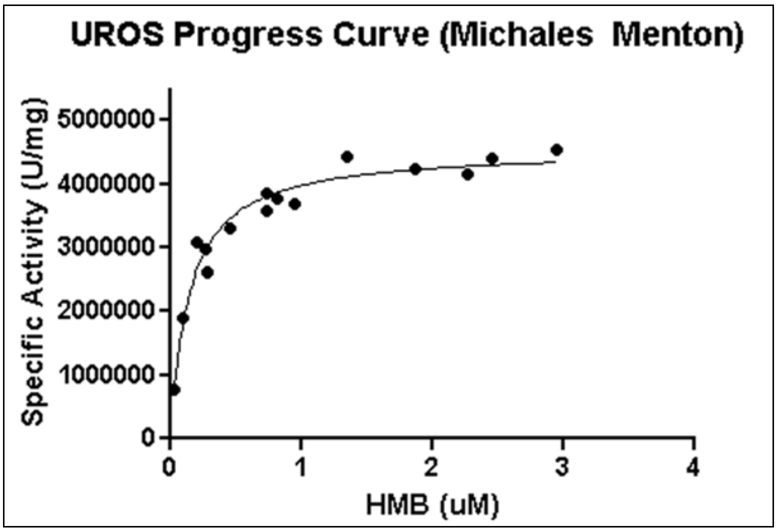


Figure 2.10 – Plot of the reaction velocity of URO-synthase as a function of the substrate concentration. The coupled-assay was modified to determine the apparent K_m of recombinant URO-synthase as described by Tsai, 1987 [59], with modifications. The enzyme has a km of 0.14 μM and a V_{max} of 4.5 x 10^6 nmol/hr/mg.

Property	URO-Synthase
Mr (Theoretical, daltons)	28,712
Mr (Mass Spec, daltons)[a]	28,808
Protein (Bradford, mg/ml)[b]	0.90 ± 0.02
Protein (DC Lowry, mg/ml)[b]	1.43 ± 0.09
Protein (Fluram, mg/ml)[b]	0.86 ± 0.01
Protein (A_{280}, mg/ml)[b]	1.32 ± 0.01
pH Optimum	8.2
K_m (μM)	0.15
K_{cat} (min^{-1})	2,240
Catalytic Efficiency ($sec^{-1}M^{-1}$)	2.5 x 10^8
Stability at 30° C, pH 7.4 ($t_{1/2}$d)	>14
Stability at 60° C (pH, $t_{1/2}$min)	8.2, <1

Table 2.4 – Summary of the properties of the human recombinant URO-synthase. [a.]Determined for the purified enzymes by mass spectrometry. [b.]Apparent protein concentrations for a 1 mg/ml solution as determined by amino acid analysis.

Overexpression and Purification of recombinant human URO-decarboxylase - Recombinant human URO-decarboxylase was purified to homogeneity in 3 steps with about 45 % yield using the pThioHis expression system. Table 2.5 summarizes a typical purification profile of the recombinant human enzyme from a 4 L culture of *E. coli* strain Top10 transformed with the Thio-URO-synthase construct, a thioredoxin fusion protein vector containing the URO-decarboxylase cDNA. The Thioredoxin fusion peptide enhanced the solubility of URO-decarboxylase and after removal by the thrombin protease the r-URO-decarboxylase was identical to the wild-type sequence with the addition of an amino-terminal glycine and serine. Enzyme purity after each step was evaluated by SDS-PAGE (Figure 2.10). Compared to the pET28-Smt3 and pET16 systems used for the expression and purification of URO-synthase and HMSB respectively, the thioredoxin fusion system did not provide an equally efficient purification step by metal affinity chromatography (lane 1). However, the thioredoxin fusion was essential for obtaining soluble enzyme expression in *E. coli*, and further purification by ion exchange (lane 2) and gel filtration (lane 3) provided a homogenous protein sample. Mono Q chromatography resulted provided a 2-fold purification step and a subsequent size selection by gel filtration resulted in a homogeneous protein sample (Table 2.3 and Figure 2.10).

Purification	**Vol**	**URO-decarboxylase Activity**		**Protein**		**Purification**		
(Step)	(ml)	(U/ml)	(Units)	(mg/ml)	(mg)	(U/mg)	(Fold)	(%)
Crude Lysate*	200	9510	1902000	12.68	2536	750	1	100
Nickel Affinity	80	22400	1790000	6.8	544	3290	4.39	94
Mono Q	145	7000	1022000	0.99	144	7100	9.47	53.7
Superdex 70	80	10600	851000	1.4	112	7600	10.1	44.7

Table 2.5 - Typical purification of recombinant human URO-decarboxylase. Cleavage of the fusion protein was carried out after the nickel affinity purification step using the thrombin.

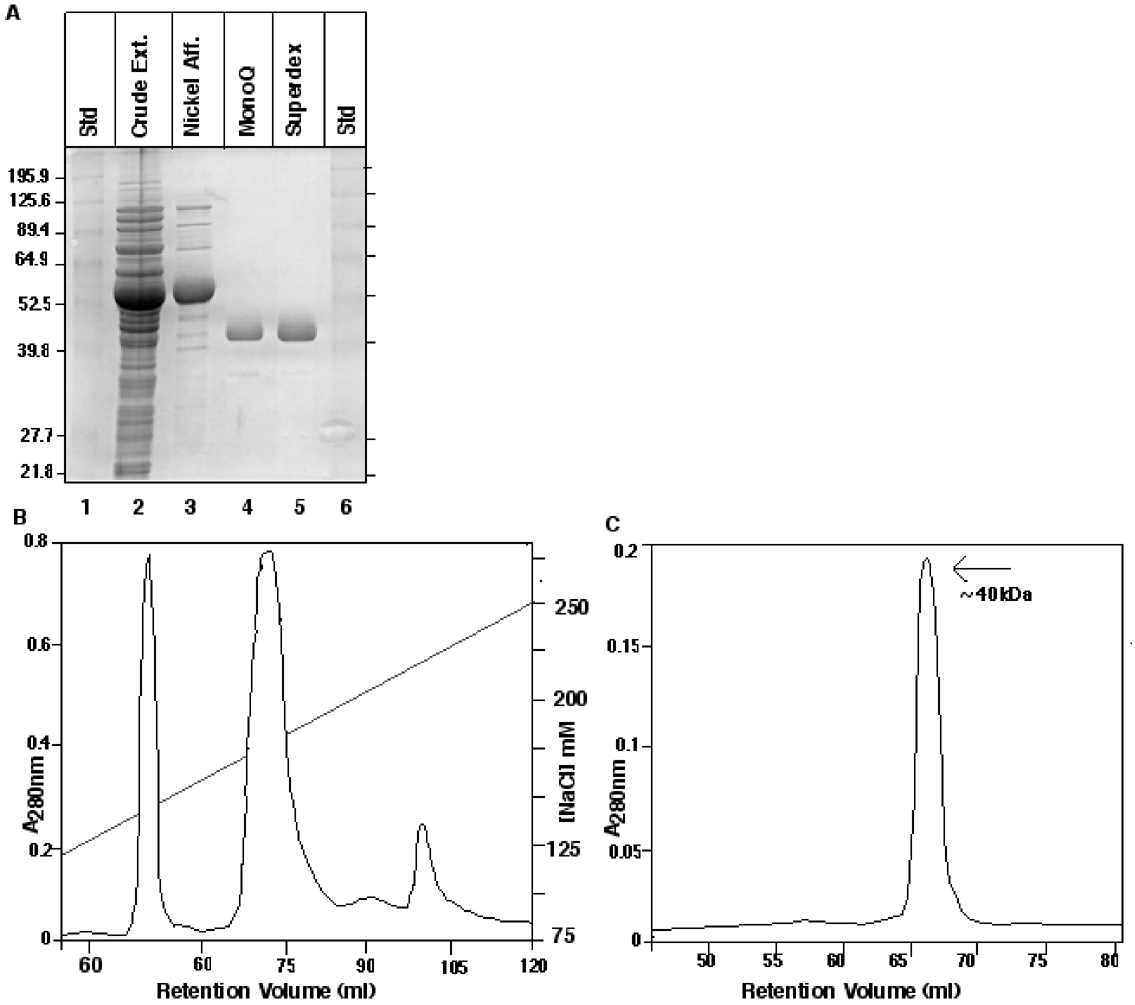


Figure 2.11 – Protein purification of URO-decarboxylase. **A** SDS-PAGE, documenting protein expression and purification of URO-decarboxylase: Lane 1 – LMW marker; lane 2 – crude lysate; lane 3 – IMAC chromatography; lane 4 – ion exchange; lane 5 –GF; lane 6 – LMW marker**.** **B** Mono Q ion-exchange chromatogram showing URO-decarboxylase (second peak) after cleavage of the thioredoxin fusion (first peak). **C** Gel filtration chromatogram of the purified URO-synthase run on a Superdex-70 (16/60) column (Amersham Biosciences). Note that URO-synthase exists as a monomer in 20 mM phosphate buffer of pH 7.4, containing 150 mM NaCl and 2 mM DTT. (Presence of DTT in the buffer is essential to maintain the monomer conformation, data not shown).

Properties of URO-decarboxylase

After cleavage of the thioredoxin-URO-decarboxylase fusion with thrombin, the enzyme has a molecular weight of 41072 kDa (Fig. 2.12), as determined by mass spectroscopy. To determine the physical and biochemical properties of purified URO-decarboxylase, the time course of the URO-decarboxylase activity was determined first, as well as the optimal amount of URO-decarboxylase to be used. The enzymatic assay can be divided in three steps. The first step is to form the porphyrinogen substrate for the assay. The second step is the reaction of the substrate with the source of URO-decarboxylase to be assayed. The third step is the separation and quantification of the reaction products. We generated the porphobilinogen enzymatically, with pure HMB-synthase and PBG as substrate. A forty-minute incubation of 6 units of HMB-synthase with 2.5 nano moles of PBG was sufficient to 20 μM URO'gen I in the reaction mix. The reaction was stopped by adjusting the pH to 6.8, which irreversibly inactivated HMB-synthase. 300 μl of this reaction was added to 200 μl of buffer containing the URO-decarboxylase to be assayed. The optimal amount of pure URO-decarboxylase to be assayed was determined to be 1-3 μg of protein. The reaction was incubated at 37^{o} C for 30 min and stopped with 3 M HCl. Both the porphobilinogen generating and the decarboxylation reaction were flushed with nitrogen and performed in the dark, working under a red safety light, to prevent the oxidation of the porphyrinogens. After the assay was terminated, the samples were exposed to UV light for 20 min, to oxidize the porphyrinogens to porphyrins. The different decarboxylation products formed were separated on a HPLC system using a 3 micron particle size BDS Hypersil column from Thermo Hypersil-Keystone (Bellefonte, PA). The aqueous solvent was 50% (v/v) methanol/NaH_2PO_4, pH 4.5; the organic phase was 100% methanol. The HPLC column was calibrated with a chromatographic marker kit standard that contains 62.5 pmol each of 8-, 7-, 6-, 5-, 4-, and 2-carboxylporphyrins. To determine the optimal pH for URO-decarboxylase enzymatic activity, we adjusted the amount of HMB-synthase used in the URO I generating step of the assay, in order to obtain a range of substrate concentration from 0.05 to 10 μmoles. The enzyme reached maximal activity at pH 6.8. Below pH 5 the enzyme had a residual activity inferior than 25% of the maximal value (Fig. 2.13).

Thermostability of recombinant human URO-decarboxylase – Fig. 2.14 shows the thermostability profile of URO-decarboxylase in 20 mM $NaPO_4$, 100 mM NaCl, 2 mM DTT, pH 6.8. The enzyme was rapidly inactivated at 60° C and 42° C (t1/2 <1and 5 min, respectively, pH 6.8). At 37° C, the half-life was only 11 min. At 4° C the enzyme retained full activity after the length of the assay. At 30° C (pH 7.4), the enzyme showed a half-life of 1 week (data not shown). For long term storage, we evaluated the stability of the enzyme at -80° C. URO-decarboxylase retained 100% activity for a period of at least two months (15% glycerol was added to prevent precipitation of the enzyme upon freezing). The properties of recombinant human URO-decarboxylase are summarized in table 2.6, including the comparison of several protein determination methods.

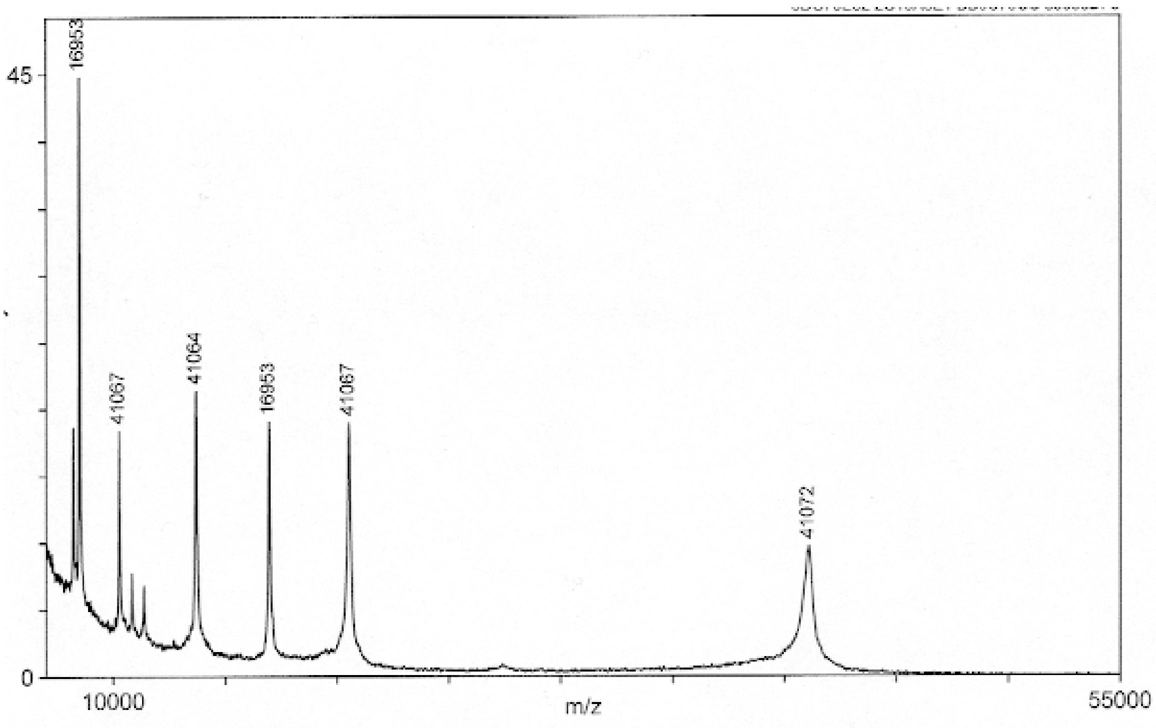


Figure 2.12 – Mass spectrometry profile of URO-decarboxylase. The enzyme has a molecular of about 41072 kDa.

Kinetic properties of URO-decarboxylase – Due to technical difficulties inherent to the assay itself and the available HPLC equipment, the kinetic properties of URO-decarboxylase were not determined.

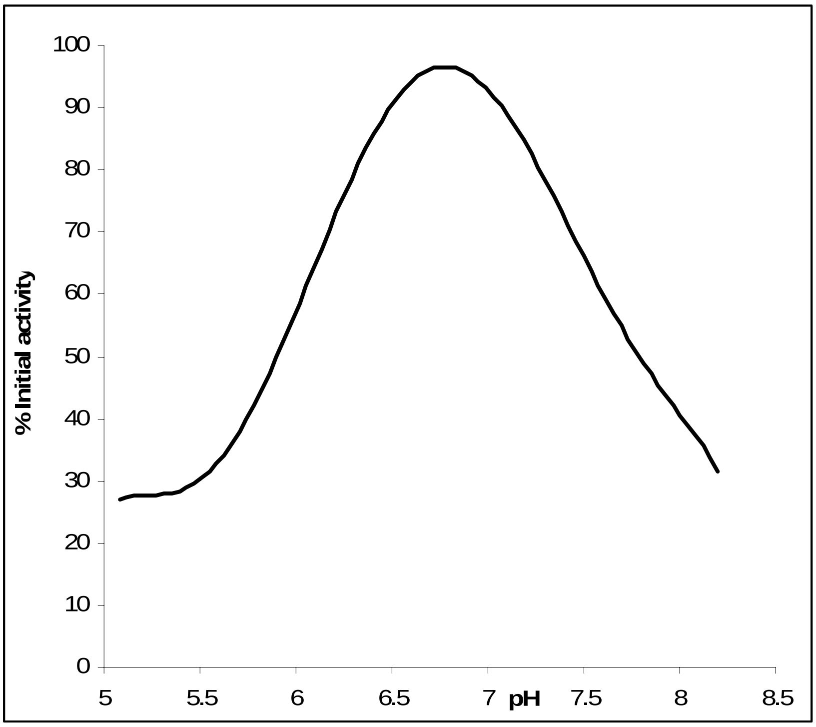


Figure 2.13 – Optimal pH of URO-decarboxylase activity. URO-decarboxylase showed maximum activity at pH 6.8.

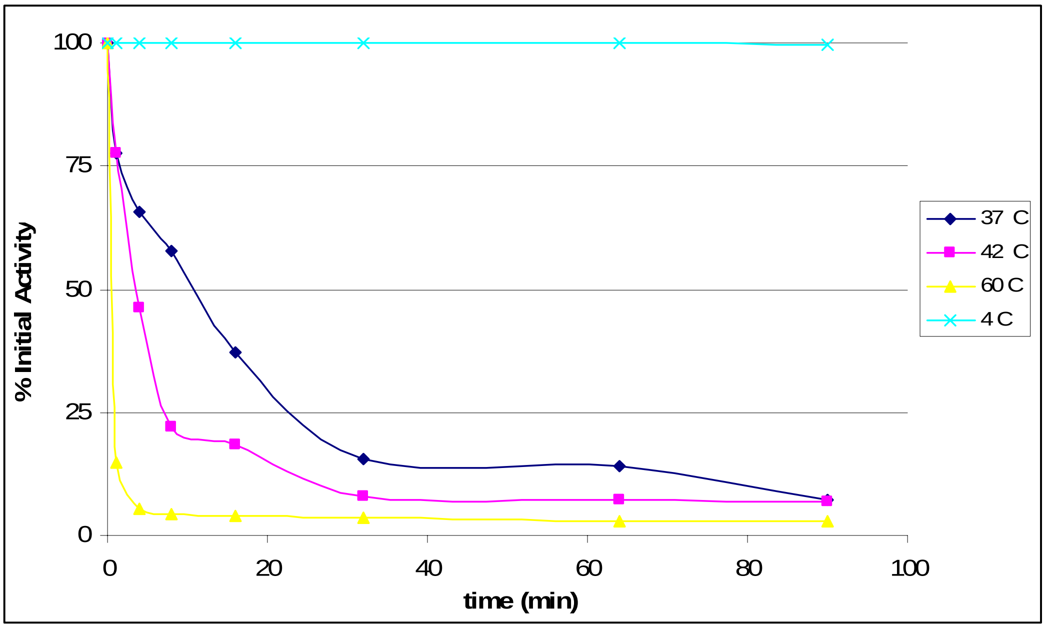


Figure 2.14 – Thermostability profile of recombinant human URO-decarboxylase. The half-lives were 11 min at 37° C, 4.5 min at 42° C and less than 1 min at 60° C.

Property	URO-Decarboxylase
Mr (Theoretical, daltons)	41,066
Mr (Mass Spec, daltons)[a]	41,072
Protein (Bradford, mg/ml)b	1.17 ± 0.05
Protein (DC Lowry, mg/ml)b	1.23 ± 0.12
Protein (Fluram, mg/ml)b	0.62 ± 0.05
Protein (A_{280}, mg/ml)[cb]	1.02 ± 0.01
pH Optimum	6.8
Stability at 30° C, pH 7.4 ($t_{1/2}$d)	>7
Stability at 60° C (pH, $t_{1/2}$min)	6.8, <1

Table 2.6 – Summary of the properties of the human recombinant URO-decarboxylase. [a] Determined for the purified enzymes by mass spectrometry. [b.] Apparent protein concentrations for a 1 mg/ml solution as determined by amino acid analysis.

Overexpression and Purification of *recombinant Ulp1*

Recombinant Ulp1 was produced in our lab as a source of protease for cleavage of the fusion between the Smt3 tag and URO-synthase. The enzyme was expressed in *E. coli* using the pET-16b vector, and purified in one step by immobilized metal affinity chromatography, as described for HMB-synthase. The removal of the His•Tag™ was not required for proteolytic activity. The IMAC eluate was concentrated to 2mg/ml and stored at -140^{o} C in 10% glycerol until further use. Note that storage at -135^{o} C was required to maintain activity for long periods of time. Enzyme storage at -80^{o} C resulted in inactivation within two months.

DISCUSSION

In this chapter we describe a simple three-step purification method to purify three of the four cytosolic human heme biosynthetic enzymes, HMB-synthase, URO-synthase and URO-decarboxylase, expressed in *E. coli*. Erythroid and housekeeping HMB-synthase enzymes were subcloned into pET16b. URO-synthase was subcloned into a modified pET28 vector that directed the expression of a fusion of URO-synthase with the Smt3 peptide, and URO-decarboxylase was subcloned into a modified pThioHis vector. *E. coli* strain BL21-CodonPlus(DE3)-RIL-X was used for the expression experiments. This host contains extra copies of *argU, ileW, leuY* and *proL* tRNAs that recognize rare arginine, isoleucine, leucine and proline codons and was used to overcome problems due to codon bias. Each vector was chosen and optimized for each protein individually. The most basic construct pET16b-HMB-synthase, expressed a protein with a stretch of leader sequence at its N-terminus. The leader sequence contains a segment of ten consecutive histidines, and a Factor Xa recognition sequence. The total length of this stretch is 21 amino acids and the ten histidine tag was used for an efficient metal affinity chromatography first step of purification. When the fusion protein was cleaved with Factor Xa, only one extra amino acid, a histidine, was left over in front of the first methionine. Initial experiments with a pET16b-URO-synthase construct allowed efficient expression and purification; however it was not possible to remove the 21 amino acid tag with Factor Xa. This was probably due to structural constraints which likely resulted in the Factor Xa recognition sequence being inaccessible to the protease.

One of our major concerns in using fusion constructs was being able to produce a protein sequence the closest to the native protein possible. For this reason, we expressed URO-synthase as a fusion to the Smt3 peptide. This fusion also had a leader N-terminal sequence containing ten histidines that allowed for a metal affinity purification step. Smt3, the yeast homologue of SUMO (Small Ubiquitin-like Modifier) belongs to the ubiquitin-like (UBL) family of proteins. It is processed to its mature form by Ulp1 protease. In this construct, the URO-synthase sequence was cloned adjacent to the Ulp1 recognition site. When URO-synthase was cleaved from Smt3 with Ulp1 protease, only a serine was left in front of the first methionine. An advantage of using this system is that Ulp1 is easily made in our lab, expressed with the pET16b system and purified in one step by metal affinity chromatography, without the need to remove the histidine tag for proteolytic activity. This protease is active over a wide range of pH and buffer conditions.

URO-decarboxylase was cloned into the pThioHis expression system and expressed a fusion to thioredoxin. The thioredoxin sequence was modified by the vendor (Invitrogen, Carlsbad, CA) to include a series of mutations that code for several histidines scattered throughout the primary sequence. In the folded thioredoxin, these histidines form a patch that binds to chelating sepharose resin. This vector was modified from its original version purchased from Invitrogen (Carlsbad, CA), to replace the enterokinase recognition site by a thrombin site. The fusion with thioredoxin allowed for increased solubility of URO-decarboxylase during expression in *E. coli*. After cleavage, URO-decarboxylase was left with two extra amino acids, a glycine and a serine. Overall, The IMAC (immobilized metal affinity chromatography) system permitted for an efficient first step of purification. Thioreducing agents like DTT or β-mercaptanol irreversibly destroy the resin, and therefore we included TCEP in the lysis and IMAC buffers, which do not reduce the metal but do keep the enzyme thiols reduced.

In addition to the metal affinity chromatography, we further purified the three enzymes by ion exchange chromatography in a high-resolution MonoQ column, followed by gel filtration in a Superdex 75 column. The ion exchange chromatography step allowed for the separation of most remaining contaminating *E. coli* proteins and the separation of URO-synthase from Smt3 and URO-decarboxylase from thioredoxin. It also allowed for the separation of the HMB-synthase-substrate intermediates. Previously, the ion exchange separation of HMB-synthase and URO-synthase was described using a low strength phosphate buffer system [50]. Due to competition of the phosphate with the protein for resin binding sites, a very low concentration of potassium phosphate was used. A major disadvantage of this buffer was that both HMB-synthase and URO-synthase showed reduced solubility in low salt buffers. To circumvent this problem, we used a HEPES buffer, which helped maintain the protein stability, and did not bind the matrix. The concentration of sodium chloride in the elution buffer was adjusted for each protein sample (200 mM for HMB-synthase and URO-synthase, and 250 mM for URO-decarboxylase). Finally, a gel filtration step in a Superdex 75 column yielded a homogenous protein preparation, by excluding any high and/or low molecular contaminants. Contrary to previous experience in this lab (personal communication), conditions were determined to maintain these proteins stable for long periods of time, even at protein concentrations as high as 80mg/ml. Precise buffer conditions were essential. The enzyme was most stable in 20 mM sodium phosphate, 200 mM sodium chloride, and 2 mM DTT, pH 7.4), although a HEPES buffer and 100 mM ionic strength was well tolerated. We determined the physicokinetic properties of the purified recombinant HMB-synthase, URO-synthase and URO-decarboxylase. Notably, these enzymes had similar properties to those described for the native enzyme except for a higher specific activity of URO-synthase, and a slightly higher km value of HMB-synthase [50, 37, 64]. The housekeeping and erythroid HMB-synthase isozymes had similar properties, and behaved similarly during the various purification steps. For that reason we present only the purification of housekeeping HMB-synthase. It was previously accepted that URO-synthase has very low heat stability. This was probably due to the thermostability assay being done at an inappropriate pH (e.g. pH 8.2) [46]. At pH 7.4, this enzyme was quite stable and maintained its full activity for a minimal period of 6 months at 4^o C. At 30^o C the enzyme was stable for period greater than two weeks. HMB-synthase was stable for a similar amount of time, and URO-decarboxylase had a half life of 1 week. This was an important requirement for the NMR studies described in the next chapter.

BIBIOGRAGHY

1. Wetmur J.G., Bishop D.F., Cantelmo C., Desnick R.J., 1986, Human delta-aminolevulinate dehydratase: nucleotide sequence of a full-length cDNA clone. Proc Natl Acad Sci U S A, 83(20): p. 7703-7.
2. Wetmur J.G., Bishop D.F., Ostasiewicz L., Desnick R.J., 1986., Molecular cloning of a cDNA for human delta-aminolevulinate dehydratase. Gene, 43(1-2): p. 123-30.
3. Raich N., Romeo P.H., Dubart A., Beaupain D., Cohen-Solal M., Goossens M., 1986, Molecular cloning and complete primary sequence of human erythrocyte porphobilinogen deaminase. Nucleic Acids Res, 14(15): p. 5955-68.
4. Romeo P.H., Raich N., Dubart A., Beaupain D., Pryor M., Kushner J., Cohen-Solal M., Goossens M., 1986, Molecular cloning and nucleotide sequence of a complete human uroporphyrinogen decarboxylase cDNA. J Biol Chem, 261(21): p. 9825-31.
5. Bawden M.J., Borthwick I.A., Healy H.M., Morris C.P., May B.K., Elliott W.H., 1987, Sequence of human 5-aminolevulinate synthase cDNA. Nucleic Acids Res, 15(20): p. 8563.
6. Grandchamp B, De Verneuil H, Beaumont C, Chretien S, Walter O, Nordmann Y., 1987, Tissue-specific expression of porphobilinogen deaminase. Two isoenzymes from a single gene. Eur J Biochem, 162(1): p. 105-10.
7. Bishop, D.F., 1990, Two different genes encode delta-aminolevulinate synthase in humans: nucleotide sequences of cDNAs for the housekeeping and erythroid genes. Nucleic Acids Res, 18(23): p. 7187-8.
8. Martasek P., Camadro J.M., Delfau-Larue M.H., Dumas J.B., Montagne J.J., de Verneuil H., Labbe P., Grandchamp B., 1994, Molecular cloning, sequencing, and functional expression of a cDNA encoding human coproporphyrinogen oxidase. Proc Natl Acad Sci U S A, 91(8): p. 3024-8.
9. Nishimura, K., S. Taketani, and H. Inokuchi, 1995, Cloning of a human cDNA for protoporphyrinogen oxidase by complementation in vivo of a hemG mutant of Escherichia coli. J Biol Chem, 270(14): p. 8076-80.
10. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1988, Human uroporphyrinogen III synthase: molecular cloning, nucleotide sequence, and expression of a full-length cDNA. Proc Natl Acad Sci U S A, 85(19): p. 7049-53.
11. Taketani S., Kohno H., Furukawa T., Yoshinaga T., Tokunaga R., 1994, Molecular cloning, sequencing and expression of cDNA encoding human coproporphyrinogen oxidase. Biochim Biophys Acta, 1183(3): p. 547-9.

12. Conboy J.G., Cox T.C., Bottomley S.S., Bawden MJ, May BK., 1992, Human erythroid 5-aminolevulinate synthase. Gene structure and species-specific differences in alternative RNA splicing. J Biol Chem, 267(26): p. 18753-8.
13. Kaya A.H., Plewinska M., Wong D.M., Desnick R.J,. Wetmur J.G., 1994, Human delta-aminolevulinate dehydratase (ALAD) gene: structure and alternative splicing of the erythroid and housekeeping mRNAs. Genomics, 19(2): p. 242-8.
14. Yoo H.W., Warner C.A., Chen C.H., Desnick RJ., 1993, Hydroxymethylbilane synthase: complete genomic sequence and amplifiable polymorphisms in the human gene. Genomics,. 15(1): p. 21-9.
15. Romana M., Dubart A., Beaupain D., Chabret C., Goossens M. Romeo P.H., 1987, Structure of the gene for human uroporphyrinogen decarboxylase. Nucleic Acids Res, 15(18): p. 7343-56.
16. Moran-Jimenez M.J., Ged C., Romana M., Enriquez De Salamanca R., Taieb A, Topi G., D'Alessandro L., de Verneuil H., 1996, Uroporphyrinogen decarboxylase: complete human gene sequence and molecular study of three families with hepatoerythropoietic porphyria. Am J Hum Genet, 58(4): p. 712-21.
17. Delfau-Larue, M.H., P. Martasek, and B. Grandchamp, 1994, Coproporphyrinogen oxidase: gene organization and description of a mutation leading to exon 6 skipping. Hum Mol Genet, 3(8): p. 1325-30.
18. Taketani .S, Inazawa J., Abe T., Furukawa T., Kohno H., Tokunaga R., Nishimura K., Inokuchi H., 1995, The human protoporphyrinogen oxidase gene (PPOX): organization and location to chromosome 1. Genomics, 29(3): p. 698-703.
19. Taketani S., Inazawa J., Nakahashi Y., Abe T., Tokunaga R., 1992, Structure of the human ferrochelatase gene. Exon/intron gene organization and location of the gene to chromosome 18. Eur J Biochem, 205(1): p. 217-22.
20. Wang A.L., Arredondo-Vega F.X., Giampietro P.F., Smith M., Anderson W.F., Desnick R.J., 1981, Regional gene assignment of human porphobilinogen deaminase and esterase A4 to chromosome 11q23 leads to 11qter. Proc Natl Acad Sci U S A, 78(9): p. 5734-8.
21. Beaumont C., Foubert C., Grandchamp B., Weil D., Van Cong N'Guyen, Gross M.S.., Nordmann Y., 1984, Assignment of the human gene for delta aminolevulinate dehydrase to chromosome 9 by somatic cell hybridization and specific enzyme immunoassay. Ann Hum Genet, 48 (Pt 2): p. 153-9.
22. de Verneuil H., Grandchamp B., Foubert C., Weil D., N'Guyen V.C., Gross M.S., Sassa S., Nordmann Y., 1984, Assignment of the gene for uroporphyrinogen decarboxylase to human chromosome 1 by somatic cell hybridization and specific enzyme immunoassay. Hum Genet, 66(2-3): p. 202-5.

23. Wang A.L., Astrin K.H., Anderson W.F., Desnick R.J., 1985, delta-Aminolevulinate dehydratase: induced expression and regional assignment of the human gene to chromosome 9q13----qter. Hum Genet, 70(1): p. 6-10.
24. Dubart A., Mattei M.G., Raich N., Beaupain D., Romeo P.H., Mattei J.F., Goossens M., 1986, Assignment of human uroporphyrinogen decarboxylase (URO-D) to the p34 band of chromosome 1. Hum Genet, 73(3): p. 277-9.
25. Potluri V.R., Astrin K.H., Wetmur J.G., Bishop D.F. Desnick R.J., 1987, Human delta-aminolevulinate dehydratase: chromosomal localization to 9q34 by in situ hybridization. Hum Genet, 76(3): p. 236-9.
26. Cotter, P.D., 1992, Assignment of human erythroid delta-aminolevulinate synthase (ALAS2) to a distal subregion of band Xp11.21 by PCR analysis of somatic cell hybrids containing X; autosome translocations. Genomics, 13(1): p. 211-2.
27. Bishop, D.F., A.S. Henderson, and K.H. Astrin, 1990, Human delta-aminolevulinate synthase: assignment of the housekeeping gene to 3p21 and the erythroid-specific gene to the X chromosome. Genomics, 7(2): p. 207-14.
28. Cox T.C., Bawden M.J., Abraham N.G., Bottomley S.S., May B.K., Baker E. Chen L.Z., Sutherland GR., 1990, Erythroid 5-aminolevulinate synthase is located on the X chromosome. Am J Hum Genet, 46(1): p. 107-11.
29. Namba H., Narahara K., Tsuji K., Yokoyama Y., Seino Y., 1991, Assignment of human porphobilinogen deaminase to 11q24.1----q24.2 by in situ hybridization and gene dosage studies. Cytogenet Cell Genet, 57(2-3): p. 105-8.
30. Astrin K.H., Warner C.A., Yoo H.W., Goodfellow P.J., Tsai SF, Desnick R.J., 1991, Regional assignment of the human uroporphyrinogen III synthase (URO-synthase) gene to chromosome 10q25.2----q26.3. Hum Genet, 87(1): p. 18-22.
31. Whitcombe D.M., Carter N.P., Albertson D.G., Smith S.J., Rhodes D.A., Cox T.M., 1991, Assignment of the human ferrochelatase gene (FECH) and a locus for protoporphyria to chromosome 18q22. Genomics, 11(4): p. 1152-4.
32. Brenner D.A., Didier J.M., Frasier F., Christensen S.R., Evans G.A., Dailey H.A., 1992, A molecular defect in human protoporphyria. Am J Hum Genet,. 50(6): p. 1203-10.
33. Cacheux V., Martasek P., Fougerousse F., Delfau M.H., Druart L., Tachdjian .G, Grandchamp B., 1994, Localization of the human coproporphyrinogen oxidase gene to chromosome band 3q12. Hum Genet, 94(5): p. 557-9.
34. Cotter P.D., Drabkin H.A., Varkony T., Smith D.I., Bishop D.F., 1995., Assignment of the human housekeeping delta-aminolevulinate synthase gene (ALAS1) to chromosome band 3p21.1 by PCR analysis of somatic cell hybrids. Cytogenet Cell Genet, 69(3-4): p. 207-8.

35. Roberts A.G., Elder G.H., De Salamanca R.E., Herrero C., Lecha M., Mascaro J.M., 1995, A mutation (G281E) of the human uroporphyrinogen decarboxylase gene causes both hepatoerythropoietic porphyria and overt familial porphyria cutanea tarda: biochemical and genetic studies on Spanish patients. J Invest Dermatol, 104(4): p. 500-2.
36. Chen, C.H., 1994, Acute intermittent porphyria: identification and expression of exonic mutations in the hydroxymethylbilane synthase gene. An initiation codon missense mutation in the housekeeping transcript causes "variant acute intermittent porphyria" with normal expression of the erythroid-specific enzyme. J Clin Invest, 94(5): p. 1927-37.
37. Anderson, P.M. and R.J. Desnick, 1980, Purification and properties of uroporphyrinogen I synthase from human erythrocytes. Identification of stable enzyme-substrate intermediates. J Biol Chem, 255(5): p. 1993-9.
38. Battersby A.R., Fookes C.J., Matcham G.W., McDonald E., 1980, Biosynthesis of the pigments of life: formation of the macrocycle. Nature, 285(5759): p. 17-21.
39. Jordan P.M., Warren M.J., Williams H.J., Stolowich N.J., Roessner C.A., Grant S.K., Scott AI., 1988, Identification of a cysteine residue as the binding site for the dipyrromethane cofactor at the active site of Escherichia coli porphobilinogen deaminase. FEBS Lett, 235(1-2): p. 189-93.
40. Jordan, P.M. and M.J. Warren, 1987, Evidence for a dipyrromethane cofactor at the catalytic site of E. coli porphobilinogen deaminase. FEBS Lett, 225(1-2): p. 87-92.
41. Warren, M.J., Jordan P. M., 1988, Investigation into the nature of substrate binding to the dipyrromethane cofactor of Escherichia coli porphobilinogen deaminase. Biochemistry, 27: p. 9020-9030.
42. Raich N., Mignotte V., Dubart A., Beaupain D., Leboulch P., Romana M., Chabret C., Charnay P., Papayannopoulou T., Goossens M., 1989., Regulated expression of the overlapping ubiquitous and erythroid transcription units of the human porphobilinogen deaminase (PBG-D) gene introduced into non-erythroid and erythroid cells. J Biol Chem, 264(17): p. 10186-92.
43. Porcher C., Picat C., Daegelen D., Beaumont C., Grandchamp B., 1995, Functional analysis of DNase-I hypersensitive sites at the mouse porphobilinogen deaminase gene locus. Different requirements for position-independent expression from its two promoters. J Biol Chem, 270(29): p. 17368-74.
44. Mignotte V., Eleouet J.F., Raich N., Romeo P.H., 1989, Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. Proc Natl Acad Sci U S A, 86(17): p. 6548-52.

45. Brownlie P.D., Lambert R., Louie G.V., Jordan P.M., Blundell T.L., Warren M.J., Cooper J.B., Wood S.P., 1994, The three-dimensional structures of mutants of porphobilinogen deaminase: toward an understanding of the structural basis of acute intermittent porphyria. Protein Sci, 3(10): p. 1644-50.
46. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1987, Purification and properties of uroporphyrinogen III synthase from human erythrocytes. J Biol Chem, 262(3): p. 1268-73.
47. Aizencang G., Solis C., Bishop D.F., Warner C., Desnick R.J., 2000., Human uroporphyrinogen-III synthase: genomic organization, alternative promoters, and erythroid-specific expression. Genomics, 70(2): p. 223-31.
48. Mathews M.A., Schubert H.L., Whitby F.G., Alexander K.J., Schadick K., Bergonia H.A., Phillips J.D., Hill C.P., 2001, Crystal structure of human uroporphyrinogen III synthase. Embo J, 20(21): p. 5832-9.
49. de Verneuil, H., S. Sassa, and A. Kappas, 1983, Purification and properties of uroporphyrinogen decarboxylase from human erythrocytes. A single enzyme catalyzing the four sequential decarboxylations of uroporphyrinogens I and III. J Biol Chem, 258(4): p. 2454-60.
50. Lash, T.D., 1991, Action of uroporphyrinogen decarboxylase on uroporphyrinogen-III: a reassessment of the clockwise decarboxylation hypothesis. Biochem J, 278 (Pt 3): p. 901-3.
51. Elder, G.H., J.A. Tovey, and D.M. Sheppard, 1983, Purification of uroporphyrinogen decarboxylase from human erythrocytes. Immunochemical evidence for a single protein with decarboxylase activity in human erythrocytes and liver. Biochem J, 215(1): p. 45-55.
52. Sassa S., de Verneuil H., Anderson K.E., Kappas A., 1983, Purification and properties of human erythrocyte uroporphyrinogen decarboxylase: immunological demonstration of the enzyme defect in porphyria cutanea tarda. Trans Assoc Am Physicians, 96: p. 65-75.
53. de Verneuil, H., S. Sassa, and A. Kappas, 1983, Purification and properties of uroporphyrinogen decarboxylase from human erythrocytes. A single enzyme catalyzing the four sequential decarboxylations of uroporphyrinogens I and III. J Biol Chem, 258(4): p. 2454-60.
54. Romeo P.H., Raich N., Dubart A., Beaupain D., Pryor M., Kushner J., Cohen-Solal M., Goossens M., 1986, Molecular cloning and nucleotide sequence of a complete human uroporphyrinogen decarboxylase cDNA. J Biol Chem, 261(21): p. 9825-31.
55. Whitby F.G., Phillips J.D., Kushner J.P., Hill C.P., 1998, Crystal structure of human uroporphyrinogen decarboxylase. Embo J, 17(9): p. 2463-71.

56. Anderson, P.M. and R.J. Desnick, 1982, Porphobilinogen deaminase: methods and principles of the enzymatic assay. Enzyme, 28(2-3): p. 146-57.
57. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182, 319-26.
58. Cadene, M. and B.T. Chait, 2000, A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins. Anal Chem, 72(22): p. 5655-8.
59. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1987, Coupled-enzyme and direct assays for uroporphyrinogen III synthase activity in human erythrocytes and cultured lymphoblasts. Enzymatic diagnosis of heterozygotes and homozygotes with congenital erythropoietic porphyria. Anal Biochem, 166(1): p. 120-33.
60. Roberts, A. G. & Elder, G. H., 1997, Purification and properties of uroporphyrinogen decarboxylase from human erythrocytes. Methods Enzymol 281, 349-55.
61. Laemmli, U.K., 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227(259): p. 680-5.

CHAPTER 3: HUMAN UROPORPHYRINOGEN III SYNTHASE - NMR RESONANCE ASSIGNMENTS, 3D SOLUTION STRUCTURE, AND MAPPING OF THE ACTIVE SITE; NMR STUDIES DID NOT SUPPORT A HEME ENZYME COMPLEX OF THE CYTOSOLIC HEME BIOSYNTHETIC ENZYMES.

SUMMARY

Uroporphyrinogen III synthase (URO-synthase), the fourth enzyme in the heme biosynthetic pathway, catalyzes the cyclization and D-ring isomerization of the linear tetrapyrrole hydroxymethylbilane (HMB), to form uroporphyrinogen (URO'gen) III, the cyclic tetrapyrrole and physiologic precursor of heme, chlorophyl, and corrin. The deficient activity of URO-synthase results in the autosomal recessive cutaneous disorder, congenital erythropoietic porphyria (CEP), in which URO'gen I accumulates. Mapping of the structural determinants that specify catalysis and, potentially, protein-protein interactions is lacking. To map the active site and the enzyme's possible interaction in a complex with hydroxymethylbilane-synthase (HMB-synthase) or uroporphyrinogen-decarboxylase (URO-decarboxylase) by NMR, the three recombinant human enzymes were expressed and purified in high yield. Using an 800 MHz instrument, assignment of the URO-synthase backbone $^{13}C_{\alpha}$ (100%), $^{1}H_{\alpha}$ (99.6%), and non-proline $^{1}H_{N}$ and ^{15}N resonances (94%) was achieved as well as 85% of the side-chain ^{13}C and ^{1}H resonances. The absence of chemical shift changes in the [^{1}H,^{15}N]-spectrum of URO-synthase mixed with the homogeneous HMB-synthase holoenzyme or URO-decarboxylase precluded occurrence of a stable cytosolic enzyme complex. NMR analyses of URO-synthase titrated with competitive inhibitors N_D-methyl-1-formylbilane (NMF-bilane) or URO'gen III, revealed resonance perturbations of specific residues lining the cleft that mapped the enzyme's active site. *In silico* docking of the URO-synthase crystal structure with NMF-bilane and URO'gen III was consistent with the perturbation results and provided a 3D model of the enzyme-inhibitor complex. Intermolecular NOEs were measured and the 3D solution structure determined. The solution structure of UROS is similar to the previously determined crystal structure, but it has a much smaller cleft between domains.

These studies provide the resonance assignment for human URO-synthase and the 3D solution structure, exclude the interaction of the enzyme with HMB-synthase or URO-decarboxylase in a cytosol complex, and localize the active site.

INTRODUCTION

Enzymes catalyzing multiple reactions in a pathway have been isolated as multifunctional enzyme complexes (e.g. glycolytic pathway [1], pentose phosphate pathway [2]). Such complexes allow efficient catalysis of consecutive reactions in a pathway, and are thus called metabolons [3]. The formation of such enzyme complexes allow for the pathway intermediates to be transferred from one enzyme to the next without complete equilibration with the surrounding medium. This process, called "metabolite channeling" [4], has several advantages: (1) it leads to an increase in the catalytic efficiencies (turnover numbers) of the overall pathway as compared with that observed of the respective, individual unaggregated systems [5-7]; (2) the concentration of the substrate in the microenvironment of the complex is kept relatively high compared to the average bulk concentration, and the limited solvation capacity of the cellular water is avoided [8-12]; (3) when the input to a sequence of metabolic reactions is changed, the time for attainment of a new steady state is reached faster for a multienzyme complex than for comparable enzyme activities in solution; (4) if substrates are unstable in aqueous environments or are shared in more than one pathway, then direct transfer to a successive active site is a mechanism for substrate preservation [8-11]; (5) Interactions between sequential proteins may have allosteric (thus regulatory) effects on the enzymes involved [13]; (6) complexes, if attached to, or part of cellular structural elements, would have the advantage that the diffusion of metabolites could take place in one or two rather than 3 dimensions [14]. Such complexes have been proposed in several biosynthetic pathways (e.g., the pyrimidine biosynthetic pathway, tryptophan biosynthesis, and the thymidylate pathway).

Enzyme complexes range from static to transient. Static complexes have a slow dissociation rate and the proteins involved usually bind strongly. Transient complexes may occur when kinetic evidence of rate enhancement indicate specific protein-protein interactions and yet no stable protein complex can be detected or isolated by traditional methods (such as gel chromatography, fluorescence spectroscopy, chemical cross-linking, and affinity gel chromatography). These type of interactions are likely mediated by the presence of product/substrate (e.g. cytochrome C and cytochrome C peroxidase [15]; ribonucleotide reductase and thioredoxin [16]; and p-aminobenzoate (PABA) synthase [17]).

Nuclear magnetic resonance (NMR), known as a method to solve 3D protein structures without requiring crystallization [18], is very well suited to the study of protein-protein interactions, especially weak interactions that are difficult to detect by other methods. Different NMR methodology can be applied to study such interactions. The

Nuclear Overhauser Effect (NOE) measures interproton distances and is proportional to the r^{-6} distance of two interacting proton nuclei. Using NOEs, a full three-dimensional structure of the complex can be determined [19]. With cross saturation experiments, one is able to quickly gain information regarding the contact interface between two proteins. A donating partner is not labeled, and the observed protein is ^{15}N labeled and perdeuterated, but its amide deuterons are exchanged back to protons. An [^{15}N, ^{1}H]-HSQC experiment measures the cross relaxation from the donor to the acceptor protein amide protons. The cross peaks that change in intensity upon the donor saturation are very likely to be close to the intermolecular interface. If the ^{15}N, ^{1}H resonances have been assigned, and the structure is known, the acceptor side of the interface is known. This process can be reversed to obtain the other interface [20]. A third, widely used method is called chemical shift perturbation mapping. The [^{15}N, ^{1}H]-HSQC spectrum of one ^{15}N-labelled protein is monitored when the unlabelled interacting partner is titrated in, and perturbations in the chemical shifts are recorded. The interaction causes environmental changes on the protein interfaces which affect the chemical shifts of nuclei in this area [20]. As a consequence of allosteric conformational changes resulting from the protein interaction, environmental changes and consequent chemical shift changes may be observed elsewhere in the protein.

The subcellular bi-compartmental organization of the heme biosynthetic enzymes evolved to maintain the highly reduced porphyrinogens in the relatively non-oxidative environment of the cytosol (to protect them from oxidation to the non-physiologic and unmetabolizable porphyrins), while the last oxidative steps of the pathway were effectively catalyzed in the highly oxidative environment of the mitochondria. HMBS and UROS catalyze the sequential conversion of four molecules of PBG to the cyclic porphyrinogen, URO’gen III, in the cytoplasm. HMBS, the third enzyme in the pathway, catalyzes the head to tail condensation of four molecules of PBG, by deamination, to form the linear tetrapyrrole HMB. In the presence of UROS, HMB is rapidly converted to URO’gen III by a reaction mechanism which involves intramolecular rearrangement of the pyrrole D ring and rapid ring closure [21]. In the absence of UROS, HMB is nonenzymatically cyclized to form URO’gen I, an isomer which is non-physiological (see Fig. 2.1). The deficient activity of UROS, which causes congenital erythropoietic porphyria (CEP), results in the accumulation of uroporphyrinogen I (URO I), which then oxidizes to uroporphyrin I, the primary pathological compound in this human disorder [22,23]. UROD is the fifth enzyme in the heme biosynthetic pathway and sequentially decarboxylates the four acetate side chains of uroporphyrinogen III.

In 1958, Bogorad first demonstrated that URO-synthase and HMB-synthase were required to synthesize URO'gen III from porphobilinogen (PBG) [24]. Subsequently, it was hypothesized that HMB-synthase and URO-synthase, which are both required to convert PBG to URO'gen III, physically interacted, perhaps in a complex of the four cytosolic heme biosynthetic enzymes [25-27]. Experimental evidence supporting a physical protein-protein complex was provided by the PBG-induced increase in the sucrose density sedimentation rate of URO-synthase when HMB-synthase was added [26] and by the binding of URO-synthase to Sepharose-bound HMB-synthase [27]. Kinetic evidence for a close association was provided by allosterism of URO'gen III formation from PBG [28], a decrease in the HMB-synthase K_m for PBG upon addition of URO-synthase [29], and a decrease in the URO-synthase K_m for HMB when it was generated by co-incubation with HMB-synthase instead of being added exogenously [30]. To date, however, neither the interaction of the four enzymes, nor the interaction of HMB-synthase and URO-synthase has been proven.

This concept could be expanded as it would be “organizationally attractive” and efficient for the four cytosolic enzymes to be localized in a complex, or “metabolon” for the conversion of ALA to COPRO’gen III. Moreover, the delivery of COPRO’gen III to the mitochondria would be extremely efficient if evolution had situated the cytosolic complex on the mitochondrial outer membrane.

Recently, the crystal structure of human URO-synthase was reported at 1.85 Å resolution using a recombinant enzyme with a 22 residue N-terminal His_{10} tag extension [31]. Efforts to identify the enzyme’s active site and to investigate its reaction mechanism were not successful due to the inability to co-crystallize the enzyme with a substrate analogue. Instead, the active site region was inferred to be in the large cleft between the two major domains of the enzyme based on the distribution of conserved residues and by analogy to the vitamin B_{12} binding site of methionine synthase, an enzyme with structural homology to URO-synthase [31]. Because it was not possible to co-crystallize the enzyme with a substrate analog, no residue-ligand contacts were identified. Selective site-directed mutagenesis of ten conserved residues in the presumed active site cleft region did not reveal an essential catalytic residue whose mutation completely eliminated activity; and six out of the ten mutations resulted in wild-type activity while four reduced the enzymatic activity to 32 to 74% of wild-type, respectively [31].

Extensive studies have been performed on the mechanism of hydroxylmethylbilane ring closure. Suggested mechanisms have included have included the formation of a spiro intermediate [32], which can proceed via a fragmentation/recombination or a sigmatropic shift [33-36], or a lactone [37a, 37b]. An important step was the realization that the cyclization of HMBS was accompanied by the inversion of ring D by an intramolecular mechanism that only involves ring D [33] and the formation of an azafulvene intermediate [32]. This azafulvene then attacks carbon 16 (the carbon bound to the methylene bridge to ring C), instead of carbon 19 (which would lead to URO I), forming a spiropyrrolenine intermediate, with ring D vertical to the plane formed by the other three rings (see scheme 4.1 for a diagram of this mechanism). To form URO'gen III from this transition state, a fragmentation/rearrangement mechanism was suggested by Matheson and Corwin [38]. The formation of the spiro-pyrrolenine provides half a turn of the ring D inversion, and a fragmentation of the bond between rings C and D would allow for the second half turn, and subsequent formation of a new carbon-carbon bond [34]. Since HMB has a short half life of 4 min [30], non-enzymatically converting to URO I, the role of the enzyme must be to 1 - accelerate the removal of the hydroxyl group from ring A, possibly by it's protonation, and activating the azafulvene; 2 – to provide stereospecificity, protecting carbon 19 from the azavulvene; 3 – either provide stereospecificity to allow bond fragmentation between rings D and C, but not D and A, or if the mechanism involves a sigmatropic shift, possibly to deform the spiro-pyrrolenine structure to facilitate the overlap of electronic orbitals. This is a fascinating enzymatic mechanism that still remains to be demonstrated.

In this work, we determined the resonance assignments of URO-synthase, determined its 3D NMR solution structure, and demonstrated the absence of a direct cytosolic intermolecular complex formation between URO-synthase and HMB-synthase or URO-decarboxylase at up to 0.3 mM. Based on *in silico* docking studies and NMR chemical shift perturbations of specific residues caused by interactions with the enzyme's linear tetrapyrrole inhibitor, N_D-methyl-1-formylbilane, and its reaction product, URO'gen III, the active site was localized to specific residues in the cleft region between URO-synthase structural domains 1 and 2. The mapping of the active site provides a basis for further studies involving molecular mechanics/quantum mechanics simulations to elucidate the role of the enzyme's dynamics and flexibility in the reaction mechanism.

MATERIAL AND METHODS

Materials - Deuterium oxide, ^{15}N-ammonium chloride, D-glucose–U-$^{13}C_6$-1,2,3,4,5,6,7-d_7, DTT-d_{10} and ^{15}N-labeled amino acids (Phe, Tyr, Leu, Ile, Val, Lys and Arg) were purchased from Cambrigde Isotope Laboratories [(CIL) Andover, MA]. PBG was purchased from Frontier scientific (Logan, UT) and IPTG from Sigma (St. Louis, MO). All other reagents are from Fisher scientific (Pittsburg, PA) except as indicated in the text.

Protein Assays - Protein assays were as described in chapter 2.

Stable Isotope Labeling of UROS - Uniformly ^{15}N- or ^{13}C-labeled enzyme was synthesized by growing pET28b-UROS-Smt3 (see Chapter 2) in M9 minimal media containing $^{15}NH_4Cl$ (Cambridge Isotope Laboratories (CIL), Andover, MA), or U-[^{13}C]-glucose (CIL), in deionized water. ^{15}N/^{2}H-labeled enzyme was prepared similarly in $^{2}H_2O$. Uniformly ^{13}C/^{15}N/^{2}H-labeled enzyme was prepared in a similar fashion using U-[^{13}C-^{2}H]-glucose (CIL) and 99.9% $^{2}H_2O$. Special Ile-, Val-, and Leu-labeled enzyme was prepared according to Medek et al [39a] and Tugarinov and Kay [39b] using $^{15}NH_4Cl$, 99.9% $^{2}H_2O$ and either 2-ketobutyric acid-1,2,3,4-$^{13}C_4$,3,3-d_2 (CIL CDLM- 4611), sodium salt, 2-keto-3-methyl-d_3 butyric acid-1,2,3,4-$^{13}C_4$, sodium salt (Isotec 596418, Sigma-Aldrich, St. Louis, MS,), and [U-^{13}C-^{2}H]-glucose or 2-ketobutyric acid-4-^{13}C,3,3-d_2, sodium salt (Isotec 589276), 2-keto-3-(methyl-^{13}C)-butyric acid-4-^{13}C, 3-d_1 sodium salt (Isotec 589063), U-[^{12}C-^{2}H]-glucose (CIL), and ^{15}N-labeled, protonated Tyr and Phe (CIL). Protein expression and purification were as described in the previous chapter.

Uniformly [^{15}N,^{2}H]-labeled or [^{15}N,^{2}H,^{13}C]-labeled samples were concentrated to about 0.3-0.7 mM by ultrafiltration (Amicon centrifugal filter devices). Final buffer composition was 20 mM sodium phosphate, pH 7.45, 100 mM NaCl, 2 mM DTT-d10 (perdeutered DTT), 0.1 mM EDTA, in 90% H_2O / 10% D_2O (NMR buffer).

Selected amino acids labeling was performed by incorporating specific ^{15}N-labeled amino acids in the growth media.

In order to specifically label the methyl groups of valines and leucine and the δ methyl groups of isoleucines, we added the labeled precursor $^{13}C_4$-3,3-D_2 α -ketobutyric acid (CIL CDLM-4611) and $^{13}C_4$-3-D_1 α -ketoisovaleric acid (Isotec 596418, Sigma-aldrich, St. Louis, MS), to the growth medium, in combination with U-^{13}C-^{2}H-glucose and ^{15}N-NH_4Cl. H(CCO)NH-TOCSY and (H)C(CO)NH-TOCSY experiments were performed in these labeled samples. A slightly different labeling scheme was used for NOE type experiments. U-^{12}C-^{2}H-glucose and ^{15}N-NH_4Cl was used with 4-^{13}C-3,3-D_2 α-ketobutyrate (Isotec

589276) and 4,4'-^{13}C, 3D$_1$ α-ketoisovalerate (Isotec 589063), and ^{15}N-Tyr-^{1}H + ^{15}N-Phe-^{1}H (CIL).

For determination of stereo-specific assignments of leucine and valine methyl groups, a biosynthetically directed fractionally (10%) ^{13}C-labeled sample was prepared [40].

In order to obtain further NOEs for structure determination, U-^{15}N-^{1}H and U-^{13}C-^{1}H samples were prepared.

Stable Isotope Labeling of housekeeping HMBS (H-HMBS) - For ^{2}H, ^{15}N uniform labeling, the H-HMBS prokaryotic expression vector pET16b-H-HMBS in E. coli strain BL21 (see Chapter 2) was grown in M9 minimal media prepared in 99.9% Deuterium oxide, containing ^{15}N-ammonium chloride as the sole nitrogen source. The bacterial culture was grown to OD_{600} = 0.7 at 37° C, induced with 1 mM IPTG and further grown overnight at 25° C. The purification procedure was as described in chapter 2. MonoQ fractions of H-HMBS containing the enzyme-substrate intermediates E, $ES_{1,}$ ES2, and ES3 were pooled together or separated and further purified by gel filtration. Purified HMBS was kept in NMR buffer described above.

N$_D$-Methyl-1-formylbilane Preparation (NMF-bilane) — NMF-bilane octamethylester (compound 5, [41] was kindly provided by Drs. Clotilde Pichon and Ian Scott (Texas A & M University). This compound (~1 mg) was hydrolyzed overnight with stirring in 100 µl of 2N KOH at 4° C, followed by room temperature stirring until dissolved – with protection from light. The buffer was exchanged to the NMR buffer as described above.

Enzymatic Synthesis of uroporphyrinogen III (URO'gen II) from Porphobilinogen (PBG) - URO'gen III was synthesized for the NMR perturbation studies from 1.7 µmoles of PBG (using 500 µg of purified recombinant human HMB-synthase (see above) and 100 µg of purified recombinant human URO-synthase in 13 ml of 100 mM Tris buffer, pH 8.2, and 2 mM DDT. The reaction mixture was incubated for 15 min at 37° C under nitrogen in the dark. The URO'gen III was separated from the enzymes by ultrafiltration (Amicon, 10,000 MW cut-off). The sample was then bound to 50 µl of DEAE resin (GE Healthcare), equilibrated with the reaction buffer, and eluted with 50 µl of 2 M NaCl. The URO'gen III eluate was desalted into the NMR buffer described above using a 10 cm x 8 mm column of Bio-Gel P-4 resin (BioRad). All buffers were degassed and bubbled with nitrogen and all procedures were performed under nitrogen in the dark using a safety light (light <600 nm blocked) to avoid oxidation of URO'gen III to uroporphyrin III.

NMR spectroscopy – NMR spectra were acquired at 303 K on a Bruker DRX 500, 600 or 800-Mhz spectrometers equipped with four rf channels, a triple resonance probehead or cryoprobe, and a triple-axis pulsed field gradient unit. NMR spectra were processed by NMRPipe/NMRDraw [42] and analyzed with NMRView [43].

Sequence-specific assignment of the backbone resonances ($^{1}H_{N}$, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$) was achieved using data from the 3D triple-resonance experiments. Sequential resonances were matched against the primary protein sequence using deuterium-decoupled triple-resonance HNCACB / HN(CO)CACB spectra recorded on a U-[^{13}C,^{15}N,^{2}H] labeled sample [44]. In addition, specific ^{15}N-labeled amino acids were incorporated into UROS by synthesizing the enzymes in minimal media containing only a specific ^{15}N-labeled amino acid. The individually labeled amino acids included Phe, Tyr, Leu, Ile, Val, Lys and Arg. These samples were used for collection of [^{15}N,^{1}H]-HSQC spectra and aided in the amide assignments of UROS.

To assign the methyl side chains of valine, isoleucine and leucine, 3D (H)C(CO)NH-TOCSY [45] and H(CCO)NH-TOCSY spectra were recorded on a {I(δ-$^{13}CH_3$),L($^{13}CH_3$,$^{12}CD_3$),V($^{13}CH_3$,$^{12}CD_3$)}-U-[^{15}N,^{13}C,^{2}H]-labeled sample ([46,47]), and 3D ^{13}C- and, ^{15}N-edited NOESY spectra were recorded on a {I(δ-$^{13}CH_3$),L($^{13}CH_3$,$^{13}CH_3$),V($^{13}CH_3$,$^{13}CH_3$),F(^{1}H),Y(^{1}H)}-U-[^{15}N,^{2}H]-labeled sample [38, 47]. Phenylalanine and tyrosine side chain protons were also assigned from the above spectra combined with 2D-TOCSY and 2D-NOESY data on the same sample. The remaining side chain resonances were assigned using a combination of 3D ^{15}N -TOCSY, ^{15}N -NOESY spectra obtained from a U-[^{15}N] labeled sample and 3D ^{13}C -NOESY performed on a U-[^{13}C] -labeled sample. Stereo-specific assignment of leucine and valine methyl groups was accomplished by analysis of the ^{13}C -^{13}C coupling patterns in a ^{13}C-HSQC spectrum using a biosynthetically directed fractionally (10%) ^{13}C-labeled sample [40]. Slowly exchanging amide protons were identified in a series of ^{15}N-HSQC spectra recorded immediately after a lyophilized protein sample was dissolved in D_2O. Although the amino-terminal serine was assigned and included in the deposited data, it was excluded from residue counts here, with residue 1 being the native initiation methionine.

Secondary structure prediction - The chemical shift data for the $H\alpha$, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ nuclei were used to predict the URO-synthase secondary structure according the chemical shift index (CSI) method [48], using the CIS program of the PREDITOR web server (http://wishart.biology.ualberta.ca/shiftor/cgi-bin/preditor_current.py [49]). This method relies on the fact that the chemical shifts of the different nuclei in the protein backbone are related both to the type of amino acid and to the nature of the secondary structure in which they are located. By comparing the actual chemical shift for a nucleus in a specific amino acid with a reference value (i.e., the random coil chemical shift for that same nucleus of the amino acid), it is possible to predict the secondary structure for nucleus residues. The alpha proton ($^1H^{\alpha}$) chemical shift typically undergoes an upfield shift with respect to the random coil value in a helix and a comparable downfield shift in an extended β-strand conformation [51]. The alpha carbon ($^{13}C^{\alpha}$) chemical shift undergoes shifts opposite to those exhibited by the $^1H_{\alpha}$ nuclei in these two types of secondary structure. Deviations of the $^{13}C^{\alpha}$ and $^1H^{\alpha}$ chemical shifts from random coil values are used to identify α-helix and β-strand elements. The secondary structure was compared to the structure described by x-ray crystallography [31].

Methyl Side Chain Assignments – Ile, val and Leu side chain ^1H and ^{13}C resonance assignments were obtained by using specially labeled amino acid precursors 2-keto-3-methyl-d_3 butyric acid-1,2,3,4-$^{13}C_4$ and 2-ketobutyric acid-1,2,3,4-$^{13}C_4$,3,3-d_2, along with ^{15}N ammonium chloride and ^{13}C-glucose. Linear ordering of ^{13}C of these side chains prevents magnetization loss at branching carbons in 3D (H)C(CO)NH-TOCSY and 3D H(CCO)NH-TOCSY experiments which correlate side chain carbons and protons, respectively, with the backbone amide nitrogen and proton and enabled us to obtain methyl assignments for many Val and Leu side chains. We could not get Ile assignments via these experiments. The missing methyl assignments for Ile Val and Leu were obtained by careful analysis of ^{13}C-HMQC-NOESY and ^{15}N-HSQC-NOESY, ^{15}N-HSQC-NOESY-^{13}C-HMQC (this takes ^{15}N-HMQC-NOESY a step further by transferring magnetization to and evolving on the ^{13}C carbon that is attached to the proton that gave the NOE correlation to the amide in the ^{15}N-HMQC-NOESY spectrum).

Stereo specific assignments of methyl side chains of leucine and valine were obtained by analysis of the ^{13}C -^{13}C coupling patterns in a ^{13}C-HSQC spectrum of a UROS sample fractionally labeled with 10% ^{13}C during biosynthesis.

3D Solution Structure Determination - NOE-derived distance restraints were obtained from ^{15}N- or ^{13}C-edited 3D NOESY spectra [50] performed on a {I(d1-$^{13}CH_3$),L($^{13}CH_3$,$^{13}CH_3$),V($^{13}CH_3$,$^{13}CH_3$),F(^1H),Y(^1H)}-U-[^{15}N,^2H]-labeled [47] as well as fully protonated and uniformly ^{15}N- or ^{13}C labeled samples. Backbone torsional (ϕ and φ) angle restraints were calculated by the program TALOS. Slowly exchanging amide protons were identified from a series of 2D ^{15}N-HSQC spectra recorded after protein sample lyophylization and ^2H_2O addition. Structures of URO-synthase were calculated using a distance geometry-simulated annealing protocol with X-PLOR [51]. Structure calculations of the URO-synthase were done with manually assigned NOE-derived distance restraints. Hydrogen-bond distance restraints were added at the late stage of structure calculations for residues with slowly exchanging amide protons having characteristic NOE patterns and for helices from TALOS predictions. The NOE-derived restraints were categorized based on the observed NOE peak intensities (Table 3.4). Ramachandran plot analysis of the final structures (residues 1-260) with Procheck-NMR [52] evaluated the structure quality.

Resonance Chemical Shift changes and Differential Line Broadening Perturbation Analyses of Uroporphyrinogen III Synthase's titrated with HMB-Synthase, URO-Decarboxylase, PBG, URO'gen III, N_D-methyl-1-formylbilane - The titration experiments were carried out with 0.3-0.7 mM [$^{15}N,^{2}H$]-labeled URO-synthase (exchangeable hydrogens having been replaced with ^{1}H during the purification process) and either HMB-synthase, URO-decarboxylase, the cyclic product - URO'gen III, or the linear competitive inhibitor, N_D-methyl-1-formylbilane (NMF-bilane). Samples were in the NMR buffer described above and the titrations were carried out in a Bruker 500 MHz spectrometer with cryoprobe, or a Bruker 600, at 303 K. An initial [$^{15}N,^{1}H$]-HSQC spectrum of the free enzyme was recorded, after which ligands were added to the enzyme to obtain URO-synthase:ligand molar ratios of 1:0, 1:1 housekeeping-HMBS (forms “A” (E, “B” (ES_1) “C” (ES_2), or “D” (ES_3) or 1:0, 2:1, 1:1, 1:2 erythroid-HMBS (mix of forms “A” and “B”); 2:1, 1:1, and 1:2 (URO-decarboxylase); the individual H-HMBS intermediates in 2:1, 1:1, 1:2 ratios (additionally 4:1 in the case of a $ES_{2,3}$ mixture); 1:1:1 (HMB-synthase plus URO-decarboxylase), 100:1, 20:1, and 10:1 (URO’gen III); 7.7:1, 4:1, and 2:1 (NMF-bilane), respectively, with a new spectrum recorded after each addition. A reverse titration was carried with 0.5 mM [$^{15}N,^{2}H$]-labeled HMB-synthase titrated with unlabeled URO-synthase, in the same NMR buffer, to obtain 2:1, 1:1, and 1:2 HMBS:UROS molar ratios.

For investigation of the perturbations of the methyl side chains of valine, leucine and isoleucine residues of URO-synthase titrated with NMF-bilane (7.5:1, 3.75:1: 1.9:1, 1:1:, 1:1.9, 1:3.5, 1:6 and 1:9 UROS:inhibitor molar ratio), [$^{13}C,^{1}H$]-HSQC spectra were recorded in a $I(\delta\text{ı}\text{-}1^{3}CH_3),L(^{13}CH_3,^{12}CD_3),V(^{13}CH_3,^{12}CD_3)\}\text{-}U\text{-}[^{15}N,^{13}C,^{2}H]$-labeled UROS sample.

Before and during analysis, the degassed samples were kept under nitrogen in the dark or under a safelight. Resonance perturbation of URO-synthase by ligands was almost always observed as differential line broadening rather than changes in the chemical shifts, consistent with slow to intermediate chemical exchange of the binding ligands [53]. In order to quantify the line broadening of the assigned resonances of URO-synthase perturbed by proteins and/or small molecule ligands, the intensity (peak height) of each resonance was plotted against the molar percentage of protein, URO’gen III, or

NMFbilane. The average background intensity of the spectra was subtracted from each peak intensity and then the intensities were normalized for dilution by the added ligand solution.

Quantification of the levels of chemical shift perturbation - For each residue's resonance plot, the resonance intensities (peak heights) were expressed as a percentage of the intensity with no added ligand. The absolute value of the slope of the linear regression curve for each normalized plot of all concentrations was used to estimate the extent of line broadening of each residue's resonance in response to each ligand. Affected residues were divided into three groups corresponding to the extent of perturbation (strongly perturbed, red; moderately perturbed, yellow; and weakly perturbed, grey).

As an example, the calculations used to establish the perturbation status of residues 33, 116 and 123 are shown in table 3.1 and 3.2. The peak intensities of each residue, at each titration point (table 3.1) were converted to the percentage of the initial intensity (free enzyme) (table 3.2). A mathematical function to calculate the slope of this data was then applied to this data, and the results are shown in the column named “slope”.

	Peak intensity at each titration point (UROS:URO’gen III molar percentage)			
Residue	**0% URO’gen**	**1% URO’gen**	**5% URO’gen**	**10% URO’gen**
Leu 36	0.81	0.65	0.44	0.22
Gly 116	1.2	1.13	0.96	0.96
Glu 123	0.79	0.65	0.43	0.40

Table 3.1 - Peak intensity for selected residues, at each titration point with URO’gen III.

	% of Initial peak intensity, at each titration point (UROS:URO’gen III molar percentage)				
Residue	**0% URO’gen**	**1% URO’gen**	**5% URO’gen**	**10% URO’gen**	**Slope**
Leu 36	100	82.1	50.5	13.1	-8.34
Gly 116	100	101.1	88.7	95.1	-0.71
Glu123	100	84.6	50.3	48.2	-5.07

Table 3.2 - Percentage of the initial peak intensity for selected residues, at each titration point with URO’gen III, after subtraction of the background and correction for dilution of the enzyme.

The absolute of each slope value was plotted against each residue in a column type chart (see Fig. 3.3 and 3.4). In this example, residues 36 was considered to be highly perturbed, residue 123 to be mildly perturbed and residue 116 to be very weakly/non-perturbed.

In Silico Docking of URO'gen III and N_D-methyl-1-formylbilane to URO-Synthase - The AutoDock program (version 3.05; The Scripps Research Institute, La Jolla, CA) [54] was used to determine the lowest free energy structures for binding of the URO'gen III product or the N_D-methyl-1-formylbilane inhibitor, to URO-synthase. The 3D structures of each ligand were built using the InsightII program (Accelrys, Burlington, MA). The URO'gen III structure was then transferred to Charmm and energy-minimized using the Charmm27 force field parameters [55]. Two additional alternative energy-minimized ligand structures were obtained for URO’gen III by repeated "heating and cooling" of the initial structure *in silico* since the tetrapyrrole ring was kept fixed during docking. The N_D-methyl-1-formylbilane structure was energy minimized using the semi-empirical AM1 method [56] in the Gaussian-98 program [57] and only one final ligand structure was selected. For all energy-minimized ligand structures, the partial charges on individual atoms were calculated by Mulliken population analysis of the AM1 wave function. The four resulting ligand structures and the URO-synthase crystal structure coordinates (PDB ID: 1JR2-chain A [31]) were used as input to AutoDock. An initial grid of 100 x 118 x 126 points (centered on the URO-synthase molecule) with a spacing of 0.7 Å encompassed the entire molecule with sufficient space to allow the ligands to freely dock anywhere on the surface. Subsequently, a smaller grid of 94 x72 x 84 points with a spacing of 0.375 Å was used to restrict the docking of N_D-methyl-1-formylbilane to the region of the cleft between domains 1 and 2. Only the polar hydrogens were retained on the protein, all waters were removed and AutoDockTools was used to set the protein non-bonded parameters. Each docking experiment (using a genetic algorithm) consisted of an initial population of 100 individuals, 256 runs, and 5×10^6 energy evaluations per run, which yielded about 1500 generations. The remaining AutoDock default settings were used. The protein structure and the pyrrole methylenes were kept rigid while the ligand propionic and acetic side-chains were allowed torsional flexibility. Docking results were analyzed using the program "dockres" (http://inka.mssm.edu/~mezei/dockres). The conformations with a docking free energy < -6.5 kcal/mol obtained after each multi-run experiment were clustered together based on an all-atom root mean square deviation (r.m.s.d.) tolerance of 5.0 Å. The clusters were ranked based on the docking energy of each cluster's lowest energy conformer. The top 10 lowest energy clusters were reported. The free energies of ligand binding estimated by AutoDock were the sum of the intermolecular and ligand torsional energies calibrated empirically from actual protein-ligand dockings with known binding constants and the K_is were calculated by AutoDock using the equation: $\Delta G_{obs}=RT \ln K_i$ [54].

Molecular dynamics simulations - The time-dependent behavior of docked UROS-ligand complexes predicted by AutoDock were simulated by molecular dynamics (MD) using the parallel molecular dynamics simulator NAMD [58], with CHARMm22 force field and TIP3 water parameters [59]. URO'gen III and NMFbilane parameters, derived from atoms in a similar environment in other parameterized molecules, were added to the CHARMm force field. A CHARMm RTF topology file was built with the program Simulaid (http://inka.mssm.edu/~mezei/simulaid/) from the pdb coordinates, and the partial atomic charges calculated previously were added to this file. Prior to MD, the complex was solvated using a Metropolis Monte Carlo simulation (MMC) (http://fulcrum.physbio.mssm.edu/ ~mezei/mmc/), resulting in a system with 6800 water molecules, and a energy minimization step was performed for 20 pico-seconds (ps). The MD simulation was carried using periodic box conditions (no water molecules are lost), with a 2 fempto-second (fs) time step, a uniform dielectric constant of 1, and a cutoff of non-bonded forces (Van Der Waals) with a switching function starting at 10 Å and reaching 0 at 12 Å. The system was gradually heated from 10 to 310 K in 30 ps, and then maintained at constant temperature and pressure (1.01325 bar). Configurations were obtained at regular time intervals and the trajectory of the system was analyzed with the Visual Molecular Dynamics program (VMD) [60]. Additional simulations at 318 K and 328 K were performed, using as input coordinates the output of the previous simulation. The UROS coordinates were either the previously published crystal coordinates (PDB 1JR2, chain A) or those of the NMR solution structure presented in this work, as indicated in the text.

Site-directed mutagenesis – A partial list of mutagenesis of specific URO-synthase active site residues has been performed with the QuickChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA), according to the manufactor's instructions**.** The forward primer (the reverse primer was the complement of the forward primer) for introduction of the following mutations were: K7A: 5'-GAAGGTTCTTTTACTGGCGGATG CGAAGGAAGATG-3'; D8N: 5'-GTTCTTTTACTGAAGAATGCGAAGGAAGATGAC-3'; D8H: 5'-GTTCTTTTACTGAAGCATGCGAAGGAAGATGAC-3'; D8E: 5'-GTTCTTTTACTG AAGGAAGCGAAGGAAGATGAC-3'; Y19F: 5'-GTGGCCAGGATCCGTTTACAGGGATT AGG-3'; Y19C: 5'-GTGGCCAGGATCCGTGTACAGGGATTAGG-3'; L36A: 5'- CCACTT TGATCCCTGTTGCATCGTTTGAGTTT-3'; R65A: 5'-CTCTGCTTCCACTGCTGCGG GGCTGGTAAAAATGAG-3'; R65E: 5'-CTCTGCTTCCACTGAGTGCGGGGCTGGTAAAA ATGAG-3'; E123A: 5'-GGAGAAACCTGTGGAAATGCAGCAAAGCTTGCAGAATATATT

TGTTCC-3’; R148A: 5’- CCTGTGGAAACCTCAAAGCTGAAATCCTGCCAAAAG-3’; R148Q: 5’- CCTGTGGAAACCTCAAACAAGAAATCCTGCCAAAAG-3’; Y168F: 5’- G AAAGCATAACTGTGTTTCAGACAGTTGCACAC-5’.

Table 3.3 – List of NMR experiments, respective samples and protein labeling schemes. E-HMBS: erythroid HMBS; H-HMBS: housekeeping HMBS.

UROS resonance assignments and Structure determination:

Experiment	Label; [protein] (mM)	NMR experiment
Backbone assignments	^{13}C^{-15}N,^{2}H-UROS 0.5	HNCACB HN(CO)CACB
	selected ^{15}N-labelled amino acids-UROS 0.5	[^{15}N, ^{1}H]-HSQC
Methyl side chain assignments	I(δI-1^{3}CH$_3$),L(^{13}CH$_3$,^{12}CD$_3$), V(^{13}CH$_3$,^{12}CD$_3$)}-U- [^{15}N,^{13}C,^{2}H]- UROS 0.5	3D (H)C(CO)NH-TOCSY 3D H(CCO)NH-TOCSY
Additional methyl side chain assignments	{I(δI-^{13}CH$_3$),L(^{13}CH$_3$,^{13}CH$_3$), V(^{13}CH$_3$,^{13}CH$_3$),F(^{1}H),Y(^{1}H) }-U-[^{15}N,^{2}H]-UROS	3D ^{13}C-HMQC-NOESY 3D ^{15}N-NOESY-HSQC 3D ^{13}C-HMQC-NOESY-^{15}N
	0.5	2D-TOCSY 2D-NOESY
Additional assignments and NOEs	(10%) ^{13}C-labeled UROS 0.5	^{13}C-HSQC
	U-[^{15}N]-UROS 0.5	3D ^{15}N –TOCSY-HMQC ^{15}N –NOESY-HSQC
	U-[^{13}C]-UROS 0.5	3D ^{13}C-HMQC-NOESY

UROS Titratrions:

Experiment	Label; [Protein] (mM)		NMR Experiment: $[^{15}N, ^{1}H]$-HSQC	
	$^{15}N,^{2}H$-UROS	H-HMBS	UROD	ratio (UROS:HMBS:UROD)
UROS titration with H-HMBS(E, $ES_{1,\ 2,3,}$) +UROD	0.6	0		1:0:0
	0.45	0.45		1:1:0
	0.32	0.32	0.32	1:1:1
	$^{15}N,^{2}H$-UROS			(UROS:UROD)
UROS titration with UROD	0.28		0	1:0
			0.14	2:1
			0.28	1:1
			0.56	1:2
	$^{15}N, ^{2}H$-UROS	H-HMBS		(UROS:HMBS or UROS)
Titration of UROS with H-HMBS-E	0.28	0		1:0
		0.14		2:1
		0.28		1:1
		0.56		1:2
Titration of UROS with H-HMBS-ES_1	0.28	0		1:0
		0.14		2:1
		0.28		1:1
		0.56		1:2
				1
Titration of UROS with H-HMBS-ES_2	0.28	0		1:0
		0.14		2:1
		0.28		1:1
		0.56		1:2

Experiment	Label; [Protein] (mM)	NMR Experiment: [^{15}N, ^{1}H]-HSQC	
	^{15}N, ^{2}H-UROS	H-HMBS	Ratio (UROS:HMBS)
Titration of UROS with H-HMBS-ES$_3$	0.28	0	1:0
		0.14	2:1
		0.28	1:1
		0.56	1:2
Titration of UROS with H-HMBS ES$_{2,3}$	0.4	0	1:0
		0.1	4:1
		0.2	2:1
		0.4	1:1
		0.8	1:2
	^{15}N, ^{2}H-UROS	E-HMBS	
Titration of UROS with E-HMBS	0.6	0	1:0
		0.3	2:1
		0.6	1:1
		1.2	1:2
Titration of labeled HMBS with UROS	UROS (unlabelled)	^{15}N ^{2}H- H-HMBS (E, ES$_1$)	
	0	0.5	1:0
	0.24		2:1
	0.447		1:1
	0.778		1:2
Titration of labeled UROS with URO’gen III	^{15}N, ^{2}H-UROS (mM)	URO’gen III	(UROS:URO’gen III)
	0.3	0	1:0
		0.0025	100:1
		0.0119	20:1
		0.0223	10:1
		0.0982	2:1

Experiment	Label; [Protein] (mM)	Ligand (mM) NMR experiment: [^{13}C,^{1}H]-HSQC	
	^{15}N, ^{2}H-UROS (mM)	NMF-bilane (mM)	ratio (UROS:NMF-Bilane)
Titration of labeled UROS with the inhibitor NMF-bilane	0.3	0	1:0
		0.038	7.7:1
		0.072	4:1
		0.14	2:1
		0.3	1:1
	I(δι-1^{3}CH$_{3}$),L(^{13}CH$_{3}$,^{12}CD$_{3}$), V(^{13}CH$_{3}$,^{12}CD$_{3}$)}-U-[^{15}N,^{13}C,^{2}H]- UROS (mM)	NMF-bilane	
Titration of labeled UROS with the inhibitor NMF-bilane	0.375	0	1:0
		0.05	7.5:1
		0.1	3.75:1
		0.2	1.9:1
		0.38	1:1
		0.7	1:1.9
		1.3	1:3.5
		2.3	1:6.1
		3.4	1:9.0

RESULTS:

Backbone Resonance Assignments

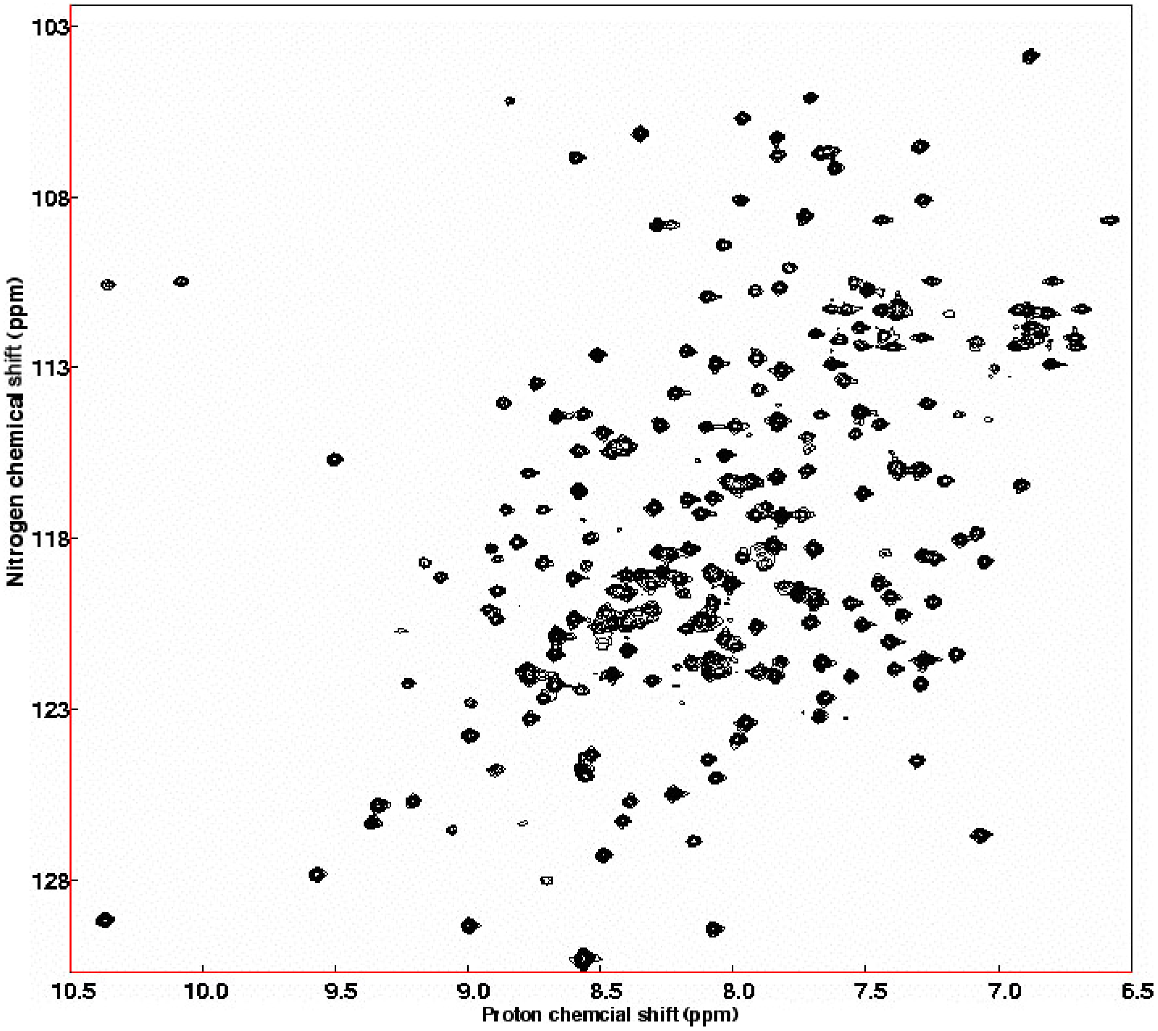


Figure 3.1 – 2D [^{1}H, ^{15}N]-HSQC spectrum of the uniformly ^{15}N-labelled UROS.

The amide ^{1}H, amide ^{15}N, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts of UROS were. [^{1}H, ^{15}N]-HSQC spectrum of UROS is shown in Fig. 3.1, in which all the peaks represent the NH groups of the protein (correlations between proton and nitrogen). Sequential assignments of amide ^{1}H, ^{15}N and α and β ^{13}C were used to determine the protein backbone resonances [61]. This approach exploits the relatively large one-bond ***J*** couplings between the backbone ^{13}C and ^{15}N nuclei and between the backbone protons and the ^{15}N and C^{α} nuclei to which they are attached. Two triple resonance spectra were recorded. The first spectrum, HN(CO)CACB, correlates the amide NH and amide ^{15}N chemical shifts of each amino acid with the C^{α} and C^{β} chemical shift of the preceding amino acid residue (Fig.

3.2A). The second spectrum, HNCACB, correlates the NH and ^{15}N chemical shifts with the intra-residue C^{α} and C^{β} chemical shifts, as well as those of the preceding residue (Fig. 3.2B).

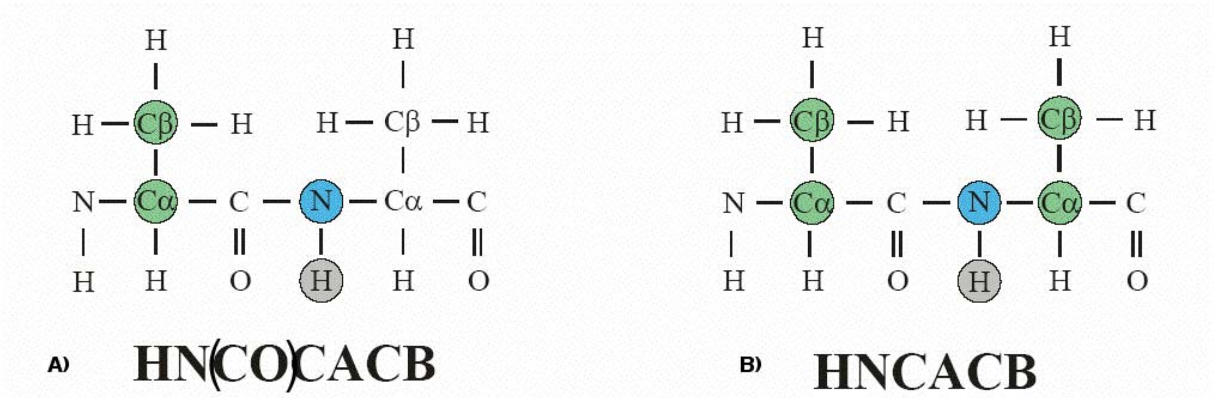


Figure 3.2 - Connectivities observed in the 3D NMR experiments. **A** - The HN(CO)CACB experiment correlates the alpha and beta carbons of the previous residue with the NH chemical shift. **B** - The HNCACB experiment correlates the alpha and beta carbons of both the previous and current residue with the NH chemical shift.

An example of the initial assignments made from these correlations is shown in Fig. 3.3. The NMR data for the recombinant human UROS permitted assignment of 100% of the backbone $^{13}C_{\alpha}$ resonances, 99.6% of the backbone $^{1}H_{\alpha}$ resonances and 94% of the non-proline backbone ^{15}N and $^{1}H_{N}$ resonances. In addition, approximately 85% of the side-chain ^{13}C and ^{1}H resonances were assigned and all assignment data was deposited in the *Biological Magnetic Resonance Data Bank* (BMRB) Database (http://www.bmrb.wisc.edu), accession number 7242 [62].

Fig 3.4 shows the [^{15}N, ^{1}H]-HSQC spectrum of UROS recorded at pH 7.45 and 303 K, with annotation of the residue corresponding to each resonance. The ^{13}C chemical shifts reported are referenced against the protonated enzyme, with the exception of leucine, valine methyl, and isoleucine d-methyl shifts, that were referenced to the I(d^{13}CH$_3$),L(^{13}CH$_3$,^{13}CH$_3$),V(^{13}CH$_3$,^{13}CH$_3$),F(^{1}H),Y(^{1}H)}-U-[^{15}N,^{2}H] enzyme.

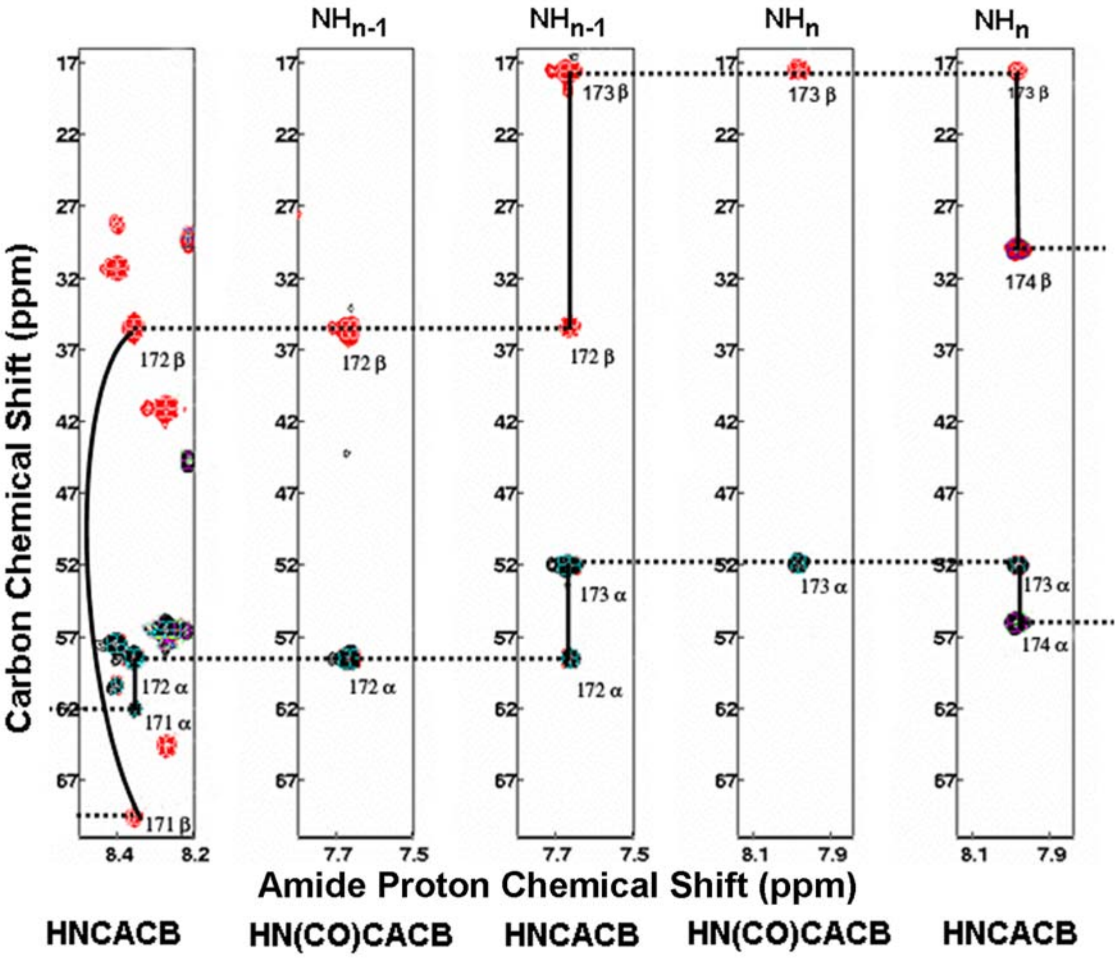


Figure 3.3 - Strip plots of HN(CO)CACB and HNCACB spectra of UROS at the ^{15}N frequencies of residues 172 through 173. Peaks originating from the same nucleus are connected by horizontal bars. In each HNCACB strip, four resonances were visible corresponding to the beta carbon of the preceding residue, the beta carbon of the current residue, the alpha carbon of the preceding residue, and the alpha carbon of the study residue. In the HN(CO)CACB strip the $^{\beta}$C (top resonance) and $^{\alpha}$C (bottom resonance) of the preceding residue were observed. Note that resonances of other residues originating in the same ^{15}N frequency were deleted from these strips for clarity of viewing.

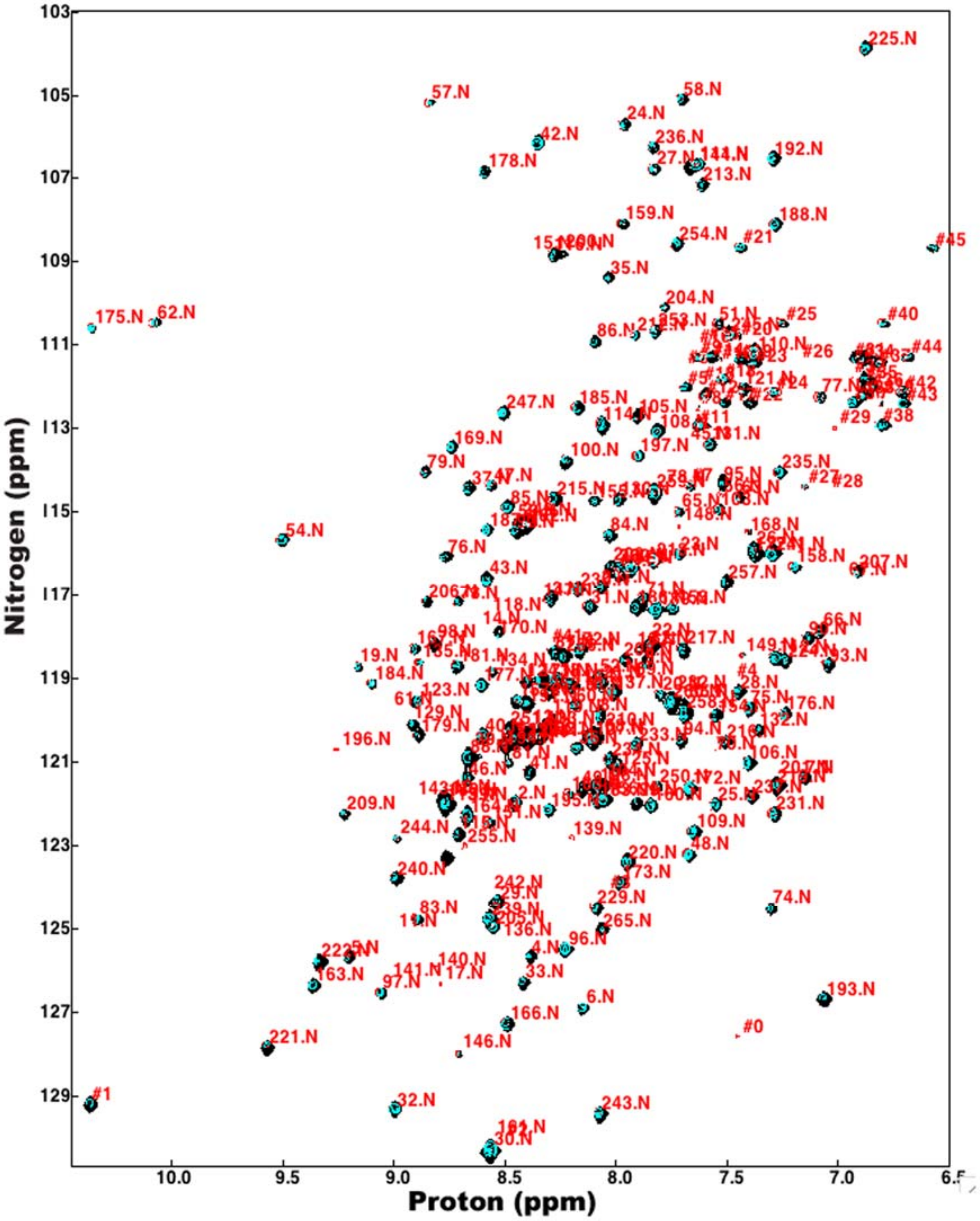


Figure 3.4 – [^{15}N, ^{1}H]-HSQC spectrum of the uniformly [^{15}N, ^{2}H]-labeled with assigned resonances shown in red labels. Backbone assignments of ^{1}H, ^{15}N nuclei were obtained for 94 % of the residues.

Secondary Structure Prediction - The chemical shift information obtained for the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ nuclei was used to identify secondary structure elements in UROS (Fig. 3.5) according to the CSI method [48]. Most of the predicted secondary structure elements match the pattern observed in the x-ray crystal structure.

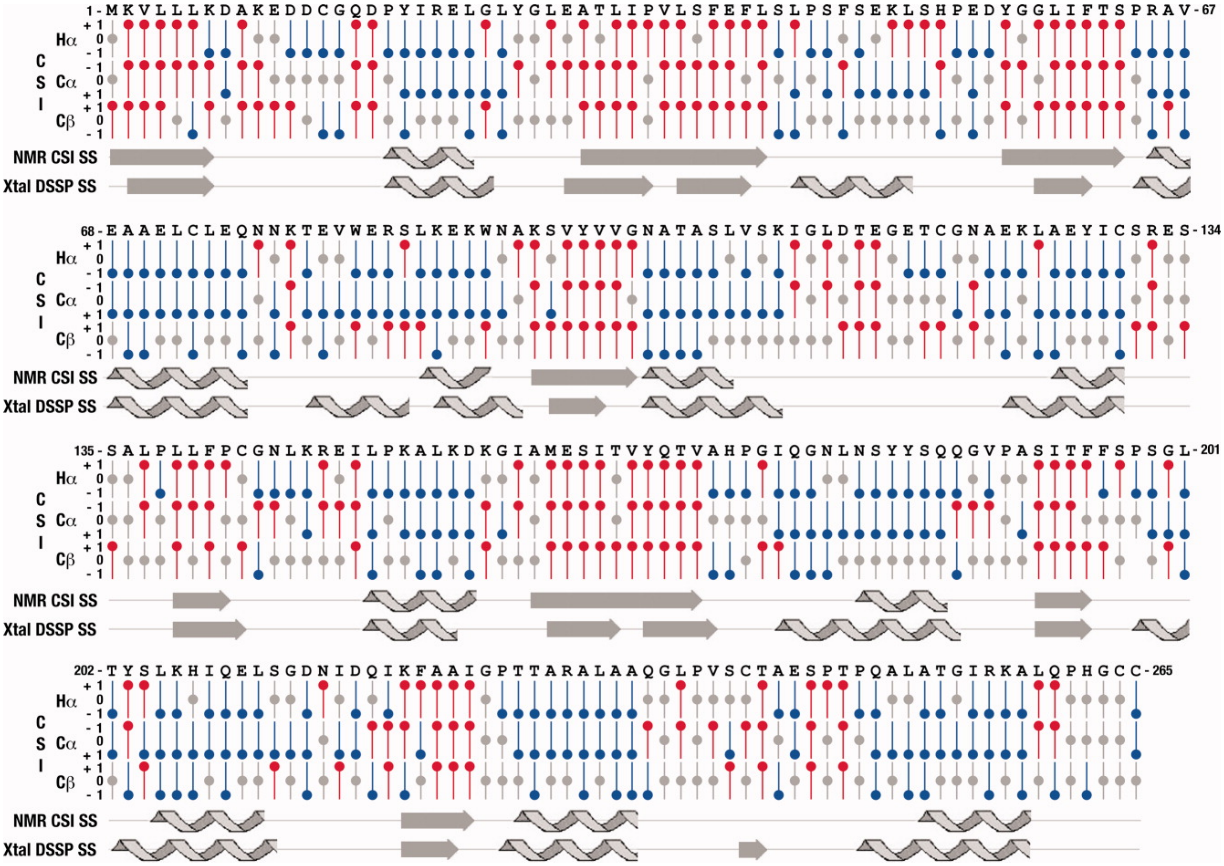


Figure 3.5 - CSI prediction of the secondary structure of URO-synthase. The CSI determined by the PREDITOR program for H_{α}, C_{α}, and C_{β} resonances are plotted under the URO-synthase amino acid sequence and the predicted secondary structures are indicated with arrows for beta structures and helices for α-helix structures. For comparison, the published secondary structure [31] calculated by DSSP (Xtal DSSP SS) for the crystal structure is shown below the CSI secondary structure prediction (NMR CSI SS). The results are displayed in a format similar to that used by the NMRview program. Note that the CSI values for Cαs were plotted in the opposite orientation to those for H_{α} and C_{β} to make it easier interpret the structure predictions.

Methyl Side Chain Assignments - Side chain ^{1}H and ^{13}C resonance assignments were determined for about eighty-five percent of the residues, that includes full assignment of valine, leucine methyl and isoleucine δ1-methyl groups. Figure 3.6 shows the ^{13}C-HSQC spectrum of a specifically labeled sample UROS sample.

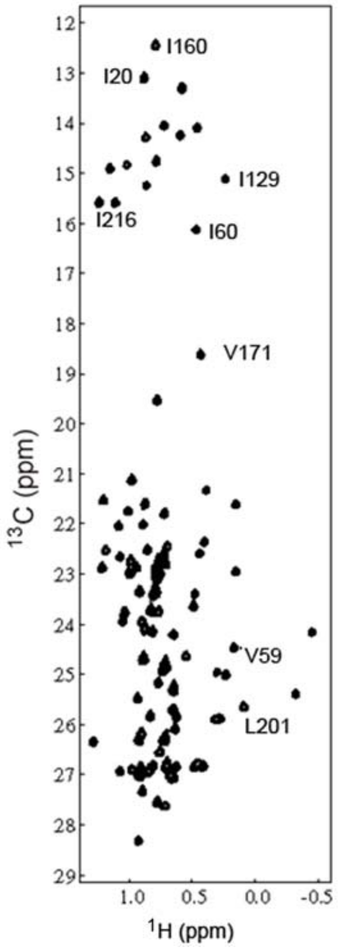


Figure 3.6 - ^{13}C-HSQC spectrum of a {I(δ-^{13}CH$_3$),L(^{13}CH$_3$,^{12}CD$_3$),V(^{13}CH$_3$,^{12}CD$_3$)}-U-[^{15}N,^{13}C,^{2}H]-labeled UROS sample. As an example, some resonances are identified with the corresponding residue.

Stereo Specific Assignments of Methyl Side Chains of Leucine and Valine - The side chains were fractionally labeled with 10% ^{13}C during biosynthesis. This resulted in the labeling scheme shown in Fig. 3.7. Analysis of the ^{13}C -^{13}C coupling patterns in a ^{13}C-HSQC spectrum of UROS of this fractionally labeled sample permitted to stereo-specifically assign the methyl side chain resonances (Fig. 3.8)

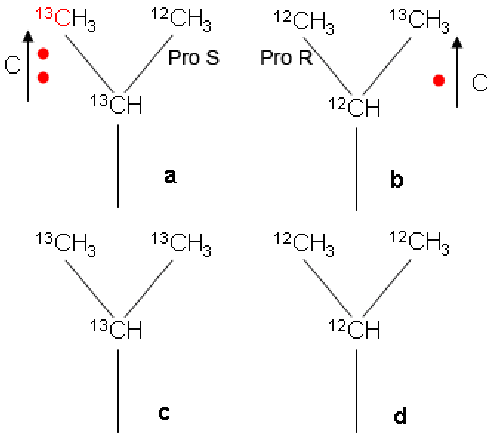


Figure 3.7 (**above**) - labeling scheme of for stereo specific assignement of Val and Leu. Fractionally labeling the sample with 10% ^{13}C results in the production of ≈9% of **a**, **b**, 1% of **c** and 81% of **d**.

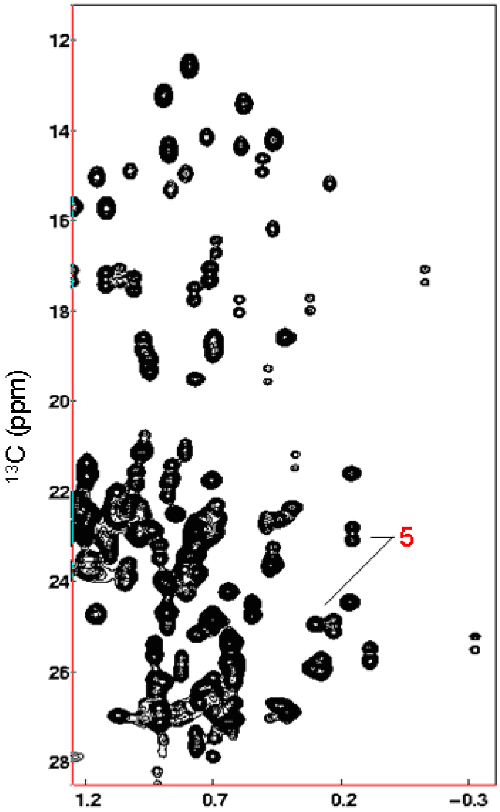


Figure 3.8 (**right**) - ^{13}C-HSQC spectrum on a $^{13}CH_3{}^{12}CD{}^{13}CH_3$ labeled sample. Pro R and Pro S groups show as singlets or doubles, according to the scheme of fig 3.8. As an example, the side chain methyl groups of Leu 5 are identified.

Spatial Restrains and Structure Calculation - NOE constrains were obtained by the proton distances obtained from the ^{13}C-HMQC-NOESY, ^{15}N-NOESY-HSQC, experiments performed on the $\{I(\delta I\text{-}^{13}CH_3), L(^{13}CH_3, ^{13}CH_3), V(^{13}CH_3, ^{13}CH_3), F(^{1}H), Y(^{1}H)\}$-U-[^{15}N,^{2}H]-labeled sample. Additional constrains were obtained from 3D ^{15}N -TOCSY, ^{15}N–NOESY experiments performed on U-[^{15}N] labeled and U-[^{13}C] -labeled samples, respectively.

The program XPLOR (http://xplor.csb.yale.edu/xplor/) was used to calculate 200 structures. 30 Structures were selected based on energy and root mean square deviation (r.m.s.d). A summary of the structure statistics is shown in table 3.4. The r.m.s.d. from the mean whole structure for non-H atoms was 0.61Å ± 0.11, 0.36 Å ± 0.05 for domain 1, 0.48 Å ± 0.08 for domains 2, and 0.10 Å ± 0.03 for the mid-sheet. The structure is well defined, except for the loop region between residues 10 and 18 (Fig. 3.9 A). The solution NMR structure of UROS is very similar to the X-ray structure [31], with an r.m.s.d. of 4.4 Å for the whole structure alignment, and 1.0 Å for domain 2 and 0.93 Å for domain 1 alignment. The higher r.m.s.d for the whole structure alignment is due to a tilting of the connecting hinges, resulting in a smaller cleft between domains on the solution structure. The enzyme consists of 2 globular domains, each organized as several parallel beta-sheets surrounded by alpha-helices. The C-terminus of the beta-sheets of each domain point to a center cleft, the putative active site. Both domains are connected by an anti-parallel beta-sheet (residues 36-41 and 168-172) (Fig. 3.9 B). Domain 1 (residue 1-35 and 173-265) has 5 beta-strands surrounded by 5 alpha-helices and has a large, likely flexible loop (residue 8-19) connecting a b-strand to an alpha-helix. The middle of this loop gave little or no NOEs (for residues 13-15) and is not well defined. Domain 2 (residues 42-167) is made of 4 β–strands surrounded by 7 a-helices and a small loop (residues 114-118) connecting two of the helices. Residues 115 and 116 show only a few NOEs and this makes this loop relatively less well defined than the rest of the structure. Hence the relatively larger r.m.s.d for domain 2 that included this loop as opposed to domain 1 where calculations left out the large loop. The surface of the protein is largely negatively charged, except for the cleft region, which is positively charged (Fig. 3.9 C). The statistics of the calculations is shown in table 3.4. The r.m.s.d for the individual domains of the assembly of the 30 final calculated structures is notably better than the r.m.s.d for the whole protein (Fig. 3.10 A and B). There might not be enough NOEs to define the relative orientation of the two domains more precisely. The reported two differently bent X-ray structures however suggested interdomain flexibility. Our titration analyses with competitive inhibitors also indicated that the two domains are flexible and this flexibility is important in catalysis. Because of the flexibility we did not hope that residual dipolar coupling measurements would be able to define the relative orientation of the two domains even better and therefore were not attempted. Ramachandran plot analysis of the final structures (residues 1-260) with Procheck-NMR [52] showed that 84.8 ± 1.1%, 12.2 ± 1.4%, 2.3 ± 0.8%, and 0.7 ± 0.4% of the residues were in the most favorable, additionally allowed, generously allowed, and disallowed regions, respectively. The corresponding values for the residues in the secondary structure regions (sheets and helices) were 97.8 ± 0.5%, 1.9 ± 0.5%, 0.3 ± 0.3%, and 0 ± 0%, respectively. The NMR and x-ray structures aligned each domain separately with r.m.s.d. around 1 Å (Fig. 3.10 C and D).

Total Experimental Restraints	4282					
Total NOE Distance Restraints	3899					
Total Ambiguous						
Total Unambiguous	3899					
Intra-residue	816					
Inter-residue	3083					
Sequential (|i-j|=1)	885					
Medium (2≤|i-j|≤4)	651					
Long Range (|i-j|>4)	1547					
Hydrogen Bond Restraints	93					
Dihedral angle restraints from Talos						
Φ	201					
Ψ	189					
Non-experimental restraints						
Talos inferred hbonds	43					
Final Energies						
E_{TOT}	464.9 ± 9.3					
E_{NOE}	32.9 ± 3.2					
E_{DIH}	2.3 ± 0.5					
E_{L-J}	-1462.1 ± 15.1					
Ramachandron Plot (%)	**Full Molecule**	**Secondary Structure**				
Favored Regions	84.8 ± 1.1	97.8 ± 0.5				
Additional Allowed Regions	12.2 ± 1.4	1.9 ± 0.5				
Generously Allowed Regions	2.3 ± 0.8	0.3 ± 0.3				
Disallowed Regions	0.7 ± 0.4	0 ± 0				
Cartesian Co-ordinate RMSDs (Å)			domain1	domain1 sstr	domain2	domain2 sstr
backbone	0.61 ± 0.11	0.53 ± 0.11	0.36 ± 0.05	0.33 ± 0.05	0.48 ± 0.08	0.31 ± 0.04
non-h atoms	0.95 ± 0.09	0.84 ± 0.09	0.77 ± 0.05	0.73 ± 0.04	0.90 ± 0.10	0.70 ± 0.04

Table 3.4 - Structural statistics of the URO-synthase NMR structure determination.

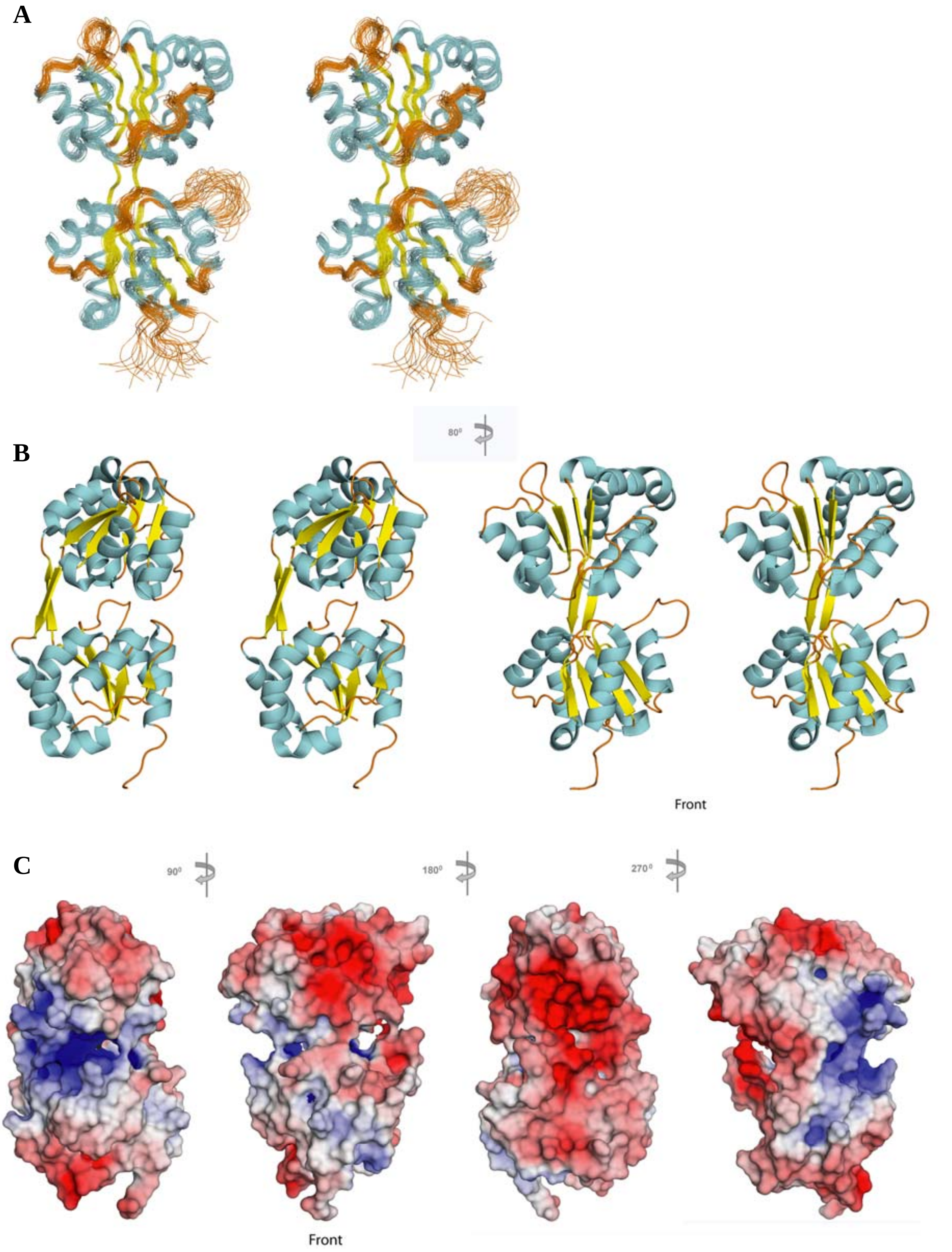


Figure 3.9 - 3D solution structure of URO-synthase. **A -** Stereogram of an overlay of the best 30 solution structures of URO-synthase (RMSD = 0.61 Å).The loops were colored in orange, helices in cyan, and beta-sheets in yellow. All figure representations were generated with the molecular visualization program Pymol [63]; **B -** Stereograms of the cartoon representation of the median structure, shown from a side view (left), and from a

80° rotation showing a frontal view. Color representation was as in; **C** - Electrostatic surface potential of URO-synthase surface. The electrostatic surface potential was calculated with the computer program APBS tools [64] and mapped onto the surface representation of URO-synthase. Color indicates potential range from -6 kT/e (red) to +6 kT/e (blue). The views from left to right are: left side, frontal, right side, back.

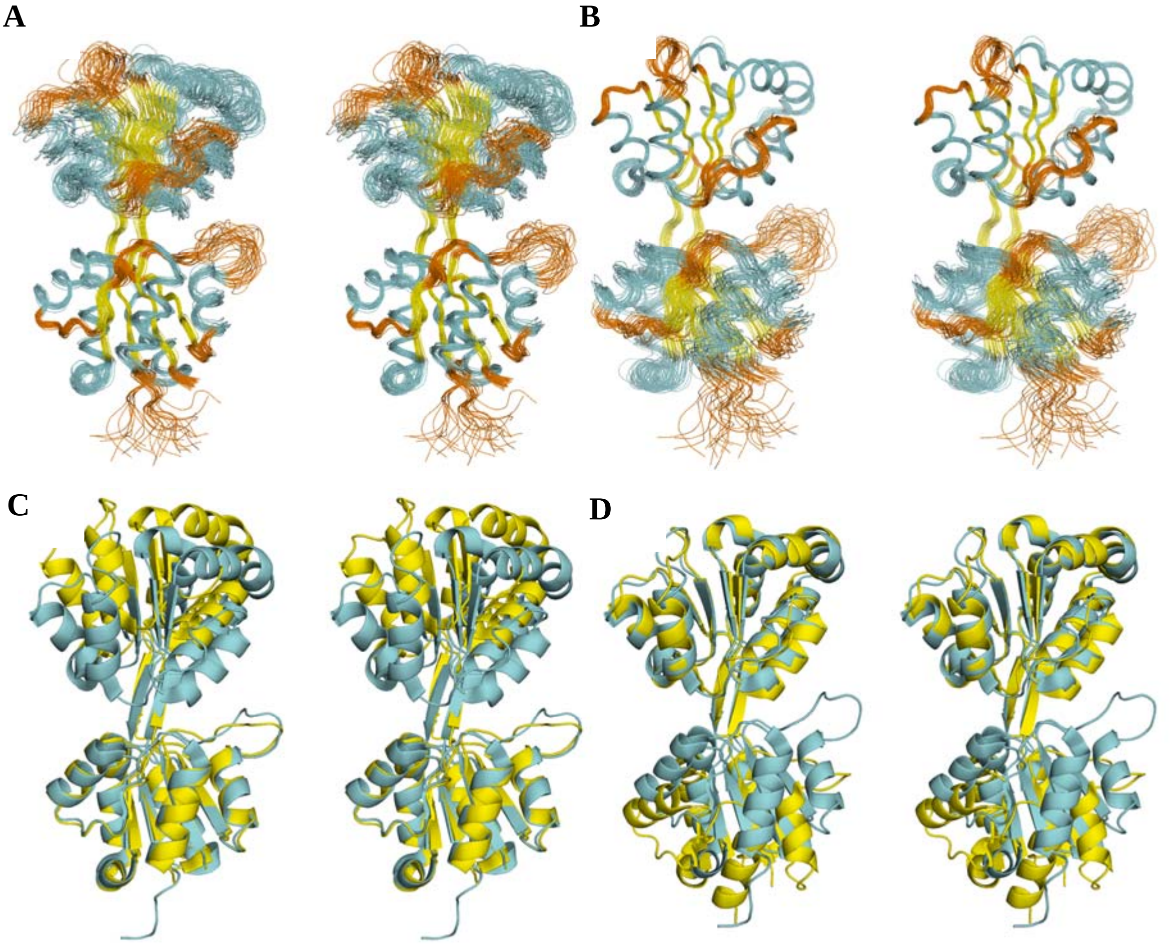


Figure 3.10 - Inter-domain flexibility of the URO-synthase structure. **A** - Stereogram view of a superimposition of the 30 best URO-synthase solution structures's domain 1 (non-H atoms r.m.s.d. = 0.77 Å / backbone = 0.36 Å); **B** - Stereogram view of a superimposition the 30 best URO-synthase solution structures's domain 2 (non-H atoms r.m.s.d. = 0.90 Å / backbone = 0.48 Å); **C** - Comparison of the URO-synthase NMR solution structure (cyan) aligned to domain 1 of the structure previously determined by x-ray crystallography (yellow) (non-H atoms r.m.s.d. = 0.895 Å / (0.727 Å for the backbone atoms); **D** - Comparison of the URO-synthase NMR solution structure (cyan) aligned to domain 2 of structure determined previously by x-ray crystallography (yellow) (non-H atoms r.m.s.d. 1.041 Å / backbone atoms r.m.s.d = 0.882 Å for backbone atoms. Note that the simultaneous

alignment of both domains resulted in a non-H atoms r.m.s.d. of 4.42 Å, and 4.357 Å for backbone atoms.

Normal Mode Analyses (NMA) of the URO-synthase Solution Structure - Since UROS-synthase structure flexibility has been suggested [31], and since there is a large shift of the two globular domains between the crystal and the solution structures, we employed Normal Mode Analyses (NMA) to study the inherent flexibility of the protein, which might have implications in the mechanism of substrate-binding and catalysis. NMA is a method to study slow motions, such as low frequency movement of domains, in contrast to molecular dynamics, which is appropriate to study fast vibrational movements in the nano second range. Calculations were performed using the NMA web-server (Webnm@) at the University of Bergen, Norway (http://www.bioinfo.no/tools/normalmodes), which implements the method of Hensen *et al* [65]. In normal mode analyses, modes 1 through 6 represent translational and rotational movements, which can be discarded for the purpose of our analyses. Table 3.5 shows the energy of the deformation of modes 7 though 20. These energies are associated with every atom and characterize rigid regions. The lowest energies are associated with larger rigid regions, and therefore domain movements.

Figure 3.11 shows the conformation extremes of mode 7 and 8. Both modes produce a rotation/twist of the domains relative to each other. The differences between these modes are in mode 7 there's a rotation of the domains that results from the tilting of the hinge connecting both domains, and in mode 8 results from the tilting of domain two at the top of the hinge, which also generates a forward movement that changes the size of the cleft between domains. Fig. 3.12 shows the vector field resulting from the atom displacements. Of note, Mode 8 better explains the movements associated with a transition from the crystal to the solution structures (Fig 3.13). This analysis shows that the enzyme structure is very flexible, and domain movements can change the size of the cleft, as previously suggested based on the structural differences between the two units packed in the crystal cell (PDB 1JR2, chain A and chain B), and between the crystal and solution structures.

Normal mode index	Deformation Energy	Normal mode index	Deformation Energy
7	1404.19726847	**14**	8779.93424587
8	2029.10807677	**15**	8620.78898735
9	3364.19818456	**16**	9883.51799372
10	3461.84529559	**17**	10506.0643986
11	4350.99792335	**18**	11790.2113041
12	4886.94355955	**19**	11588.6046826
13	7017.58076624	**20**	13648.8377312

Table 3.5 - Deformation Energies output of the Normal Mode Analyses at the Webnm@ server.

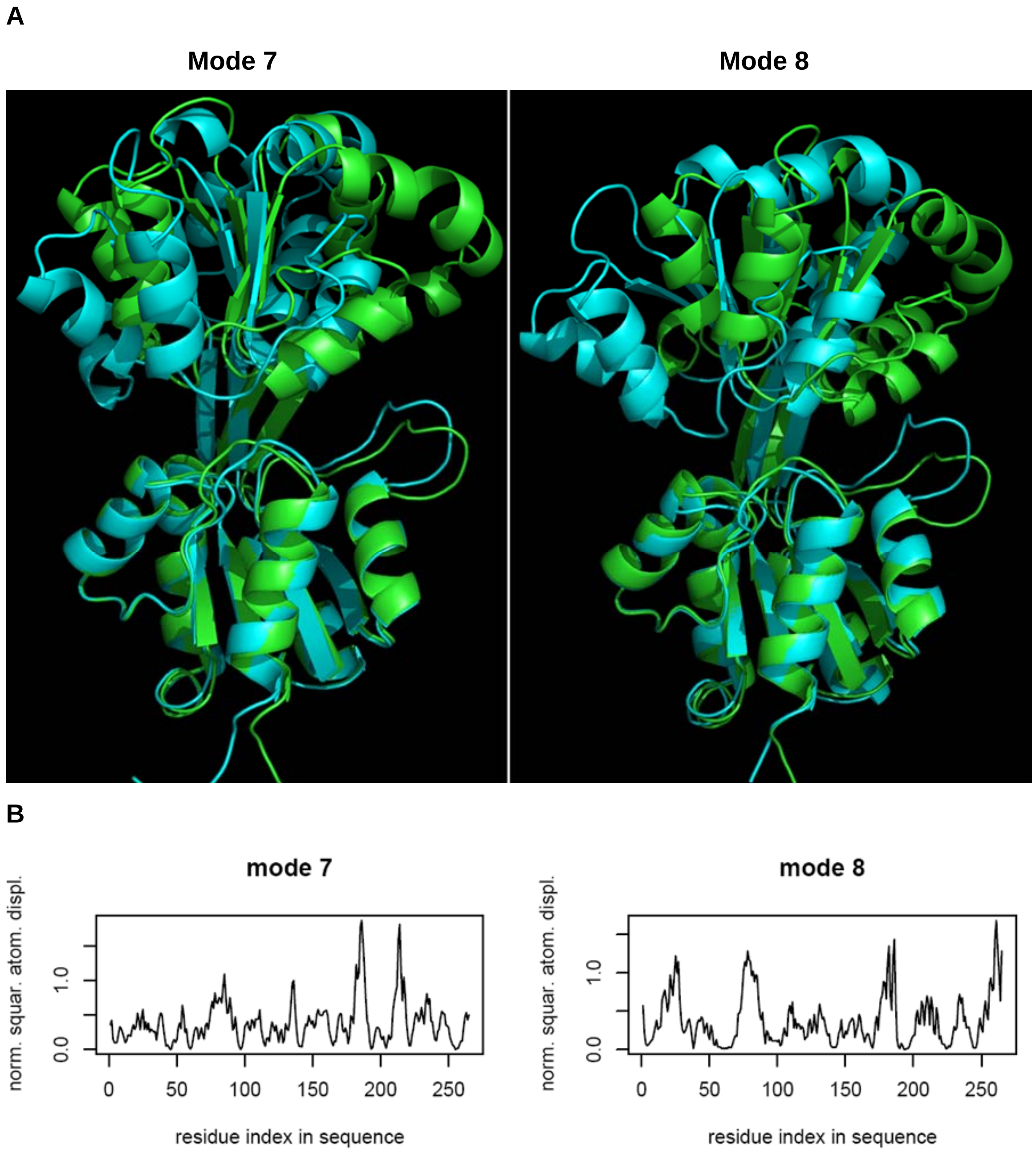


Figure 3.11 - Conformation extremes of Mode 7 and 8. **A** - Both modes 7 and 8 generate a twist kind of movement of the domains relative to each other, and mode 8 also generates a forward movement, which decreases the size of the cleft between domains. **B** – r.m.s.d. of each amino acid between the NMA structures.

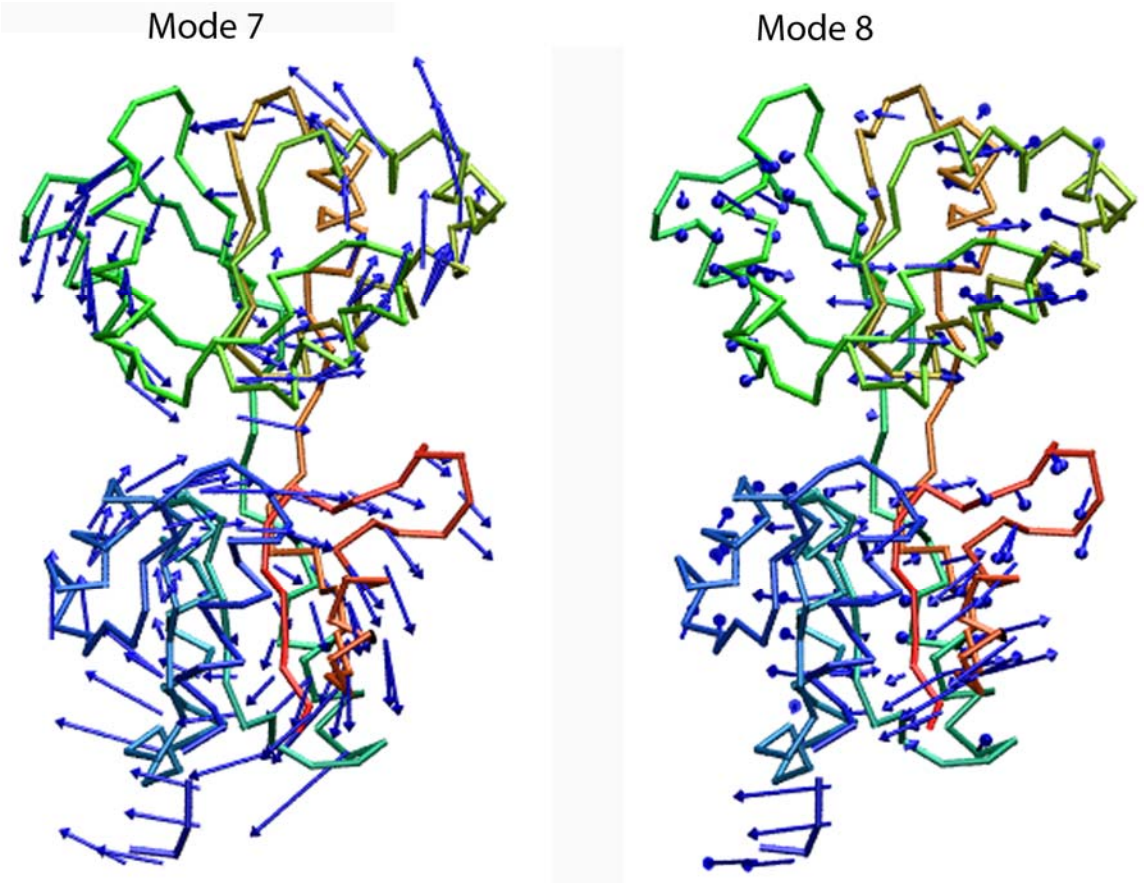


Figure 3.12 - The vector field representations calculated by the Webnm@ server. The vector field is calculated over cubic regions with an edge length of 3 Angstroms, containing on average 1.3 C-alpha atoms. The vector field defined on a regular lattice at the center of each cube is the mass-weighted average of the displacements of the atoms in the cube.

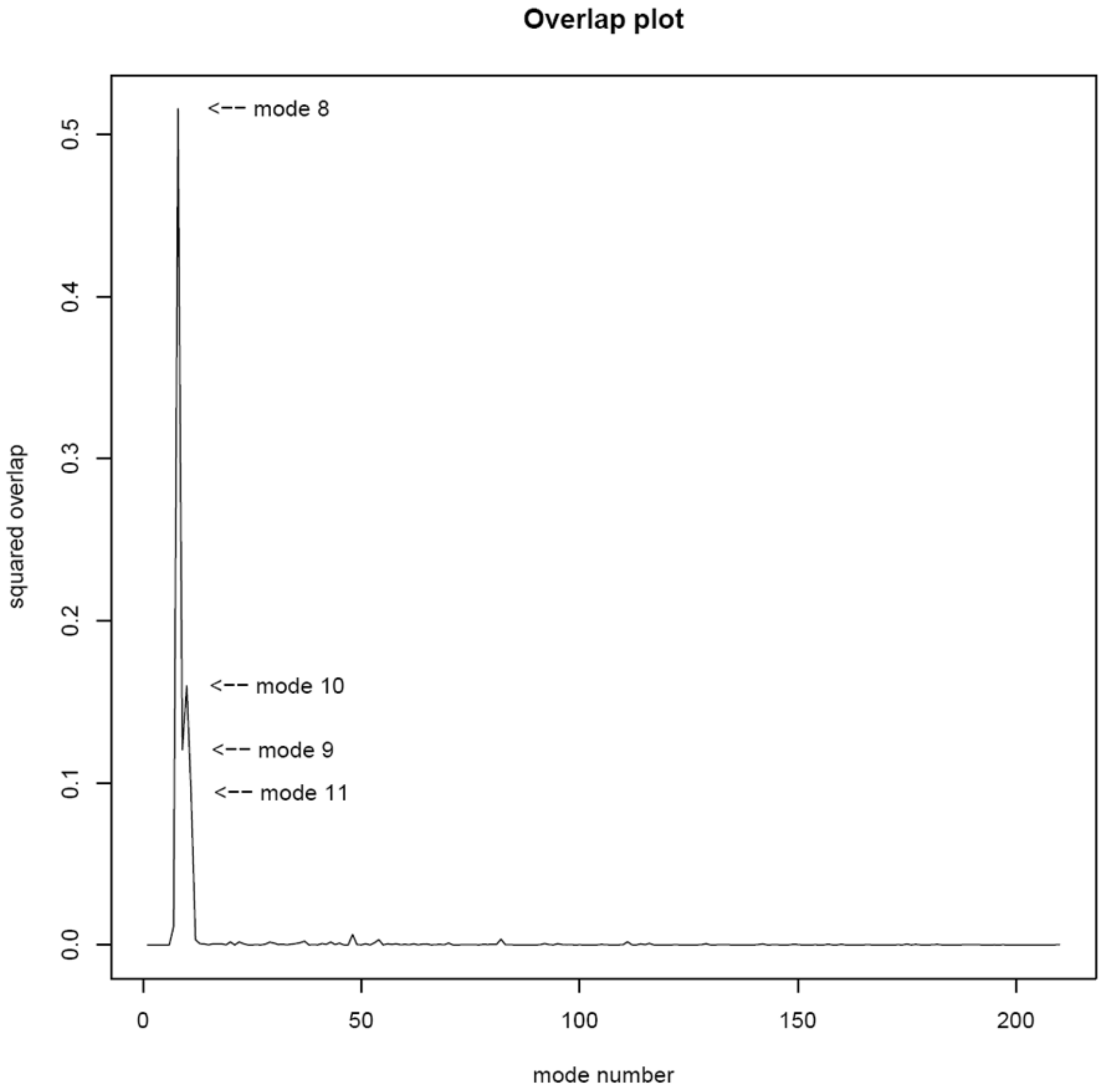


Figure 3.13 - The calculation of the dot product (overlap) between the difference vector and the full set of normal modes for the crystal and NMR solution structures. The Overlap identifies mode 8 as contributing mostly to the structural difference.

Does UROS Interact in a Complex With HMBS and/or UROD?

Titration of Stable Isotope-Labeled URO-Synthase with HMB-Synthase or URO-Decarboxylase - The URO-synthase [^{15}N,^{1}H]-HSQC spectra of [^{15}N,^{2}H]-labeled URO-synthase were recorded and compared before and after addition of HMB-synthase and/or URO-decarboxylase. In these titrations we first used HMB-synthase forms "A" (holoenzyme containing only the covalently-bound dipyrromethane cofactor) and/or “B” [66] (the enzyme containing in addition to the dipyrrole cofactor, one bound pyrrole unit [67]. Both the Housekeeping and Erythroid isozymes were tested. As shown for one region of the spectrum in Figure 3.14A, only the resonance for Asn-77 was minimally perturbed when URO-synthase was titrated with housekeeping HMB-synthase form A (in the rest of the spectrum, Tyr-19, Thr-62, and Ala-222 were also very slightly perturbed). When titrated with URO-decarboxylase, only resonances for residues Glu-48, and Ser-212 were slightly perturbed (Fig 3.14B). These weak perturbations that did not cluster together were presumably artifacts due to slight differences in pH or ionic strength between the URO-synthase and added enzyme. A titration of a mix of UROS+HMBS with UROD did not show perturbation (data not shown).

Similarly, the titration with housekeeping HMBS form “B” (ES_1), and Erythroid HMBS (a mix of forms “A” and “B”) did not result on significant perturbations of the spectrum (Fig 3.14 C and D, respectively). A few resonances were slightly line broadened and/or shifted, but these were located in the cleft region between domain 1 and 2, inaccessible to a large protein such as HMBS. As will be seen on the next experiments, these were not due to protein-protein contacts.

An inverse experiment was performed, in which we studied the perturbation of [^{15}N, ^{2}H]-labeled H-HMBS (form A) resonances when titrated with unlabeled UROS. However, the quality of the [^{15}N, ^{1}H]-HSQC spectrum of HMBS was of much inferior quality than the spectrum of UROS (due to line broadening) and the perturbation of specific resonances was difficult to analyze (Figs. 3.15 and 3.16).

Thus, these results did not support a stable, direct interaction between URO-synthase and the adjacent cytosolic heme biosynthetic enzymes, HMB-synthase and URO-decarboxylase.

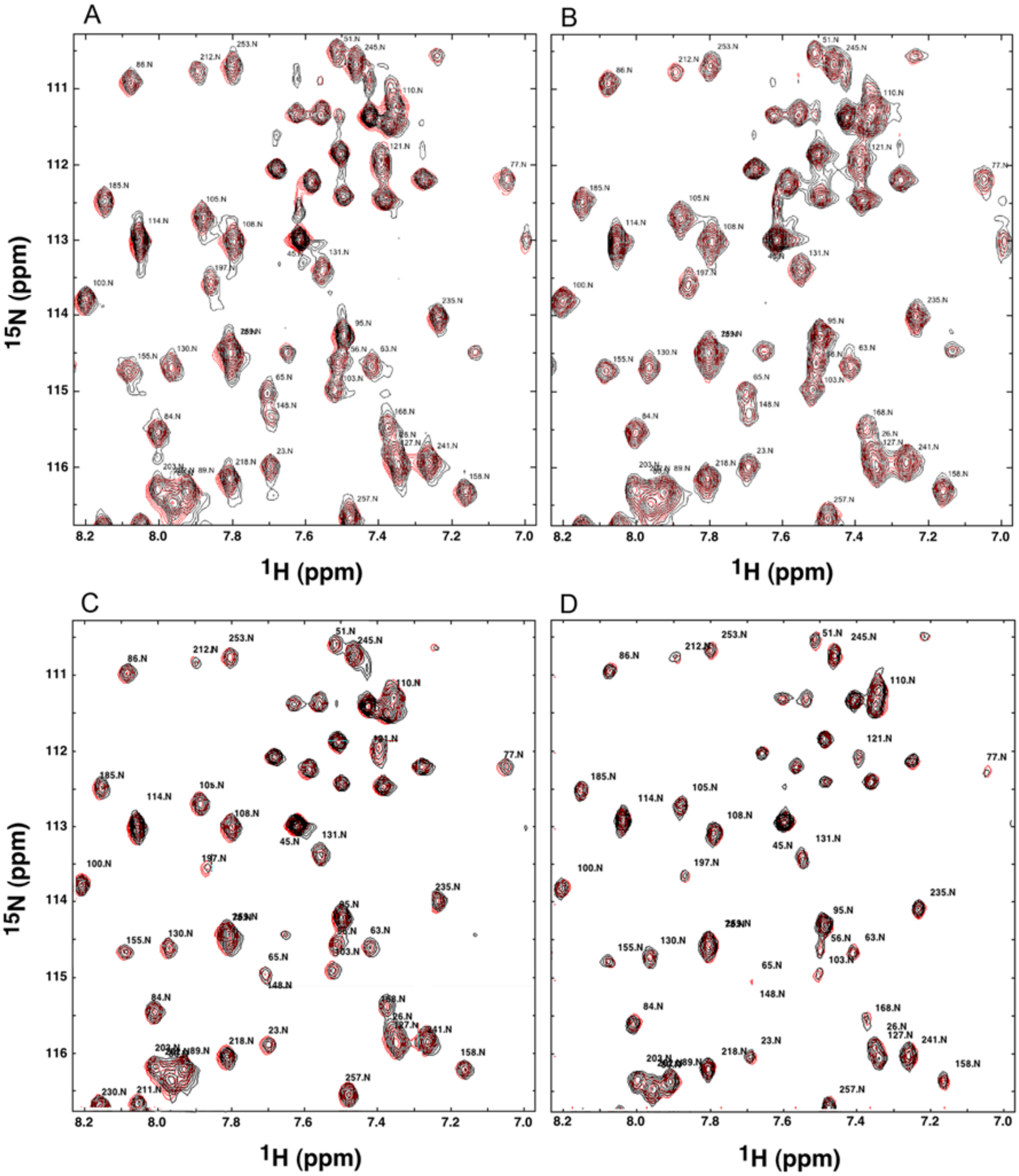


Figure 3.14 - $[^{15}N,^{1}H]$-HSQC spectra of labeled URO-synthase mixed with the prior or subsequent enzyme in the pathway. **A** - $[^{15}N,^{1}H]$-URO-synthase (resonances shown in red) was titrated with increasing concentrations of HMB-synthase form A (ES) (shown in black overlay for the 1:2 mixture) as described in methods. For the region of the spectrum shown, only Asn-77 was slightly line-broadened at the highest concentration (0.6 mM) of added enzyme. **B** - Titration with URO-decarboxylase as in A, with the 1:2 mixture with URO-decarboxylase in black overlay; only Ser-212 was slightly line-broadened at the highest enzyme concentration (0.63 mM). Arrows identify the resonances of residues 77 and 212. **C** - Titration with form "B" (ES_1) of housekeeping HMBS. **D** - Titration with form A of erythroid HMBS enzyme.

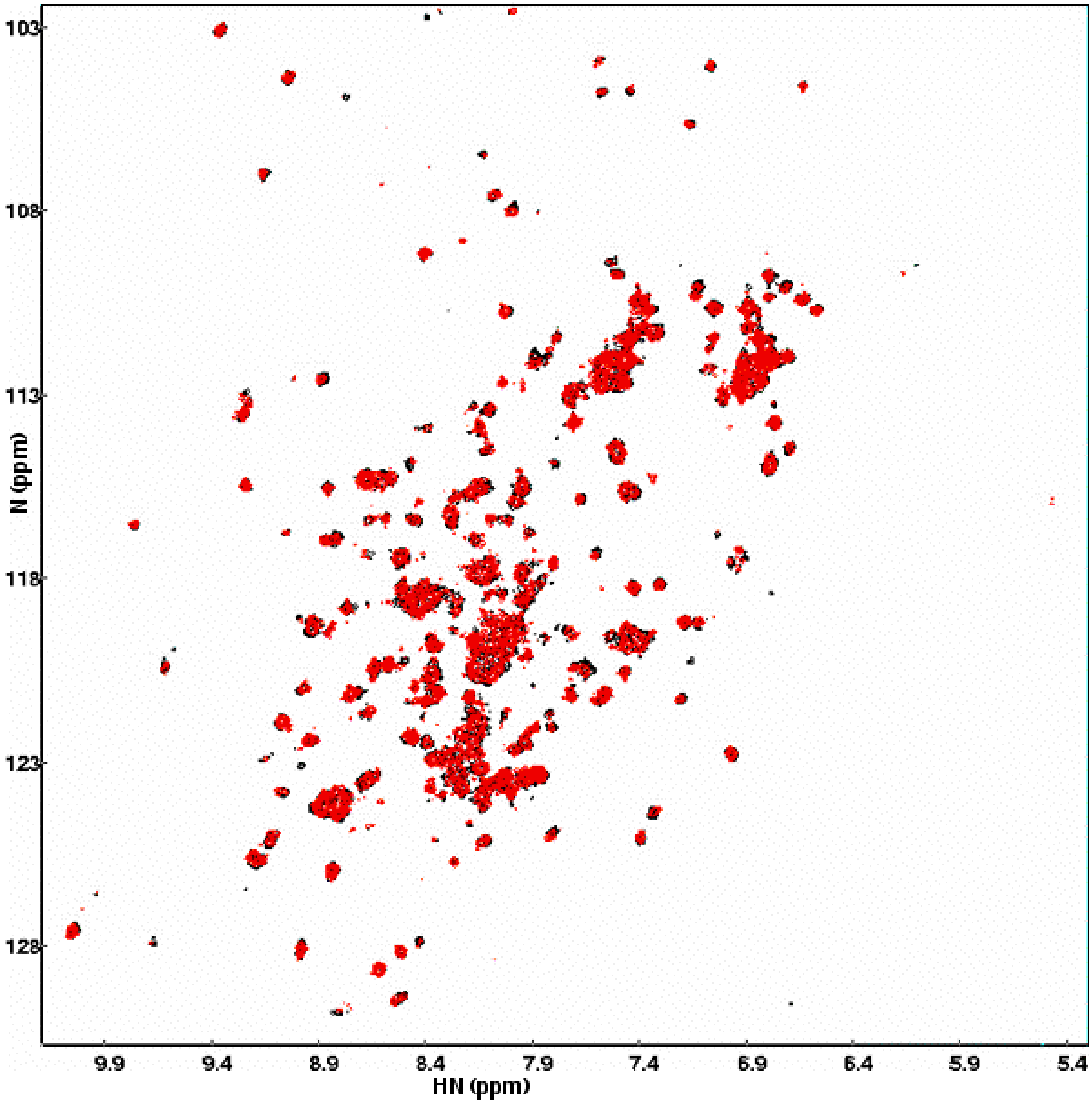


Figure 3.15 - Superposition of the [^{15}N, ^{1}H]-HSQC spectrum of the free, [^{15}N, ^{2}H]-labeled H-HMBS (0.5 mM, black) with the spectrum of HMBS mixed with UROS (H-HMBS: UROS 1.2 ratio, red).

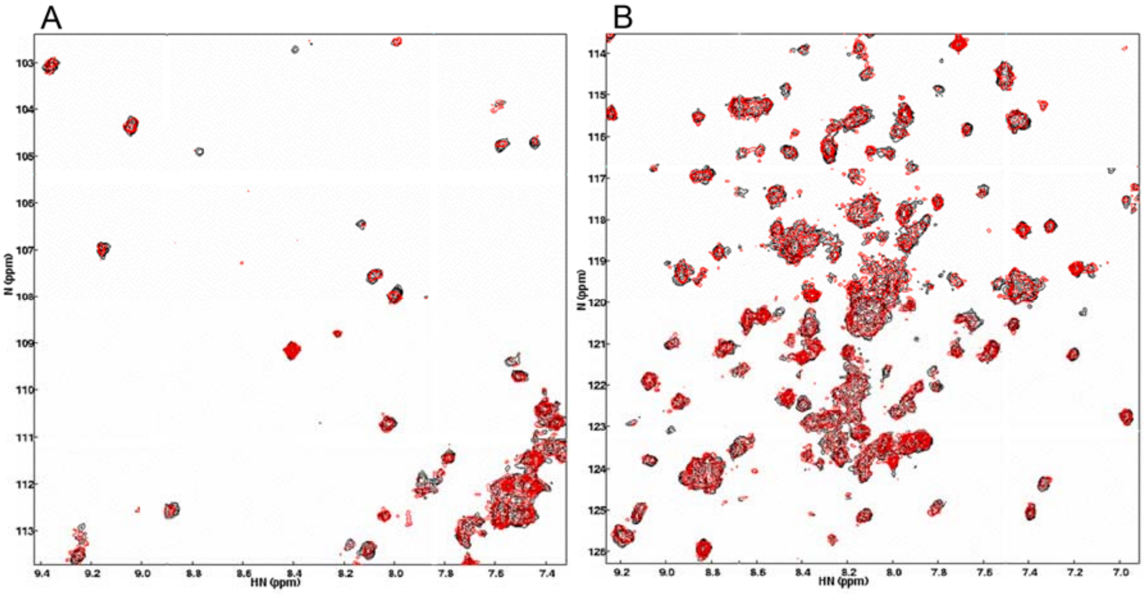


Figure 3.16 - Superimposition of selected regions of the [^{15}N, ^{1}H]-HSQC spectra of free [^{15}N, ^{2}H]-labeled HMBS (0.5 mM, black) with the spectrum of labeled H-HMBS mixed with UROS (H-HMBS:UROS 1.2 ratio, red. **A** – 103-114ppm (N) region. **B** – 114-126 ppm (N) region.

Does UROS Form a Complex with the HMBS's Enzyme-Substrate Intermediates?

Since we found no evidence for a direct protein-protein interaction between UROS and HMBS, we hypothesized that UROS may interact with enzyme-substrate intermediates of HMBS (Fig. 3.17) and this interaction may be mediated by the HMBS-tethered pyrrole chain (Fig. 3.18). It has been observed that the conformation of HMBS changes during catalysis [68]. The reactivity of a buried cysteine (Cys-134) with N-ethylmaleimide increases progressively as HMBS proceeds from the holoenzyme E through ES$_3$. It is therefore conceivable that a HMBS conformation only present at the later stages of product formation could allow UROS to either bind HMBS, or the flexibility of the loop that holds the di-pyrrole cofactor could expose the growing pyrrole, and possible bind UROS, as depicted in figure 3.18. It would be expected that the tethered UROS would experience perturbation of some residues close to HMBS, in addition to residues perturbed by binding of the pyrrole polymer. Therefore, we separately titrated UROS with the purified H-HMBS enzyme–substrate intermediates (Fig 3.17, [69-72], see chapter 2): free enzyme - E; enzyme with one PBG molecule bound to the dipyrrole cofactor – ES$_1$; enzyme with two cofactor-bound PBG molecules - ES$_2$; enzyme with 3 cofactor-bound PBG molecules – ES$_{3,}$).

Both HMBS's forms C (ES_2) and D (ES_3) induced similar perturbation of UROS's resonances. However, during the experiments, sufficient amounts of URO'gen III were formed to account for the perturbation of UROS resonances (about 15 μM). Although these enzyme-substrate complexes are stable for several weeks at 4^{o} C, in NMR conditions, at 30^{o} C, URO'gen III was generated from substrate-loaded HMB-synthase by catalytic conversion of the released HMB by URO-synthase [73]. Therefore, we sought to distinguish the perturbation of URO's residues induced by URO'gen III, from those induced by the HMBS-substrate intermediates, if any.

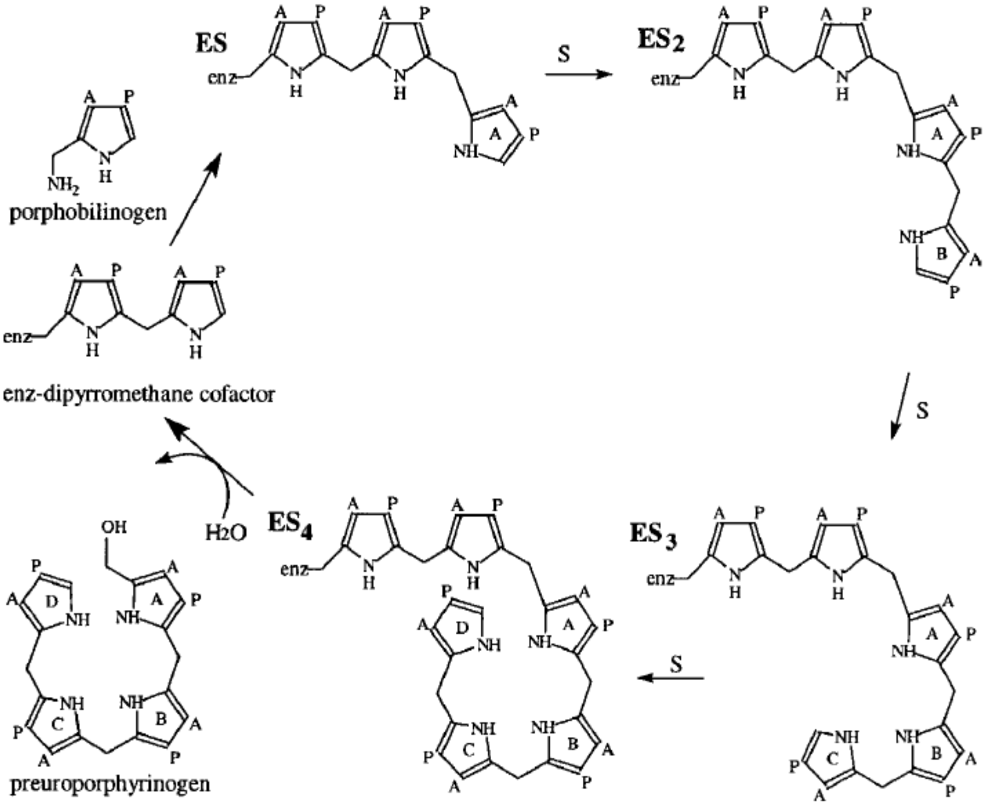


Figure 3.17 - The catalytic cycle of HMBS. The reaction involves the elongation of the enzyme-bound dipyrromethane cofactor by the stepwise addition of four molecules of PBG through ES_1, ES2 and ES3 to give ES4, an enzyme-bound "hexapyrrole". The linear HMB product, also called pre-uroporphyrinogen, is released from ES_4 by hydrolysis, regenerating the unchanged dipyrromethane cofactor. The dipyrromethane cofactor arises initially from the reaction of Cys-261 of the HMBS *apo*-enzyme with pre-uroporphyrinogen, and the resulting ES_2 complex then enters the catalytic cycle. A, –CH2CO2H; P, –CH2CH2CO2H. From Anderson *et al* [66] and Jordan *et al* [74].

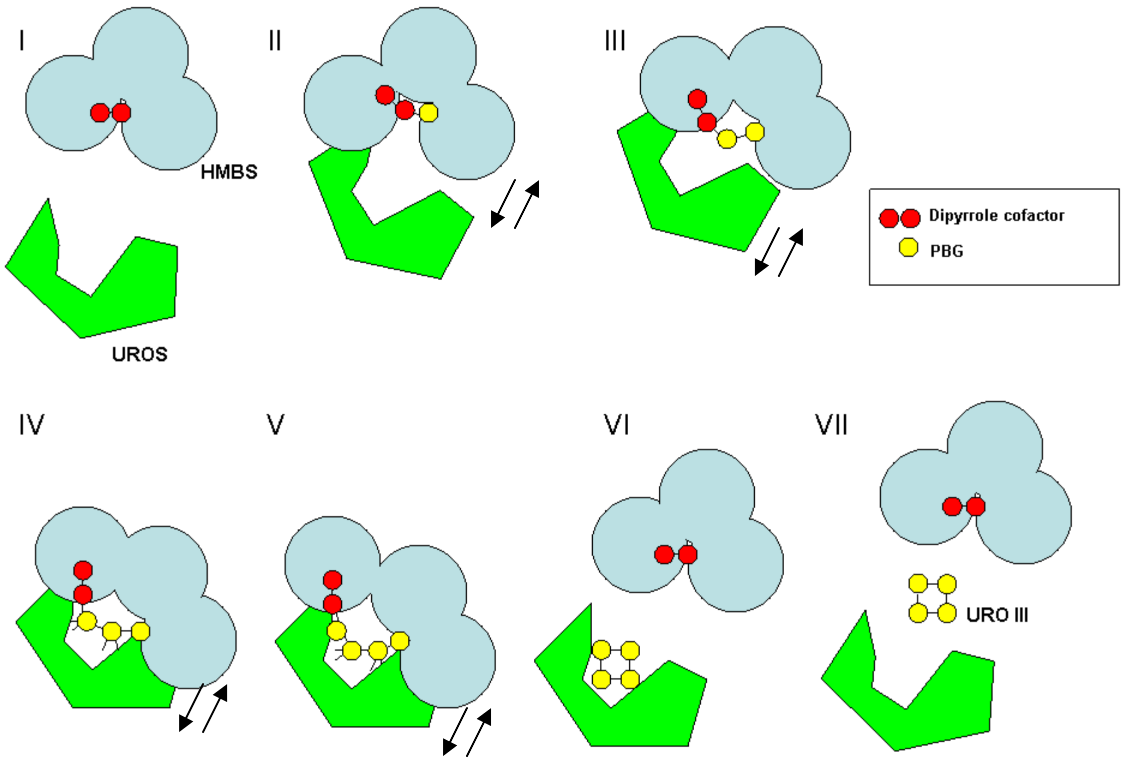


Figure 3.18 - Scheme representing the hypothetical UROS interaction with HMBS enzyme-substrate complexes, as mediated by the tethered growing pyrrole chain. I: UROS and HMBS co-exist freely in the cytoplasm; II-III: ES_1 and ES_2 enzyme-substrate intermediates may interact weakly with UROS; IV: UROS binds ring A and B of the HMBS-bound pyrrole chain, which functions as a tether between both enzymes. VI: the linear tetrapyrrole HMB is released from HMBS, remaining attached to UROS, which immediately cyclizes it to URO'gen III, which is then released into the surround medium (VII). (Globules of HMB represent domain 1-3)

In both experiments, perturbed residues were localized almost exclusively in the cleft region between domain 1 and 2. If UROS was tethered to HMBS during the later stages of HMB formation, we would expect that additional residues of UROS, in areas adjacent to the cleft would, be perturbed by HMBS. Fig 3.19 shows an overlay of the HSQC spectra of UROS titrated with 50, to 200 % HMBS (molar percentage) on the left, and on the right, the spectrum of UROS titrated with 1 to 13% URO'gen III. Mismatched resonance colors indicate line broadening of perturbed residues. Careful analysis of the spectra did not reveal any UROS resonance perturbed by HMBS that was not also perturbed by URO'gen III.

Thus, all of the perturbed residues in the titration with HMBS forms C and D can be explained by the formation of URO'gen III during the experiment (15 μM of URO'gen III), not supporting an interaction of UROS with HMBS, either directly or mediated by the substrate, under the tested experimental conditions. This however does not exclude an *in vivo* complex involving other proteins acting as a scaffold.

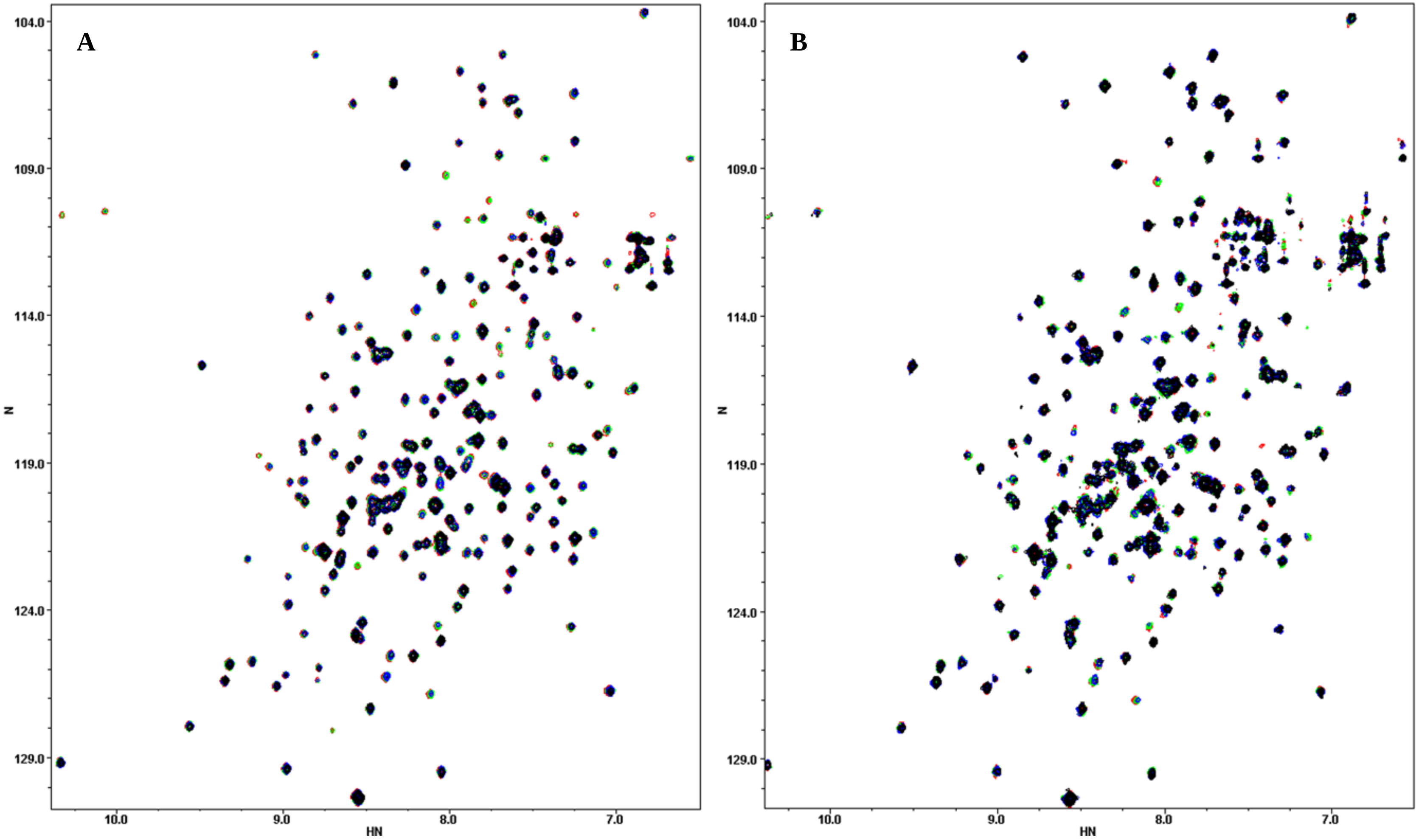


Figure 3.19 - [^{15}N,^{1}H]-HSQC spectra of labeled URO-synthase mixed with HMBS (A) or URO'gen III (B). **A** - [^{15}N,^{1}H]-URO-synthase (resonances shown in red) was titrated with increasing concentrations of HMB-synthase (green – 2:1, blue -1:1, black 1:2, UROS:HMBS-ES3 molar ratio); **B** - [^{15}N,^{1}H]-URO-synthase (resonances shown in red) was titrated with increasing concentrations of URO'gen (green - 1%, blue – 5%, black – 13%) .

URO-Synthase Active Site Mapping

Prediction of the putative Active Site of URO-Synthase by Docking Analyses - *In silico* docking to the A-chain crystal structure of human URO-synthase [31] was performed as described in Methods using three alternative energy-minimized structures for URO'gen III (Fig. 3.20A) and one for the linear tetrapyrrole, N_D-methyl-1-formylbilane (Fig. 3.20B*)*, both competitive inhibitors of substrate binding [41,75]. Each structure was used in separate docking experiments which held the protein and the pyrrole bridge methylenes in a rigid conformation while allowing conformational change of the ligand side-chains to robustly assess local possibilities for energy minimization of ligand binding.

The lowest energy clusters for all three URO'gen III structures docked in the cleft between URO-synthase domains 1 and 2 and the best three were centered above Asp-8, with ring D left of the hinge β-sheets, closest to residue Arg-148. (Fig. 3.21 A). The cluster with the most conformers was the "top ranking" cluster whose lowest energy conformer had a $\Delta G_{docking}$ = -16 kcal/mol, a $\Delta G_{binding}$ = -7.2 kcal/mol, a predicted binding constant (K_i) of 5.7 μM, and fit tightly into the domain 1 interface of the cleft (Fig. 3.22). Docking analyses using the crystal structure’s chain B did not provide as well-clustered results, consistent with the suggestion that alternative conformations of the enzyme may promote product release [31].

While most of the enzyme’s surface was negatively-charged, the cleft region occupied by the negatively-charged ligand was largely positive. Residues with one or more atoms within 4 Å of this URO'gen III conformer included Asp-8, centered below the hydrogens on the tetrapyrrole nitrogens, Tyr-19 near the tetrapyrrole B-ring propionate, Ala-9, Ser-37 and Phe-38 near the A-ring propionate, Leu-36 near the methylene bridge between rings A and D, Glu-11 and Leu-36 near the B-ring propionate, Arg-65 and Thr-170 near the A-ring acetate, Arg-148 and Thr-170 near the D-ring acetate, Phe-196, Thr-227 and Thr-228 near the C-ring propionate, Ser-197 and Ser-199 near the D-ring propionate, and Pro-246 near the C ring acetate. Likewise, the energy-minimized NMF-bilane structure with fixed torsions of the A to C ring methylene bridges docked with its top 10 clusters in positions similar to those of URO’gen III (Fig. 3.21*B*). Most clusters were nearly congruent with the position of the best URO'gen III conformer, but rotated clockwise one pyrrole unit. While the three top-ranked clusters had only 2-5 members each, the fourth top-ranked cluster had 19 members and was very similar in position to the seventh-ranked cluster, which had 10 members (Fig 3.21B). The $\Delta G_{docking}$ of the best conformer of these two clusters was -15 kcal/mol and the $\Delta G_{binding}$ was -4.6 kcal/mol corresponding to a predicted K_i of 400 μM (about 100-fold weaker than that predicted for the best URO'gen III conformer).

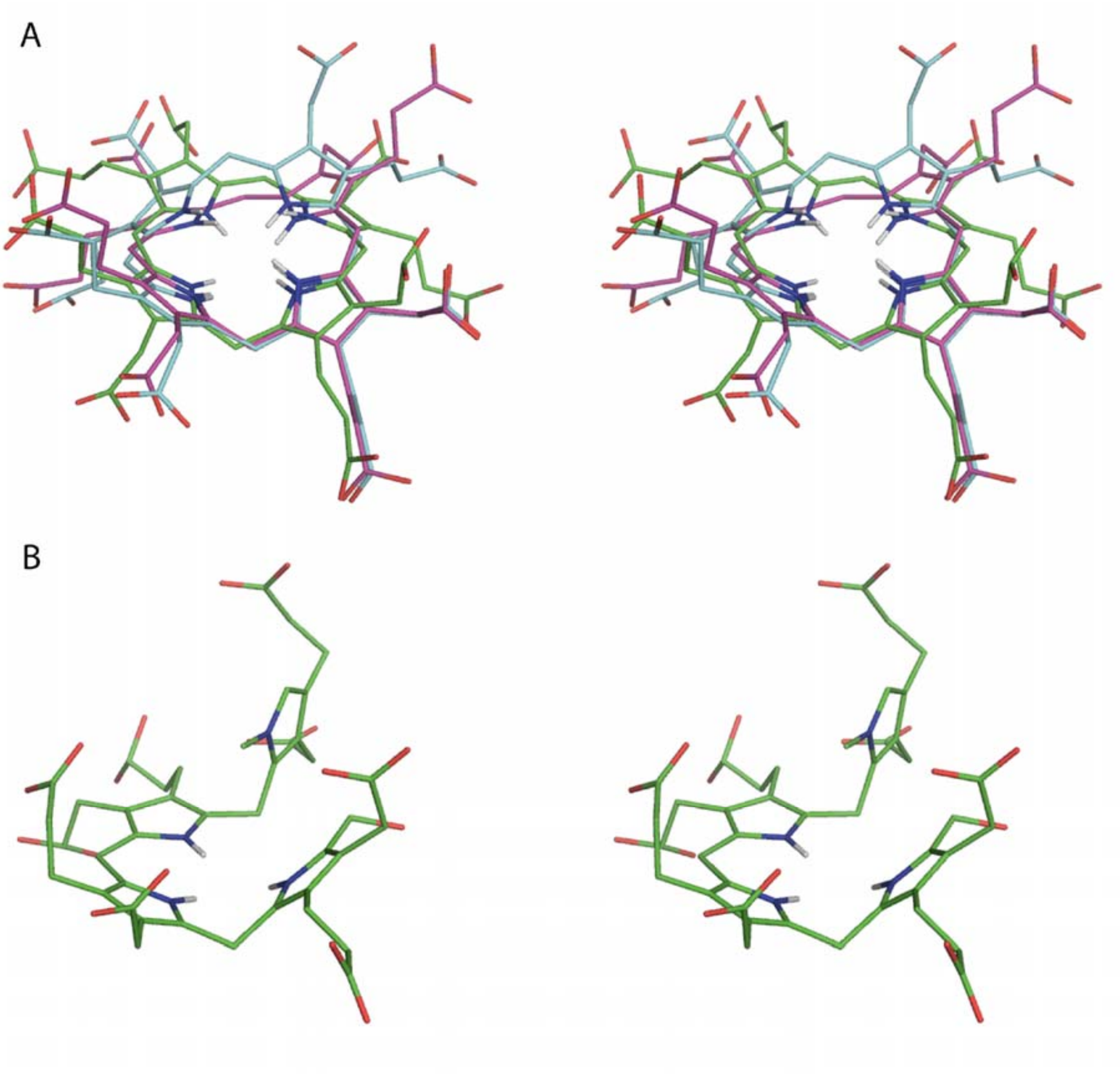


Figure 3.20 - Energy minimized structures used as input for AutoDock. **A -** Three alternative conformers of URO'gen III (see Materials and Methods); **B -** N_D-methyl-1-formylbilane. This figure was generated with the computer program Pymol (64).

Residues within 4Å of these two conformers included the same residues as found for URO'gen III docking except for absent Ser-37 and Phe-38 contacts and added contacts including Lys-7 near the A-ring propionate, Thr-62 and Pro-64 near the D-ring propionate, Ser-63 and Tyr-168 near the C-ring propionate, and Gly-200 near the B-ring propionate. Thus, the docking results for both the URO'gen III and NMF-bilane supported the location of the center of the tetrapyrrole ring directly above Asp-8 with the peripheral acid groups contacting both domain 1 and 2 residues lining the surface of the cleft between the domains.

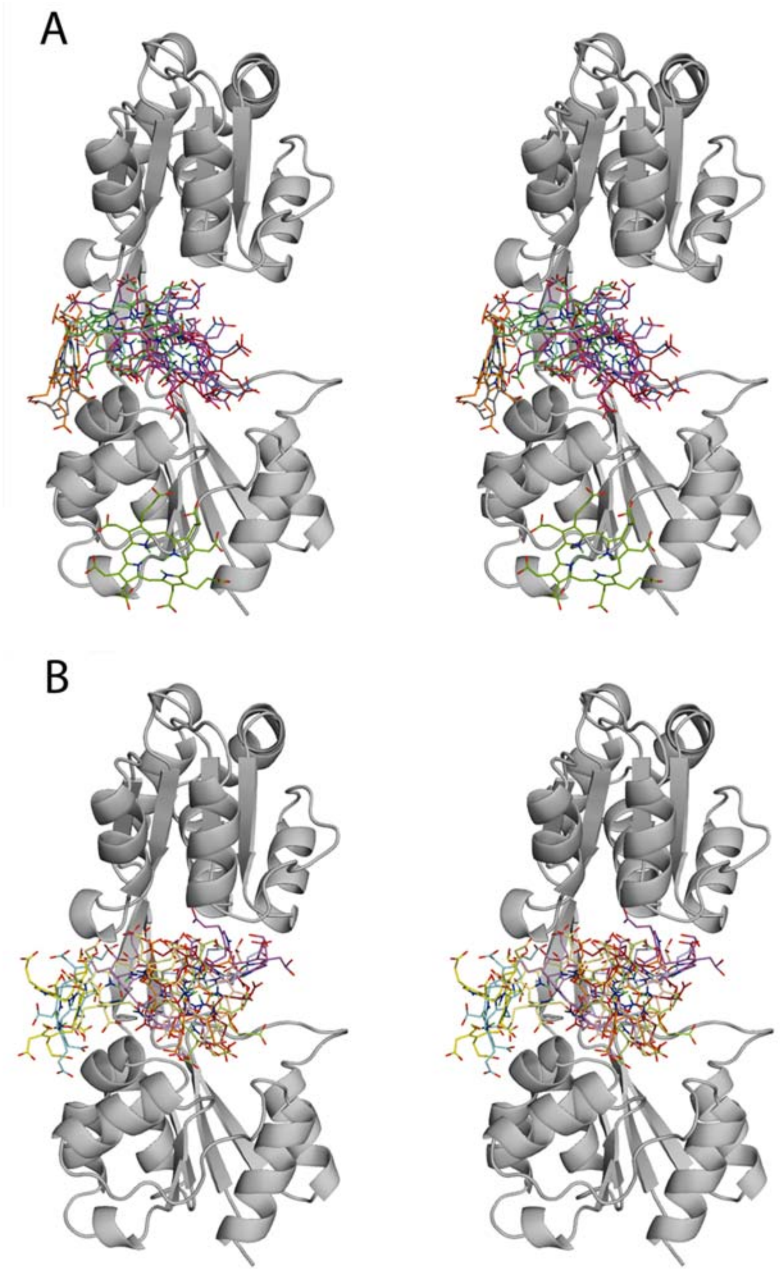


Figure 3.21 - Crystal structure model of URO-synthase docked with ligands. **A** - URO'gen III; **B** - N_D-methyl-1-formylbilane**.** The energy minimized ligands were docked to the crystal structure [31] of URO-synthase (see Materials and Methods). Each cluster of docked ligands was ranked based on the lowest energy conformer in each cluster. The lowest energy conformer from each of the top 10 clusters, are shown.

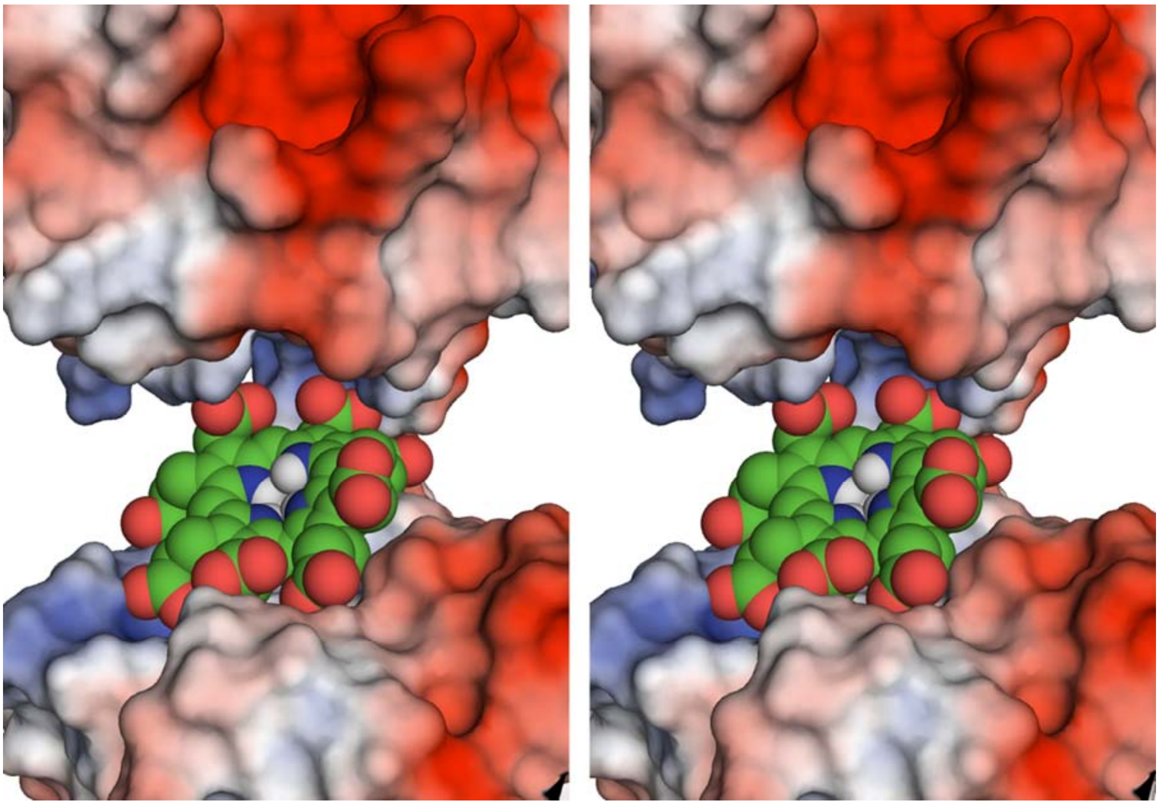


Figure 3.22 – Stereo image of the complex between URO-synthase and the lowest energy docked URO'gen conformer. The URO-synthase surface was colored according to its electrostatic potential, ranging from -4 kT/e (red) to +4 kT/e (blue), calculated with the computer program APBS tools [64]. The docked URO'gen conformer was represented as spheres (atom colors: C = green, O = red, N = blue, H = white).

Identification of the URO-synthase Active Site by NMR Perturbation Mapping - To experimentally determine the location of the active site and to identify active site residues, chemical shift perturbation studies with [^{15}N,^{2}H]-labeled URO-synthase were performed by titrating [^{15}N,^{2}H]-labeled URO-synthase with unlabeled URO'gen III or NMF-bilane in 10% D_2O/90% H_2O. These studies revealed that specific assigned backbone resonances in the [^{15}N,^{1}H]-HSQC spectrum underwent line-broadening in a ligand-concentration dependent manner, indicating a slow to intermediate exchange rate of binding. The changes in peak intensities were determined as a measure of line-broadening to map the ligand binding sites on the protein. An overlay of a region of unperturbed and increasingly perturbed URO-synthase [^{15}N,^{1}H]-HSQC spectra (with titration by URO'gen III) is shown in Figure 3.23*B*. To quantitate the extent of perturbation, the slopes of peak intensity versus ligand concentration (Fig. 3.23C) for each resonance were plotted (Fig. 3.23*A* and *B*). Of the 235 assigned resonances, 194 could be monitored for changes in peak intensities, the remaining 46 overlapped with each other or were too weak to quantitate. Titration with URO'gen III induced line broadening in 88 resonances and NMF-bilane titration resulted in 94 line-broadened resonances.

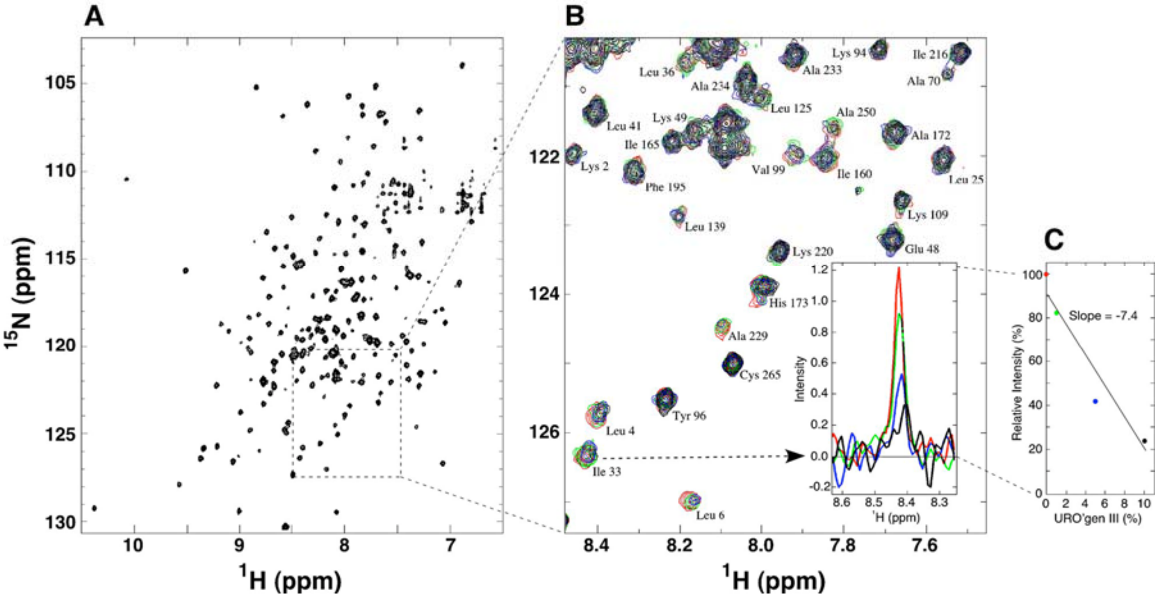


Figure 3.23 - Perturbation of the [^{15}N, ^{1}H] HSQC spectrum of URO-synthase by titration with URO'gen III. **A** - Spectrum of free URO-synthase. **B** - Detailed region of panel A showing the perturbation (line-broadening) of URO-synthase resonances in overlaid spectra collected upon titration with URO'gen III as described in Methods (red = free enzyme; green = 1% URO'gen III [molar ratio of URO'gen III:URO-synthase as a percentage; e.g. 1% = 3 mM URO'gen III]; Blue = 5%; Black = 10%). The inset shows the 1D intensity spectrum of the line-broadened Ile-33 residue using the same color code. **C** - Extent of line-broadening estimated for residue Ile-33 as the slope of a plot of resonance intensity vs ligand % as described in Methods.

The residues most strongly perturbed by URO'gen III involved 18 backbone resonances in the [^{15}N,^{1}H]-HSQC spectrum (Fig. 3.24*A*). These residues included Leu-6, Lys-7, Asp-8, Ile-33, Val-35, Leu-36, Thr-62, Ser-63, Arg-65, Gly-100, Thr-103, Arg-148, Leu-151, Phe-196, Ser-197, Gly-200, Ala-229, and Leu-251. Of these, Leu-6, Lys-7, Ile-33, Leu-36, Thr-62, Ser 63, Arg-65, Gly-100, Thr-103, Phe-196, Ser-197, Gly-200, and Leu-251 were highly conserved (92-100% identity in 15 eukaryotes (see Appendix 3.1). The locations of these residues in the URO-synthase 3D crystal structure [PBD code 1JR2, Chain A [31] are indicated in red in Figure 3.25. These strongly perturbed resonances were all from residues at or near the surface forming the major crevice between the two globular domains of the enzyme, and thus, map the active site. Additional residues that showed moderate line-broadening (Figs. 3.24A and 3.25, indicated in yellow), tended to be in the interior of the enzyme and presumably represented residues experiencing environmental alterations due to conformational change of the protein upon ligand binding. Of note, most of the strongly perturbed residues occurred in clusters that contained other adjacent strongly and/or moderately perturbed residues (Fig. 3.24).

The linear substrate analogue, NMF-bilane, strongly perturbed the same residues that were strongly perturbed by URO'gen III, except for Thr-103 (moderately perturbed) and Leu-251 (weakly perturbed) (Fig. 3.24B).

Approximately 10 times more NMF-bilane was required to give similar extents of perturbation (50% NMF-bilane was equivalent to 5% URO'gen III). Notably, the clusters of adjacent perturbed residues were nearly identical for both URO'gen III and NMF-bilane (compare Figs. 3.24A and B), the main differences being the relative intensities of individual cluster members.

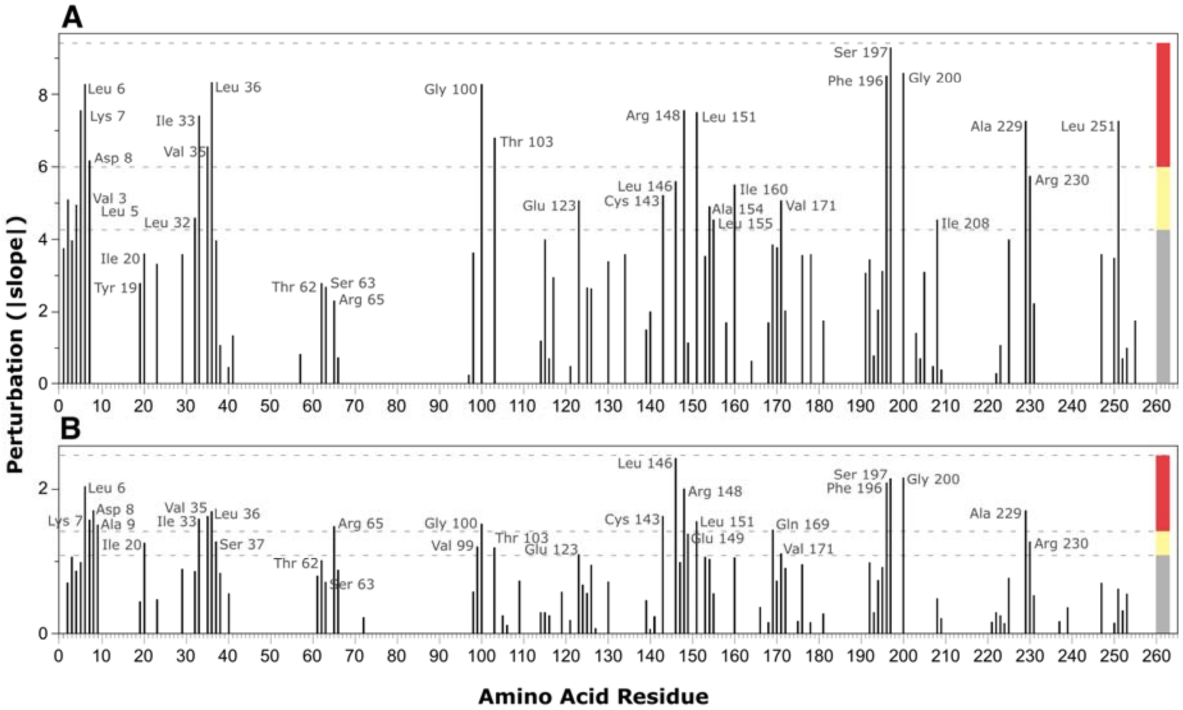


Figure 3.24 - Perturbation of URO-synthase resonance intensities by **A** - URO'gen III and **B** - N_D-methyl-1-formylbilane. The absolute values of the slopes of the intensity vs percent ligand plots describe the extent of line-broadening and were calculated as described in Methods and as shown in Fig 3.23. The most strongly perturbed residues were between the slope values indicated by the red bar while moderately perturbed

residues were defined by the yellow bar boundaries. Weakly perturbed residues are demarcated by the grey bar.

Although the peak intensities for the clusters that included Tyr-19, Ile-20, and Arg-65 were lower than the yellow-coded cut-off values in Figure 3.24A, they were yellow-coded in Figure 3.25 since they represented unique clusters and were significantly perturbed by NMF-bilane. Of these, Tyr-19 and Arg-65 were highly conserved in eukaryotes. Some differences in relative perturbation intensities caused by the closed-ring URO'gen III and the linear NMB-bilane ligands were observed, possibly due to their structural differences. Residues which were more perturbed by URO'gen III were Leu-151, Val-171, Ile-208 and Leu-251 while those more perturbed by NMF-bilane were Ala-9, Arg-65, Leu-146 and Gln-169.

Notably, the docking results (based on the 3D crystal structure of URO-synthase) for the lowest energy URO'gen III conformer coincided with the URO'gen III perturbation results which localized the active site to a specific cleft region between domains 1 and 2 (Fig. 3.26). Note that of the 18 most strongly perturbed residues, 12 (Lys-7, Asp-8, Tyr-19, Leu-36, Thr-62, Ser-63, Arg-65, Gly-100, Thr-103, Arg-148, Phe-196, and Ser-197) form surfaces directly above or below the AutoDock-predicted location of URO'gen III binding. The domain 1 surface and domain 2 Arg-148/65 surfaces are in close proximity to the ligand while the domain 2 surface formed by Thr-62, Ser-63, Gly-100, and Thr-103 is more distant from the ligand in the crystal structure. The lowest energy N_D-1-methyl-formylbilane conformer docked in a similar conformation URO'gen III, except that ring D was positioned vertically and able to reach the Ser-63 region (Fig 3.27). Therefore, the NMR perturbation data predicts that the solution structure of the active site region for the ligand-bound form is more closed than the crystal structure shown in Figure 3.26. Contacts were also predicted, but could not be confirmed, for unassigned residues Pro-64, Pro-198, Ser-199, Thr-227, Thr-228, and Pro-246. Thus, the NMR perturbation experiments confirmed the results of the *in silico* docking studies and mapped the location of the URO-synthase active site.

Intermolecular NOEs between the NMF-bilane inhibitor and the reverse labeled $^{13}C/^{1}H$-methyl protons of valine and leucine, and the $^{13}C/^{1}H$-δ-methyl isoleucine protons of otherwise deuterated URO-synthase were evaluated for confirmation of the results of the resonance perturbation studies. Unfortunately, of the 63 methyl-labeled residues, only one (Leu-36) was predicted to be within 5 Å of the carbon-attached protons of the inhibitor in the *in silico* docking studies (data not shown). However, no NOE was observed for this residue. Future NMR solution-structure studies may clarify this finding.

In addition to the nitrogen HSQC, we recorded a carbon HSQC on a sample whose methyl side chains of the 59 valine, leucine and isoleucines residues were ^{13}C-labelled. Of those, 25 were perturbed (Fig. 3.28). The pattern of residues with perturbed side chains is similar to that of backbone amides, in which most perturbations occur in domain 1, and on the left side helices of domain 2 (view relative to Fig. 3.28).

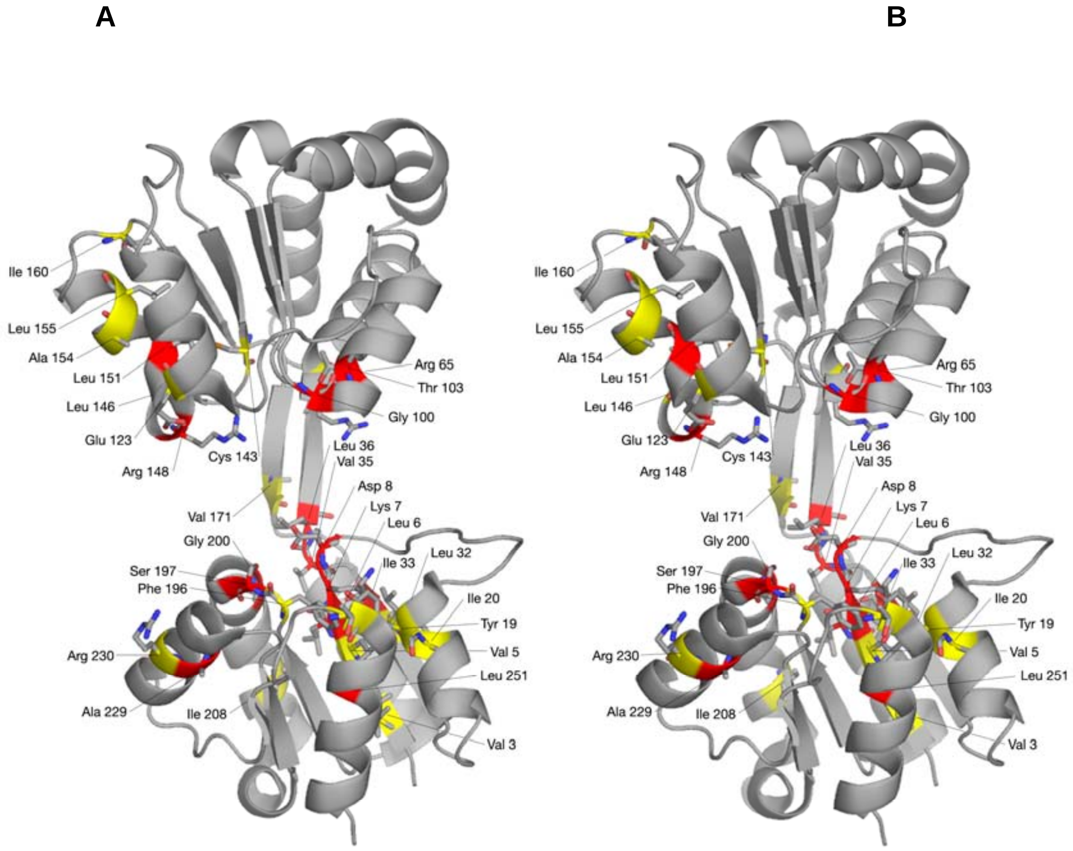


Figure 3.25 - Crystal structure of URO-synthase showing the residues that were moderately to strongly perturbed during titration with **A** - URO'gen III or **B** - N_D-methyl-1-formylbilane. The alpha carbon color was assigned as in Fig 3.24.

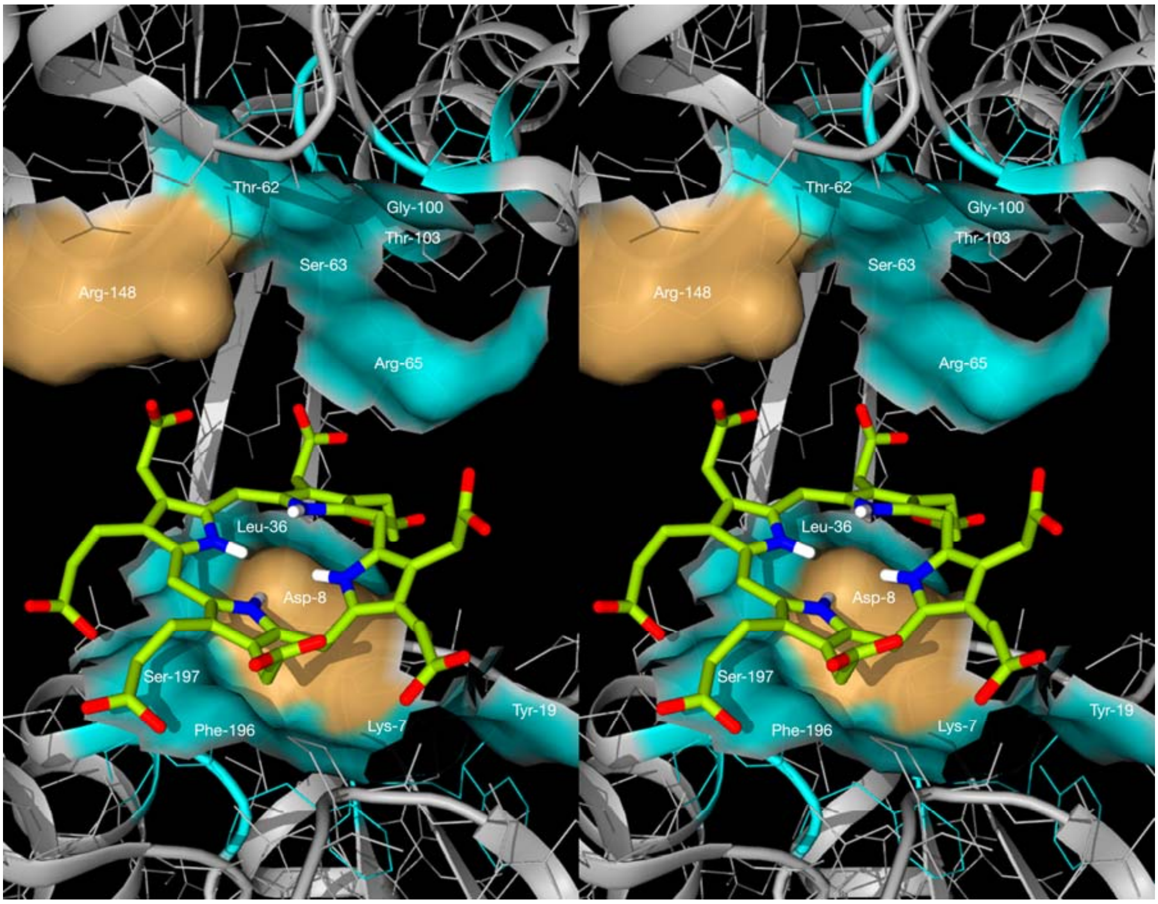


Figure 3.26 - Localization of the URO-synthase Active Site. The cleft region between domains 1 and 2 of the 3D crystal structure is shown in stereo, highlighting the surfaces formed in the cleft between domains 1 and 2 by residues whose NMR resonances were highly-perturbed by URO'gen III binding. The surfaces for residues exhibition 92-100% sequence identity in eukaryotes shown in Appendix 3.1 is teal colored and residues with 54 to 62% identity in sand color.

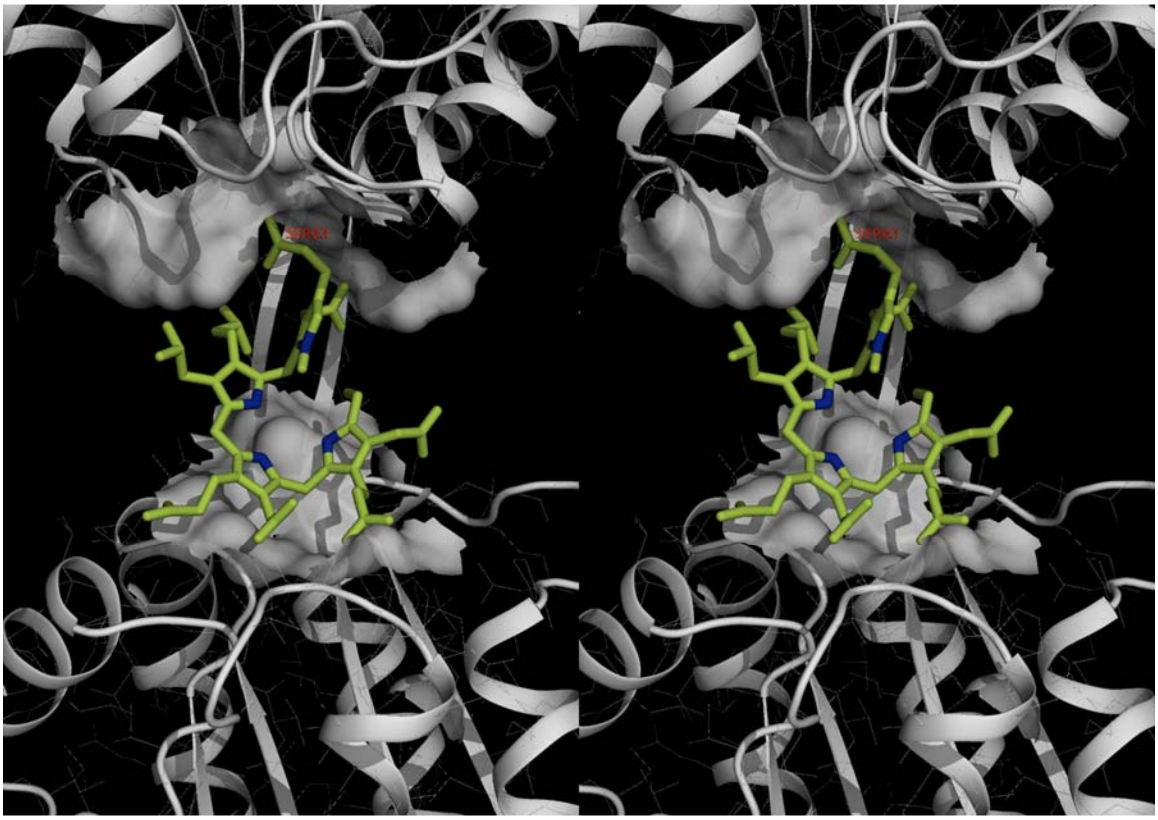


Figure 3.27 - Localization of the URO-synthase Active Site. The cleft region between domains 1 and 2 of the 3D crystal structure is shown in stereo. highlighting the surfaces formed in the cleft between domains 1 and 2 by residues whose NMR resonances were highly-perturbed by N_D-1-methyl-formylbilane binding. Residue Ser-63 is labeled, to compare the ligand's rind D proximity to this region, in contrast to the URO'gen III conformation shown in Fig 3.26.

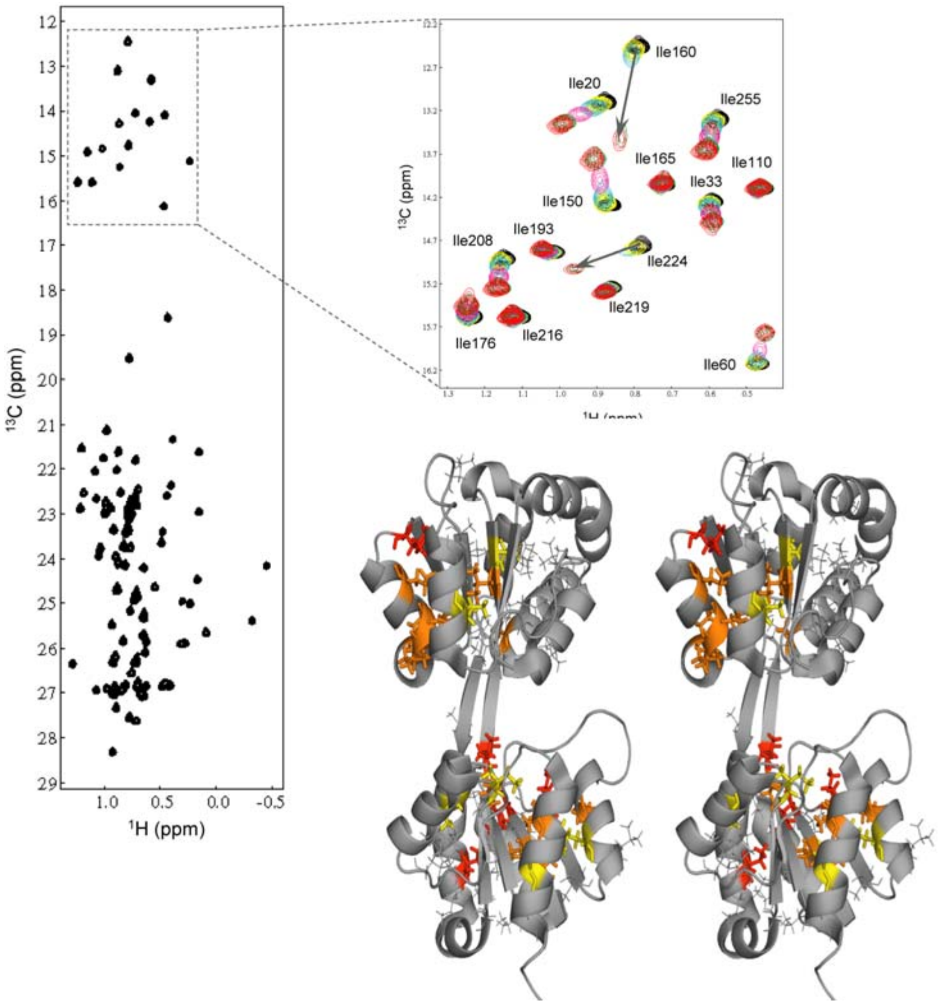


Figure 3.28 – Perturbation of UROS by the inhibitor N(D)-methyl-1-formylbilane. **A** - UROS was titrated with the inhibitor and an [^{13}N,^{1}H]-HSQC spectrum of a {I(δ-$^{13}CH_3$),L($^{13}CH_3$,$^{12}CD_3$),V($^{13}CH_3$,$^{12}CD_3$)}-U-[^{15}N,^{13}C,^{2}H]-labeled UROS sample was recorded for the free enzyme (left) and at each titration point. The spectra were superposed and color coded according to the inhibitor concentration (right): Black: free enzyme; yellow – 13% (molar ratio) of inhibitor; cyan – 27%; magenta – 50%; green – 107%; red – 200%. Arrows indicate the new position of the resonance in the spectra collected upon titration. **B** - Mapping of perturbed residues in the NMR-determined 3D structure of UROS (Titration data not shown). Color coding reflects the level of perturbation (high, medium, low) as described in the text.

Mutagenesis of Specific Active-site Residues – table 3.6 shows the mutagenesis of selected active-site residues, predicted to be important for substrate binding by the previous docking studies. These is only a partial list, and future work will include other residues to be mutated, such as Thr-103, and combined residues with possible redundant function Ser-197/Ser-199, and Ser-62/Thr-63.

Mutation	URO-synthase activity[a]
Wild type	100%
K7A	39% (*k*m = 3.33 μM)
D8N	20%
D8H	16%
D8E	60%
Y19F	80%
Y19C	1%
L36A	16%
R65A	<1%
R65E	<1%
R65Q	<1%
E123A	24% (*k*m = 0.67 μM)
R148A	114% (*k*m = 1 μM)
R148Q	107%
Y168F	49%

[a] Expressed as percentage of the wild type control

Table 3.6 - Mutant URO-synthase activities

Energy Minimization and Molecular Dynamics (MD) Simulation of the UROS-URO'gen III Complex Predicted by AutoDock – AutoDock only treats the flexibility of the ligand, while the protein is kept rigid, and consequently imposes limitations on the number of potential protein-ligand conformations. This could be particularly problematic in the case of URO-synthase, since the flexibility of the domains has been suggested [31]. Therefore we sought to search more deeply the allowed conformational space by molecular dynamics (MD) simulations, which allow flexibility of both the protein and the ligand, while simulating the time-depended behavior of the complex by solving Newton's equation of motion. Simulations were carried with the computer program NAMD. The enzyme-ligand complex was solvated and was simulated using periodic box conditions (no water molecules are lost during the simulation), at constant pressure and temperature. An energy minimization step was initially performed to optimize any distorted geometries and to set the atom velocities to 0, followed by a 6 ns molecular dynamics simulation at 37° C, 2 ns at 45° C and 2 ns at 55° C. Fig 3.29 shows the structure of the complex after the 10 ns simulation. After several nanoseconds of simulation, the complex conformation is very similar to that predicted by AutoDock. Overall, the deviation during the simulation was below 1 Å (data not shown). The simulation resulted in minor adjustments of the ligand's torsion angles which permitted a better fitting in the protein's active site cleft, but URO'gen III remained centered above Asp-8. The acetate and propionic side chains were slightly re-oriented to make contact with Tyr-19, Ser-63, Arg-65, and Arg-148. Of note, in a separate simulation using a conformation of the ligand docked slightly off-center of Asp-8 resulted in the ligand leaving the active site within 1 ns of simulation.

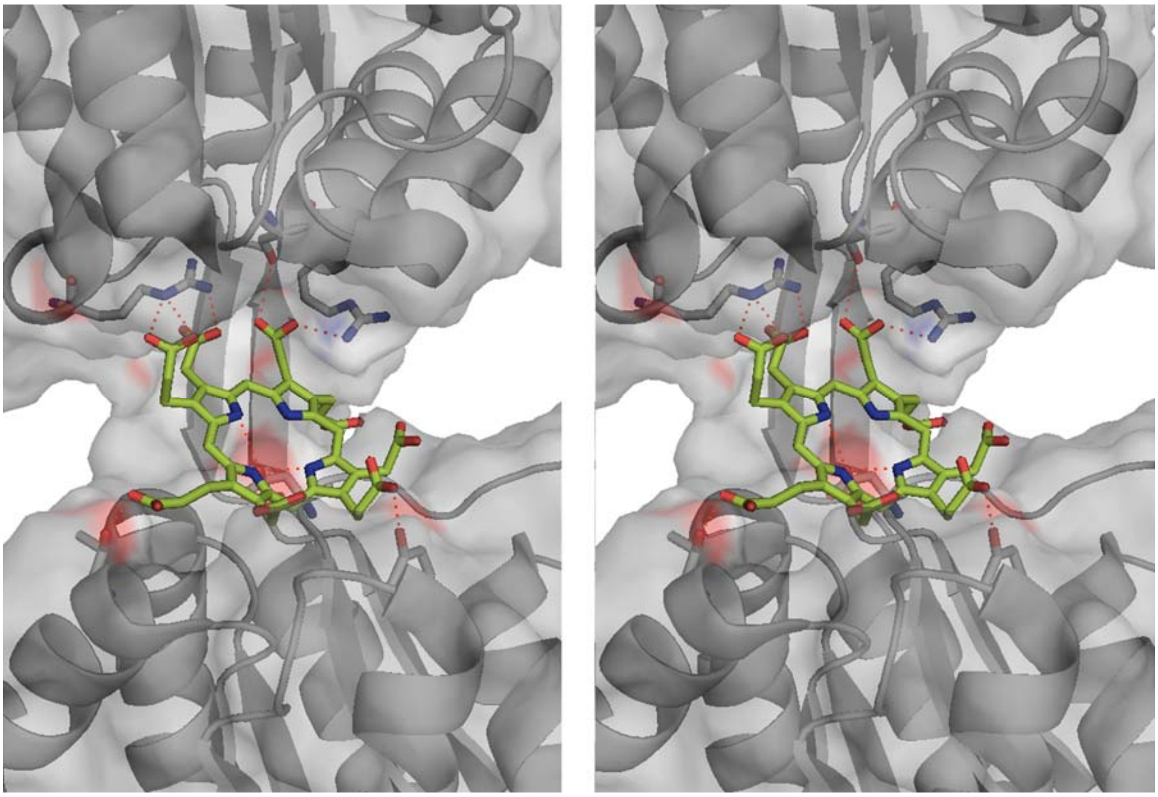


Figure 3.29 - Optimization of the URO'gen III binding site by molecular dynamics simulation. The UROS-URO'gen III complex conformation predicted by AutoDock was energy minimized and the time dependent behavior was simulated for 6 ns at 37° C, followed by 2 ns at 45° C and 2 ns at 55° C. The conformation at the end of the 10ns simulation is shown, and does not differ significantly from the AutoDock conformation, with only a few side chain adjustments to optimize hydrogen bonds (shown as dashed red lines)

A simulation of the free enzyme, or the enzyme-URO'gen III complex (crystal structure coordinates) did not result in a significant alteration of the protein conformation. However, molecular dynamics only simulates vibrations in the nanosecond range, while large domain movements, as shown in figure 3.11 by normal mode analysis, may take place in the micro and millisecond scale and the short time of MD simulations might not capture these large moments. Furthermore, the opening/closing of the cleft might be triggered by the presence of ligand, or the extremely open conformation seen in the crystal structure might be due to crystallization conditions and not be a state usually sampled in solution. Note that a crystal structure is available for URO-synthase of *Thermus thermophilus*, in which the cleft is extremely closed (PDB: 1WCW). Our NMR titration experiments showed that residues in domain 2 are perturbed by URO'gen III. Since the cleft in the crystal structure is too large to make contacts with URO'gen III with both domains, the cleft size must be smaller in solution to bind tightly to the ligand.

Simulation of the UROS-URO'gen III complex using the NMR solution structure – The 3D structure determined by NMR has a cleft between domain 1 and 2 that is too closed to accommodate the URO'gen III without clashes between the protein and ligand atoms, and therefore could not be used as an input structure for AutoDock. One alternative could be to use "open" conformations resulting from the normal mode analyses. However, energy minimization of the solution NMR structure opened the cleft just enough to fit the ligand coordinates of the best x-ray crystal structure-URO'gen III docked conformer, with only minimal atom clashes. The ligand coordinates were transferred to the NMR solution structure and this complex was energy minimized in order to optimize its geometry and avoid atom clashes between the protein and the ligand. A subsequent 4 ns molecular dynamics simulation was performed. Fig. 3.30A shows the conformation of the UROS-URO'gen III complex at the end of this simulation. Fig 3.30B shows the deviation of the protein and ligand structures from the initial input structure, during the 4ns MD simulation. After an initial equilibration, the complex reached a stable conformation at about 1ns. However, it's possible that a longer simulation (in the order of tens of ns) could significantly change this conformation. During the simulated period, the cleft size increased to accommodate the ligand (compare the structures before and after simulation in Fig 3.31). In the resulting conformation, the ligand was in contact with residues in both domain 1 and 2, including the strongly perturbed Gly-100, predicted to be an important contact residue by the previous experiments. The ligand was displaced slightly from its initial position above Asp-8, and localized in the region between Asp-8/Phe-196. The contacts predicted by docking experiments and those observed in the NMR titrations were maintained, including contacts between the ligand acetic/propionic side chains and Lys-7, Arg-65, Gly-100, Arg-148, Phe-196, Ser-197, and Ser-199. Therefore, a simulation with the NMR structure better explains the perturbation of residues in both domains.

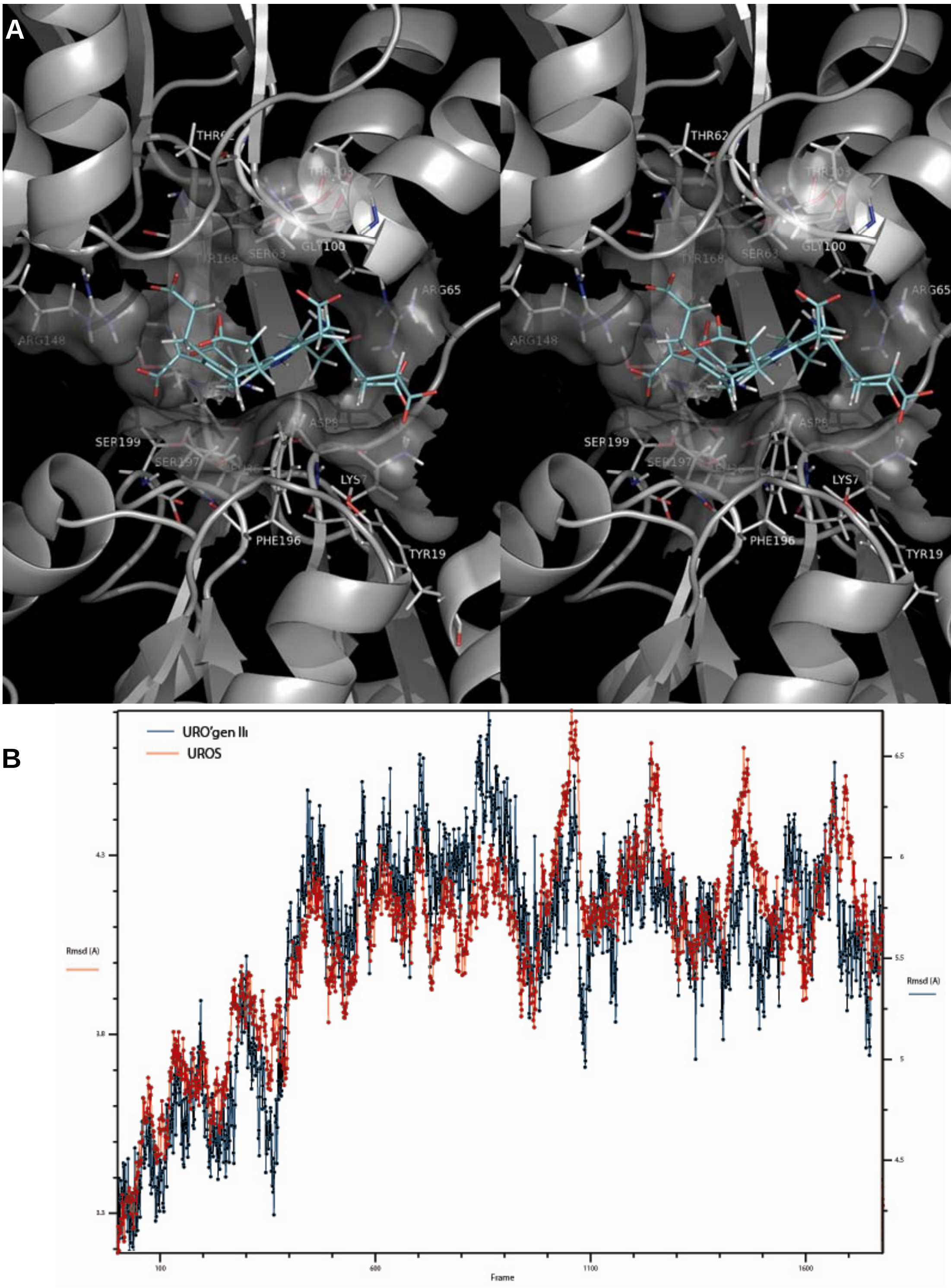


Figure 3.30 - MD simulation of the URO'gen III-UROS (NMR) complex. The Ligand coordinates of the complex predicted by AutoDock (shown in Fig 3.26) were used as input, as described in the text, and the complexed structure was simulated for 4ns. **A** - Post-simulation conformation. The ligand is represented as colored sticks, while the protein (gray) backbone is represented as a cartoon, and the active site residue's side chains are shown as sticks. A transparent represententation of the protein surface is shown for better

visualization of the active site cleft space. **B** - R.m.s.d of the molecular dynamics trajectory. Snapshots (frames) were collected at approximately 2.2 ps intervals. Note that the reference structure for r.m.s.d. calculation was the pre-energy minimization conformer, and for that reason, there is a 2 Å deviation in the initial frame. After an initial equilibration, the complex reached a stable conformation at about 1ns.

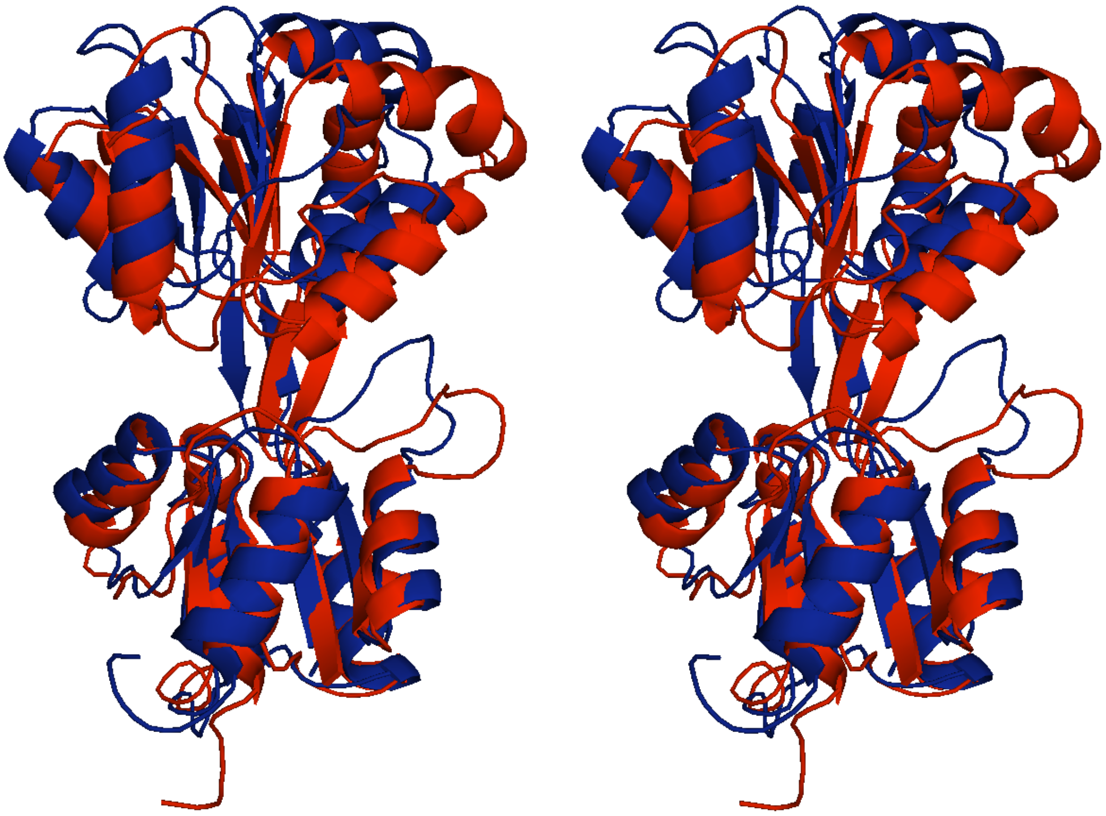


Figure 3.31. Comparison of the NMR solution structure before (red) and after (blue) simulation with the Docked URO'gen III (ligand has been removed from this illustration). The structure post-simulation has a wider cleft opening, since the protein adjusted the interdomain distance to accommodate the ligand.

DISCUSSION:

NMR studies of recombinant human URO-synthase were undertaken to determine its backbone and side-chain resonance assignments, to identify its active site in solution, and to investigate the possible interaction of the cytosolic heme biosynthetic enzymes in a biosynthetic complex or "metabolon". For these studies, milligram quantities of recombinant human HMB-synthase, URO-synthase, and URO-decarboxylase were expressed and purified. Importantly, recombinant human URO-synthase was stable at pH 7.45 for over two weeks at 30° C which was required for the NMR studies.

Although the ~29 kDa human URO-synthase was relatively large for NMR structural analysis, the complete assignment of all backbone carbons, all but one backbone alpha hydrogen, 94% of the backbone $^{1}H_{N}$ and backbone non-proline ^{15}N resonances, and 85% of the non-proline side-chain resonances (BMRB accession number 7242) was achieved by use of an 800 MHz NMR spectrometer with a cryoprobe. The URO-synthase backbone resonance assignments permitted localization of the enzyme's active site for binding of the competitive inhibitors URO'gen III and NMF-bilane, and studies of potential protein-protein interactions with HMB-synthase or URO-decarboxylase by 2D [^{15}N,^{1}H]-HSQC resonance perturbation of URO-synthase residues.

Previous physical and kinetic evidence suggested that HMB-synthase and URO-synthase were physically interactive in a complex, and possibly, that the four cytosolic heme biosynthetic enzymes were efficiently organized in an enzyme complex or metabolon [25-27]. However, our NMR perturbation studies did not find evidence for complex formation when URO-synthase was incubated with equimolar concentrations of either holo-HMB-synthase (the "A", or "ES" form containing only the dipyrrole cofactor) or URO-decarboxylase (Fig. 3.14), or with enzyme-substrate intermediates B through D. Thus, these NMR studies did not support the occurrence of a URO-synthase complex with HMB-synthase or URO-decarboxylase. It is still possible that URO-synthase could bind to the nascent linear tri- or tetra-pyrrole HMB-synthase/substrate intermediates directly, although 3D structure studies of HMB-synthase indicated that the dipyrrole cofactor, and perhaps the enzyme-substrate intermediates, are buried in the structure [68, 76-77]. Since the *in vitro* concentrations of these cytosolic heme biosynthetic enzymes were at least 150-fold higher than those in liver [78], it is unlikely that these proteins interact even weakly *in vivo.*

Previous crystallographic studies suggested that the URO-synthase active site was in the cleft between the two major domains based on partial homology to another porphyrin-binding protein and the distribution of conserved residues [31]. However, since it was not possible to co-crystallize the enzyme with substrate analogues, the inferred localization could not be confirmed or refined to specific contact residues. In contrast, the NMR perturbation studies reported here localized the URO-synthase active site to a specific region of the inter-domain cleft and identified specific residues that were strongly perturbed when the isotopically-labeled enzyme was titrated with either its cyclic tetrapyrrole product, URO'gen III, or the linear tetrapyrrole substrate analogue, NMF-bilane (Fig. 3.24). Additionally, weak to moderate resonance perturbations were observed for buried residues that were in close spatial contact, but otherwise distant in the primary structure, indicating that conformational changes occurred upon ligand binding.

Additional evidence in support of the URO-synthase active site residues identified by perturbation mapping included 1) the occurrence of strongly perturbed putative active site residues within clusters of perturbed residues, 2) the congruent *in silico* docking predictions for both URO'gen III and NMF-bilane, 3) the active sites' positive electrostatic surface charge which presumably enabled steering of the negatively-charged substrate to this site, and 4), the high degree of conservation of the predicted active site residues among eukaryotic URO-synthases (Appendix 3.1). Thus, within each cluster, the residues that were the most perturbed, at the cleft surface, most conserved, and necessary for enzymatic activity would be the best candidates for contact residues. As detailed below, these four lines of corroborating evidence identified certain residues that likely comprise the active site.

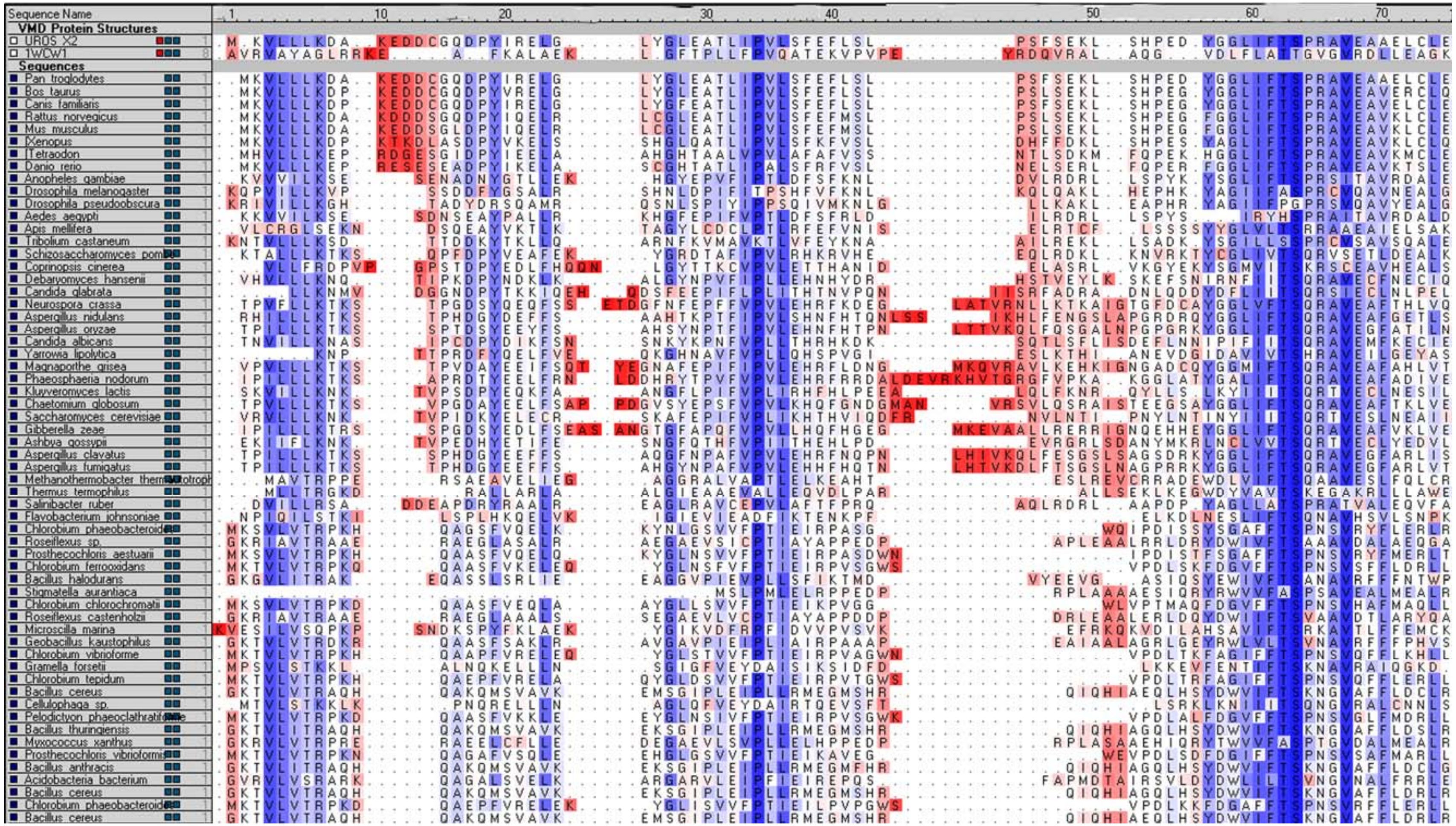
Notably, Leu-6/Lys-7/Asp-8, Tyr-19/Ile-20, Ile-33/Val-35/Leu-36, Thr-62/Ser-63/Arg-65, Gly-100/Thr-103, Arg-148, Val-171, Phe-196/Ser-197, Ala-229/Arg230 and Leu-251 were the most perturbed residues in their respective clusters identified by titration with both ligands. These clusters of perturbed residues presumably represent active site contact regions. Moreover, surface residues Asp-8, Tyr-19, Leu-36, Arg-65.a Arg-148, Phe-196, and Ser-197, were highly likely contact residues in the crystal structure model since they were both the most perturbed residues within their clusters and were within 4Å of both the docked URO'gen III and NMF-bilane ligands.

In addition, the positively-charged residues Lys-7, Arg-65, and Arg-148 were major contributors to the overall positive charge of the active site and all three were perturbed by both ligands. Although the electrostatic surface potential of URO-synthase was mostly negative, the cleft region was highly positively-charged which would presumably favor steering of the ligands to sites experimentally determined by NMR perturbation studies.

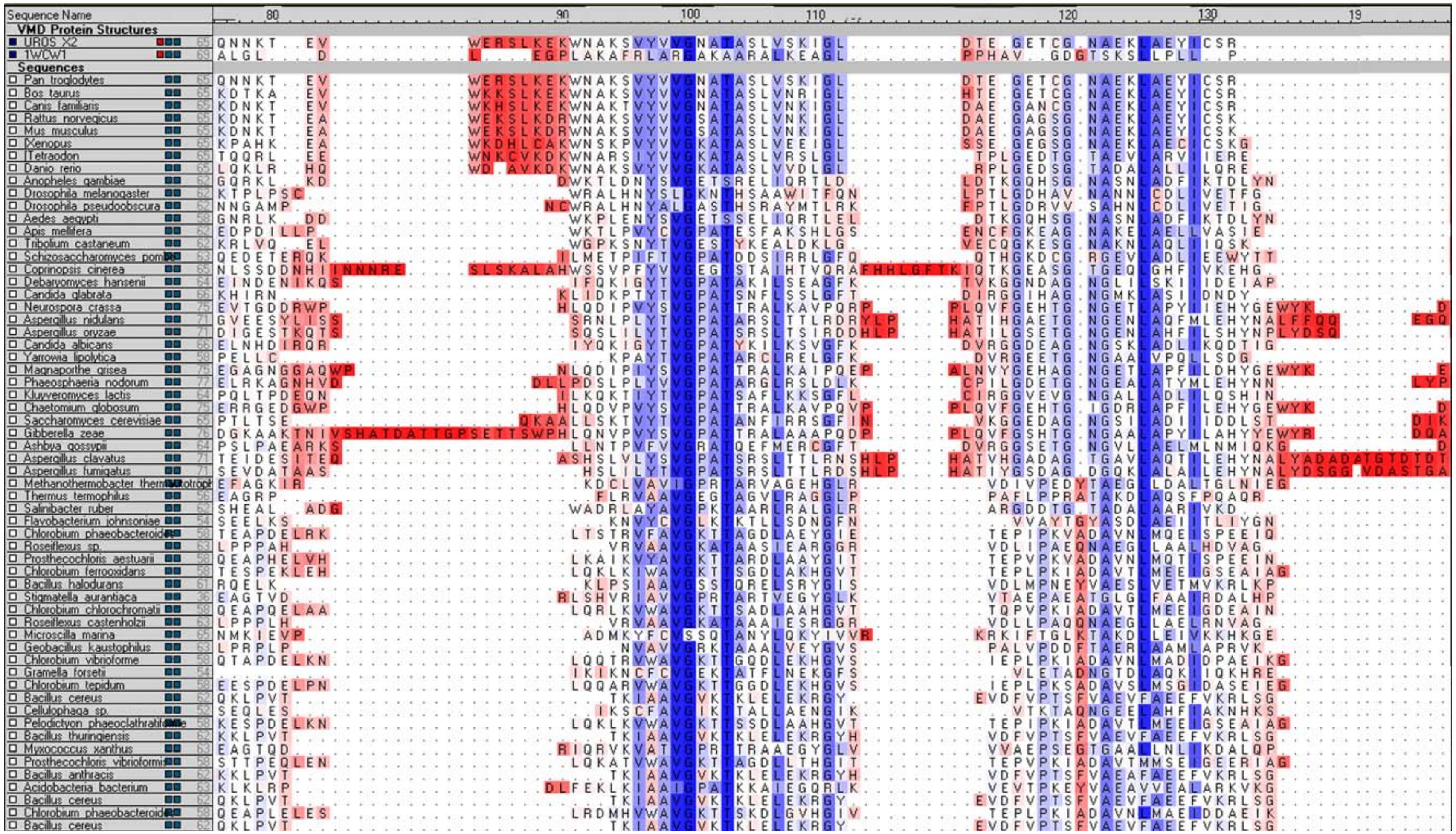
Finally, it is significant that surface residues in the cleft region perturbed by one or both ligands were highly conserved (≥ 90%) in eukaryotes, including yeast. These highly-conserved residues included Lys-7, Tyr-19, Leu-36, Arg-65, Gly-100, Thr-103, Tyr-168, Phe-196, Ser-197, Gly-200 and Leu-251. Of these, Matthews et al. [31] site-specifically mutated three and found reduced enzymatic activity for Arg-65 and Tyr-168 (74% and 49% of wild-type, respectively), but not for Ser-197. In addition, mutagenesis of the highly conserved Thr-228 reduced the enzyme's activity to 32% of wild type; although Thr-228 was not assigned by NMR, the adjacent subsurface residue, Ala-229, was strongly perturbed by both ligands supporting the likely role of Thr-228 as an active site residue. Our mutagenesis studies showed a severe reduction of the R65A mutation (<1% residual activity), contrasting to Mathews [31] result. Other reduced activity mutants included K7A (30%), D8N and D8H (14% ad 16%, respectively) and Y19C (1%). Interestingly, the K7A mutant had a *k*m 10-fold higher than the wild type enzyme, suggesting this residue is important for substrate binding. Mutagenesis of D8E only reduced activity to 60%, indicating the importance of the acidic side chain, and Y19F had an activity of 80%, indicating the importance of the aromatic side chain at this position. This is consistent with the fact that, as will be discussed later, several organisms have a phenylalanine or a tyrosine in this region, either at position corresponding to human residue 19 or 241. Additionally, we found reduced activity for E123A (24%), but not for R148A or R148Q which had over 100% activity (although R148A had a slightly increased *k*m). Since the HMB substrate of URO-synthase non-enzymatically auto-catalyzes its own ring closure without D-ring isomerization resulting in URO'gen I, it is possible that the primary function of the enzyme is to constrain the catalysis to favor formation of the III-isomer. Thus, one may not expect to identify any single residue that is absolutely essential for catalytic activity, as previously suggested [31]. Mutants that showed less than 1% residual activity, such as mutants of Tyr-19 and Arg-65, can be explained as important for substrate binding, orientation, and/or shaping of the active site. Leu-36 was a strongly perturbed residue and mutagenesis to the shorter alanine significantly reduced enzymatic activity (16%). A possible model of ligand binding is proposed in Figure 3.26, which shows the predicted docking of the lowest energy conformer of URO'gen III to the active site composed of the conserved, perturbed, surface residues. The teal-colored perturbed residues were >90% conserved in eukaryotes, while the sand-colored perturbed residues Asp-8 and Arg-148 were less well conserved conserved (about 50%). It should be noted however that for large ligands, some active site residues may be less- or non-conserved, yet involved in catalysis or binding [79,80]. Furthermore, sequence conservation is not a good reflection of structure or function conservation, as the same functional group can be shifted throughout evolution to an adjacent location in the tri-dimensional space, but far in the primary sequence. While most of the domain 1 contact residues predicted by docking were also perturbed by ligand binding, there were additional perturbed cleft-surface residues in domain 2 (e.g. Gly-100, Thr-103, and Glu-123), indicating that the tetrapyrroles also bind to the upper region of the cleft. This observation is supported by the molecular dynamics simulation of the URO-synthase/URO'gen III complex using the NMR solution structure. In this case, the ligand makes close contact with Glu-100, and in the region of Thr-103. These results are consistent with the previous suggestion, based on the two conformations of URO-synthase in the crystal unit cell and the NMR solution structure, that the protein is flexible and able to adjust its inter-domain distance [31] and the normal modes analyses (NMR) which showed large hinged movements of the domains. Therefore, enzyme conformation in the NMR structure allows both domains to efficiently bind the substrate, explaining the NMR perturbation of residues seen in both domains 1 and 2.

In sum, these NMR studies established the nearly complete backbone and 85% side-chain resonance assignments for human URO-synthase and experimentally localized its active site to the cleft between domains 1 and 2, identifying putative contact residues involved in the binding of its substrate and product. *In silico* docking and molecular dynamics simulation provided a model of the enzyme-URO'gen III complex, consistent with the NMR experimental data. In addition, these NMR studies of the interaction of URO-synthase with HMB-synthase or URO-decarboxylase did not support the long-held hypothesis that the cytosolic heme biosynthetic enzymes formed a protein-protein complex for efficient metabolism of ALA to coproporphyrinogen III.

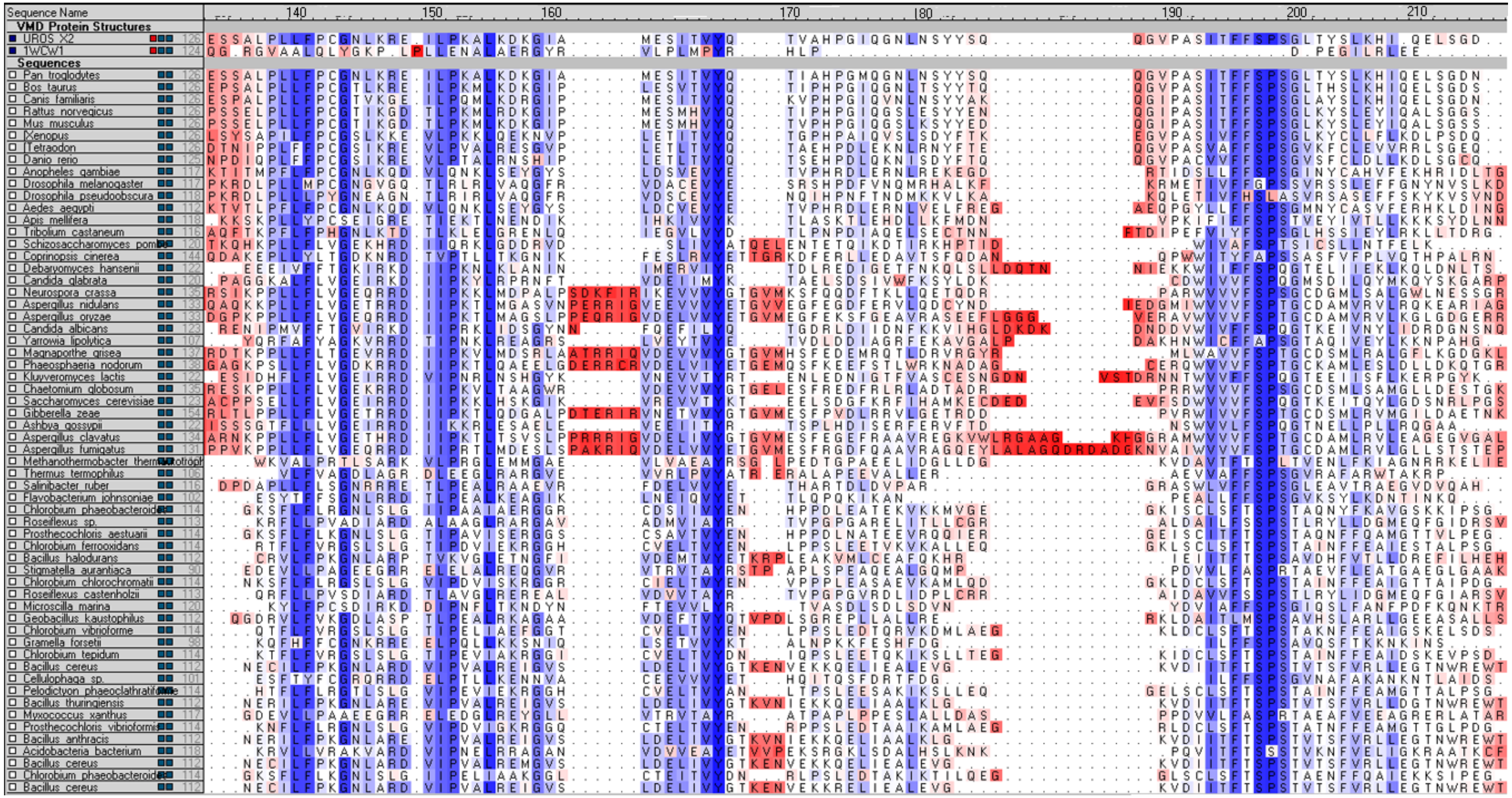
Appendix 3.1 (next pages) - Structure base sequence alignment of URO-porphyrinogen III sequences. The human (1JR2, name UROS_x2) and Thermophilus (1WCW, named 1WCW1) structures were aligned and the sequence alignment of all species were optimized to align with the know crystal structures, using the module multiseq of the molecular visualization program VMD. The red to blue gradient represents the conservation at any given position, being blue the most conserved, and red, the least. The SwissProt code of each sequence is indicated. The sequences used have the following SwissProt accession codes: Vertebrata - Human: P10746; Pan tryglodites: XP_508102.2; Bos taurus: XP_877187, Canis familiaris: XP_854146; Rattus novergius: NP_001012068; Mus musculus: NP_033505; Xenopus laevis Q66KP4; Tetrahedron nigroviridis: Q4SEB3; Danio rerio: BAE95178.1. Arthopoda - Anopheles gambiae: Q7Q3V6; Drosophila melanogaster: Q9VF38M; Aedes aegypti: Q17JM0; Apis mellifera: XP_392378.1; Tribolium castaneum: XP_967605.1; Drosophila pseudoobscura: EAL32042.1. Fungi - Schizosaccharomyces pombe: NP_594008.1; Coprinopsis cinerea: EAU87695.1; Debaryomyces hansenii: XP_459567.1; Candida glabrata: XP_449246.1; Neurospora crassa: XP_959715.1; Aspergillus nidulans: XP_661322.1; Aspergillus oryzae: BAE57583.1; Candida albicans: XP_714390.1; Yarrowia lipolytica: XP_505208.1; Magnaporthe grisea: XP_364722.1; Phaeosphaeria nodorum: EAT89826.1; Kluyveromyces lactis: XP_452517.1; Chaetomium globosum: XP_001219659.1; Saccharomyces cerevisiae: NP_014921.1; Gibberella zeae: XP_386563.1; Ashbya gossypii: NP_985207.1; Aspergillus clavatus: XP_001268334.1; Aspergillus fumigatus: XP_751158.1. Bacteria: Cryptococcus neoformans: XP_567721.1; Salinibacter rubber: YP_445674.1; Flavobacterium johnsoniae: ZP_01246588.1; Chlorobium phaeobacteroides ZP_00531421.1; Roseiflexus sp: ZP_01357250.1; Prosthecochloris aestuarii ZP_00592993.1; Chlorobium ferrooxidans: ZP_01386963.1; Bacillus halodurans: NP_243911.1; Stigmatella aurantiaca: ZP_01465383.1; Chlorobium chlorochromatii: YP_379812.1; Roseiflexus castenholzii: ZP_01533069.1; Microscilla marina: ZP_01692844.1; Geobacillus kaustophilus: YP_148497.1; Chlorobium vibrioforme: Q59335; Lawsonia intracellularis: YP_594495.1; Gramella forsetii: YP_863235.1; Bacillus cereus: YP_085786.1; Cellulophaga sp. : ZP_01049603.1; Pelodictyon phaeoclathratiforme: ZP_00589304.1|; Bacillus thuringiensis: YP_896780.1; Myxococcus xanthus: YP_630876.1; Prosthecochloris vibrioformis: ZP_00661215.1; Chlorobium tepidum: NP_662314.1; Bacillus anthracis: NP_846908.1; Acidobacteria bacterium: YP_591637.1; Bacillus cereus: NP_980847.1; Chlorobium phaeobacteroides: YP_912238.1; Bacillus cereus: ZP_00237475.1



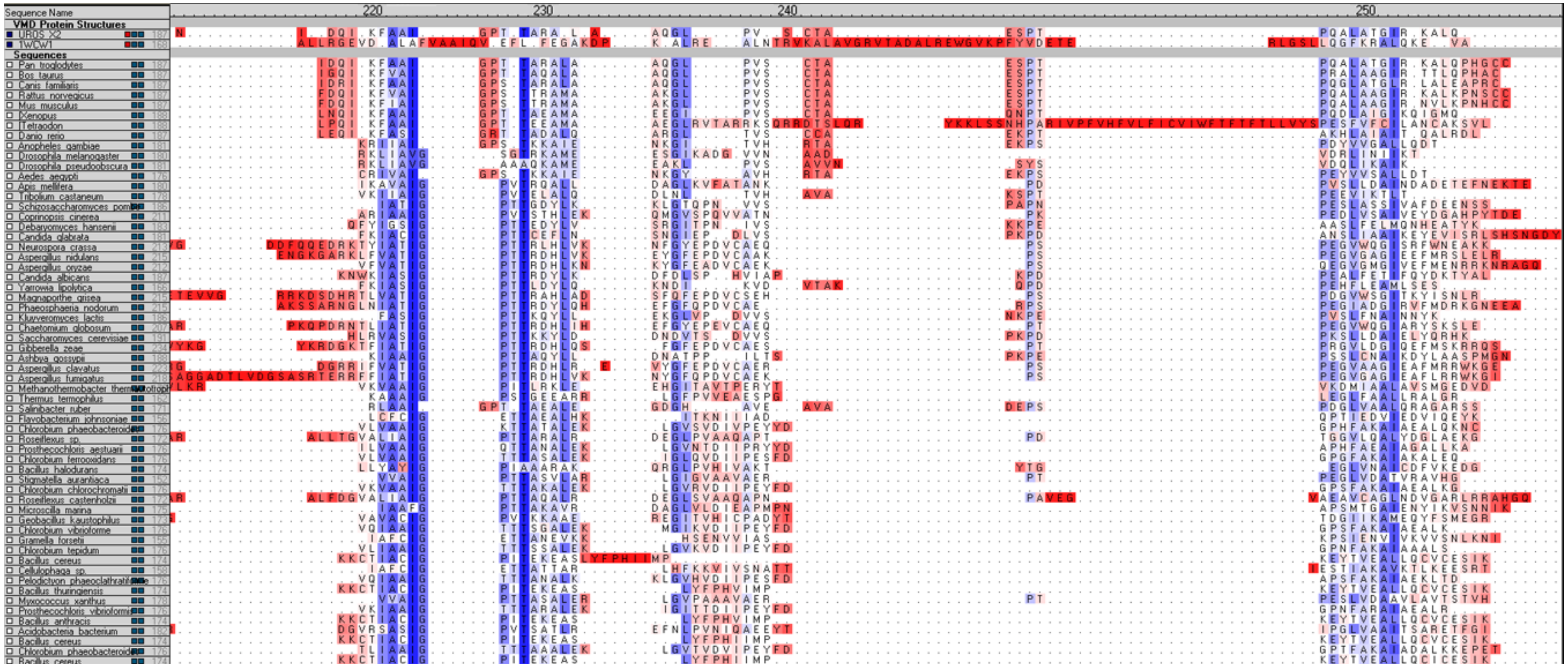
Appendix 3.1, cont. (legend on page 133)



Appendix 3.1, cont. (legend on page 133)



Appendix 3.1, cont. (legend on page 133)



Appendix 3.1, cont. (legend on page 133)

BIBLIOGRAPHY

1. Shmelev, V.K. and T.P. Serebrenikova, 1997, A study of supramolecular organization of glycogenolytic enzymes in vertebrate muscle tissue. Biochem Mol Biol Int, 43(4): p. 867-72.
2. Debnam P.M., Shearer G., Blackwood L., Kohl DH., 1997, Evidence for channeling of intermediates in the oxidative pentose phosphate pathway by soybean and pea nodule extracts, yeast extracts, and purified yeast enzymes. Eur J Biochem, 246(2): p. 283-90.
3. Srere, P.A., 1985, The metabolon. Trends in Biochemical Sciences, 10(3): p. 109-110.
4. Ovadi, J., 1991, Physiological significance of metabolic channelling. J Theor Biol, 152(1): p. 1-22.
5. Seewald M.J., Kraemer A., Farkasovsky M., Korner C., Wittinghofer A., Vetter I.R., 2003, Biochemical characterization of the Ran-RanBP1-RanGAP system: are RanBP proteins and the acidic tail of RanGAP required for the Ran-RanGAP GTPase reaction? Mol Cell Biol, 23(22): p. 8124-36.
6. Houben, K.F. and M.F. Dunn, 1990, Allosteric effects acting over a distance of 20-25 A in the *Escherichia coli* tryptophan synthase bienzyme complex increase ligand affinity and cause redistribution of covalent intermediates. Biochemistry, 29(9): p. 2421-9.
7. Gaertner, F.H., M.C. Ericson, and J.A. DeMoss, 1970, Catalytic facilitation in vitro by two multienyzme complexes from *Neurospora crassa*. J Biol Chem, 245(3): p. 595-600.
8. Srere, P.A. and K. Mosbach, 1974, Metabolic compartmentation: symbiotic, organellar, multienzymic, and microenvironmental. Annu Rev Microbiol, 28(0): p. 61-83.
9. Williams, L.G., S.A. Bernhardt, and R.H. Davis, 1971, Evidence for two discrete carbamyl phosphate pools in *Neurospora*. J Biol Chem, 246(4): p. 973-8.
10. Matchett, W.H., 1974, ndole channeling by tryptophan synthase of *neurospora*. J Biol Chem, 249(13): p. 4041-9.
11. Mowbray, J. and V. Moses, 1976, The tentative identification in *Escherichia coli* of a multienzyme complex with glycolytic activity. Eur J Biochem, 66(1): p. 25-36.
12. Clegg, J.S., 1984, Properties and metabolism of the aqueous cytoplasm and its boundaries. Am J Physiol, 246(2 Pt 2): p. R133-51.
13. Miles, E.W., S. Rhee, and D.R. Davies, 1999, The molecular basis of substrate channeling. J Biol Chem, 274(18): p. 12193-6.

14. Srere, P.A., 1987, Complexes of sequential metabolic enzymes. Annu Rev Biochem, 56: p. 89-124.
15. Vitello, L.B. and J.E. Erman, 1987, Binding of horse heart cytochrome c to yeast porphyrin cytochrome c peroxidase: a fluorescence quenching study on the ionic strength dependence of the interaction. Arch Biochem Biophys, 258(2): p. 621-9.
16. Holmgren, A., 1989, Thioredoxin and glutaredoxin systems. J Biol Chem,. 264(24): p. 13963-6.
17. Roux, B. and C.T. Walsh, 1993, p-Aminobenzoate synthesis in *Escherichia coli*: mutational analysis of three conserved amino acid residues of the amidotransferase PabA. Biochemistry, 32(14): p. 3763-8.
18. Zuiderweg, E.R., 2002, Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry, 41(1): p. 1-7.
19. Garrett D.S., Seok Y.J., Peterkofsky A., Gronenborn A.M., Clore G.M., 1999, Solution structure of the 40,000 Mr phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. Nat Struct Biol, 6(2): p. 166-73.
20. Hall D.A., Vander Kooi C.W., Stasik C.N., Stevens S.Y., Zuiderweg E.R., Matthews RG., 2001, Mapping the interactions between flavodoxin and its physiological partners flavodoxin reductase and cobalamin-dependent methionine synthase. Proc Natl Acad Sci U S A,. 98(17): p. 9521-6.
21. Warren, M.J. and A.I. Scott, 1990, Tetrapyrrole assembly and modification into the ligands of biologically functional cofactors. Trends Biochem Sci, 15(12): p. 486-91.
22. Romeo, G., M.M. Kaback, and E.Y. Levin, 1970, Uroporphyrinogen 3 cosynthetase activity in fibroblasts from patients with congenital erythropoietic porphyria. Biochem Genet, 4(6): p. 659-64.
23. Deybach J.C., de Verneuil H., Phung N., Nordmann Y., Puissant A, Boffety B., Congenital erythropoietic porphyria (Gunther's disease): enzymatic studies on two cases of late onset. J Lab Clin Med, 1981. 97(4): p. 551-8.
24. Bogorad, L., 1958, The enzymatic synthesis of porphyrins from porphobilinogen. I. Uroporphyrin I. J Biol Chem, 233(2): p. 501-9.
25. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1987, Purification and properties of uroporphyrinogen III synthase from human erythrocytes. J Biol Chem, 262(3): p. 1268-73.
26. Higuchi, M. and L. Bogorad, 1975, The purification and properties of uroporphyrinogen I synthases and uroporphyrinogen III cosynthase. Interactions between the enzymes. Ann N Y Acad Sci, 244: p. 401-18.

27. Frydman, R.B. and Feinstein, G. 1974, Studies on porphobilinogen deaminase and uroporphyrinogen 3 cosynthase from human erythrocytes. Biochim Biophys Acta, 350(2): p. 358-73.
28. Sancovich, H. A., Battle, A. M. & Grinstein, M., 1969, Porphyrin biosynthesis. VI. Separation and purification of porphobilinogen deaminase and uroporphyrinogen isomerase from cow liver. Porphobilinogenase an allosteric enzyme. Biochim Biophys Acta 191, 130-43.
29. Battersby, A. R., Fookes, C. J. R., Matcham, G. W., McDonald, E. & Gustafson-Potter, K. E. (1979). Biosynthesis of the natural porphyrins: Experiments on the ring-losure steps and with the hydroxy-analogue of porphobilinogen. J. Chem. Soc., Chem. Comm., 316-319.
30. Tsai, S.F., D.F. Bishop, and R.J. Desnick, 1987, Coupled-enzyme and direct assays for uroporphyrinogen III synthase activity in human erythrocytes and cultured lymphoblasts. Enzymatic diagnosis of heterozygotes and homozygotes with congenital erythropoietic porphyria. Anal Biochem, 166(1): p. 120-33.
31. Mathews M.A., Schubert H.L., Whitby F.G., Alexander K.J., Schadick K., Bergonia H.A., Phillips J.D., Hill C.P., 2001, Crystal structure of human uroporphyrinogen III synthase. Embo J, 20(21): p. 5832-9.
32. Pichon, C., Atshaves, B.P., Stolowich, N.J., and Scott, A.I.,1994, Evidence for an intermediate in the enzymatic formation of uroporphyrinogen III. Bioorg. Med. Chem. 2, 153–168
33. W.M. Stark, Hawker C. J., Hart G.J, Philipides A, Petersen P. M., Lewis J. D.,Leeper F. J., Battersby A. R., J., 1993, Biosynthesis of Porphyrins and Related macrocycles. Part 40.I." Synthesis of a Spiro-lactam Related to the Proposed Spiro-intermediate for Porphyrin Biosynthesis: Inhibition of Cosynthetase, Chem Soc, Perkin Trans. 1, 2875
34. Craig J. Hawker, Alan C. Spivey, Finian J. Leeper and Alan R. Battersby, *1998,* Biosynthesis of porphyrins and related macrocycles. Part 48.1,2 The rearrangement of 2*H*-pyrroles (pyrrolenines) related to the proposed spiro-intermediate for porphyrin biosynthesis, *J. Chem. Soc., Perkin Trans. 1, 1509-1517*
35. Tietze, L.F. and Geissler, H. 1993, Investigations in the biosynthesis of the pigments of life: calculations on the formation of uroporphyrinogen III from hydroxymethylbilane and description of a new mechanism for the D-ring inversion. Angew.Chem. Int. Ed. Engl. 32, 1040–1042
36. Battersby A. R., Hodgson G. L., Hunt, E., McDonald E., Saunders, 1982, Biosynthesis of porphyrins and related macrocycles. Part VI. Nature of the

rearrangement process leading to the natural type III porphyrins, J. Chem Soc,Perkin Trans 1, 2027.

37a. Scott, I., 2003, Discovering Nature’s Diverse Pathways to Vitamin B12: A 35-Year odissey. JOC, v. 68, 7, 2529-2539

37b. Scott A.I., 1990, Mechanistic and evolutionary aspects of vitamin B12 bioynthesis, Pure and applied chemistry, 62:1269

38. Matheson J. H., Corwin A. H., 1961, Biosynthesis of Pyrrole Pigments: A Mechanism for Porphobilinogen Polymerization, J. Am. Chem. Soc, 83, 135.

39a. Medek, A., Olejniczak, E. T., Meadows, R. P., and Fesik, S. W., 2000, An approach for high-throughput structure determination of proteins by NMR spectroscopy. J Biomol NMR, 18: p. 229-238.

39b Tugarinov, V., and Kay, L. E., 2003, An isotope labeling strategy for methyl TROSY spectroscopy. J Am Chem Soc 125, 13868-13878.

40. Neri D., Szyperski T., Otting G., Senn H., Wuthrich K., 1989, Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional 13C labeling. Biochemistry, 28(19): p. 7510-6.

41. Pichon C., Atshaves B., Xue T., Stolowich N.J., Scott A.I., 1994, Studies of URO'GEN III synthase with Modified Bilanes. Bioorg Med Chem letters, 4(9): p. 1105-1110.

42. Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J., Bax A., 1995, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR, 6(3): p. 277-93.

43. Johnson, B.A., Blevins, R. A., 1994, NMRView: a computer program for the visualization and analysis of NMR data. Journal of Biolmolecular NMR, 4: p. 603-614.

44. Yamazaki, T., Lee, W., Arrowsmith, C. H., Mahandiram, D. R., and Kay, L. E., 1994, A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. J. Am. Chem. Soc., 116(26): p. 11655-11666.

45. Logan T.M., Olejniczak E.T., Xu R.X., Fesik S.W., 1993, A general method for assigning NMR spectra of denatured proteins using a 3D HC(CO)NH-TOCSY triple resonance experiments. J. Biolmol. NMR, 3: p. 225-231.

46. Goto, N.K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E., 1999, A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. J Biomol NMR,. 13: p. 369-374.

47. Tugarinov V., K.L., 2003, Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. J Am Chem Soc, 125(45): p. 13868-78.

48. Wishart, D.S. and B.D. Sykes, 1994, The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. J Biomol NMR, 4(2): p. 171-80.

49 Berjanskii M.V., Neal S., Wishart D.S., 2006, PREDITOR: a web server for predicting protein torsion angle restraints. Nucleic Acids Res. 34(Web Server issue):W63-9

50. Clore, G. M. & Gronenborn, A. M.,1994, Multidimensional heteronuclear nuclear magnetic resonance of proteins, Methods Enzymol 239, 349-363

51. Brünger, A. T.,1993, X-PLOR, version 3.1. A system for X-ray crystallography and NMR (Yale University Press, New Haven).

52 Laskowski R.A., MacArthur M.W., Moss D.S. & Thornton J.M., 1993, PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst., 26, 283-291.

53 Cavanagh, J., Fairbrother, W.J., Palmer, A.G.I., Skelton, N.J.,Protein NMR spectroscopy. Principles and practice. NewYork: Academic Press; 1996, p: 587.

54. Morris, G., Goodsell, D., Haliday, R., Huey, R., Hart, W., Belew, R., and Olson, A., 1998, Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy, J. Computational Chemistry 19, 1639-1662.

55. MacKerell, A. D., Jr., Bashford, D., Bellot, M., Dunbrack, R. L., J.D., E., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E. I., Roux, B., Schlenkrich, B., Smith, J., Stote, R., Straub, S., Watanabe, M., Wiorkiewicz-Kuczera, J., and Karplus, M. ,1998, All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins , J. Phys. Chem. B. 102, 3586-3616.

56. Dewar, J. S., Zoebisch, E. G., Healy, E. F., and Stewart, J. J. P., 1985, AM1: a new general purpose quantum mechanical molecular model J. Am. Chem. Soc. 107, 3902-3909.

57. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakarzewsky, V. G., Montgomery, J. A., Stratman, R. E., Burant, J. C., Dapprich, S., Millam, J. C., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Menucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ciolowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I.,

Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzales, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A.,1998 Gaussian-98, Revision A.1. In., Gaussian, Inc., Pittsburg, PA

58. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot,C., Skeel, R.D., Kale, L., Schulten, K., 2005, Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 26:1781-1802.

59 Mahoney, M. W., and Jorgensen, W. L., 2000, A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions, *J. Chem. Phys.* **112**,8910-8922.

60. Humphrey, W., Dalke, A. and Schulten, K., 1996, "VMD - Visual Molecular Dynamics", J. Molec. Graphics, vol. 14, pp. 33-38.

61. Ikura, M., L.E. Kay, and A. Bax, A, 1990, novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry, 29(19): p. 4659-67.

62. Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. , 1991, A relational database for sequence-specific protein NMR data. J Biomol NMR 1, 217-236.

63. DeLano, W. L., 2002, The PyMOL Molecular Graphics System. In., DeLano Scientific, San Carlos, CA, USA.

64. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. ,2001, Electrostatics of nanosystems: Application to microtubules and the ribosome Proc Natl Acad Sci U S A 98, 10037-10041.

65. Hinsen K., 1998, Analysis of domain motions by approximate normal mode calculations. Proteins 33:417-429.

66. Anderson, P.M. and R.J. Desnick, 1980, Purification and properties of uroporphyrinogen I synthase from human erythrocytes. Identification of stable enzyme-substrate intermediates. J Biol Chem, 255(5): p. 1993-9.

67. Jordan, P.M. and Warren, M.J., 1987, Evidence for a dipyrromethane cofactor at the catalytic site of *E. coli* porphobilinogen deaminase. FEBS Lett., 225(1-2): p. 87-92.

68. Louie, G.V., et al., 1992, Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. Nature, 359(6390): p. 33-9.

69. Awan, S.J., et al., 1997, Reconstitution of the holoenzyme form of *Escherichia coli* porphobilinogen deaminase from apoenzyme with porphobilinogen and preuroporphyrinogen: a study using circular dichroism spectroscopy. Biochemistry. 36(30): p. 9273-82.

70. Hart, G.J., Miller A.D., Battersby A.R., 1988, Evidence that the pyrromethane cofactor of hydroxymethylbilane synthase (porphobilinogen deaminase) is bound through the sulphur atom of a cysteine residue. Biochem J, 252(3): p. 909-12.
71. Radmer, R.,Bogorad L., 1972, A tetrapyrrylmethane intermediate in the enzymatic synthesis of uroporphyrinogen. Biochemistry, 11(5): p. 904-10.
72. Jordan, P.M. and Woodcock S.C., 1991, Mutagenesis of arginine residues in the catalytic cleft of *Escherichia coli* porphobilinogen deaminase that affects dipyrromethane cofactor assembly and tetrapyrrole chain initiation and elongation. Biochem J,. 280 (Pt 2): p. 445-9.
73. Warren, M.J., Jordan P. M., 1988, Investigation into the nature of substrate binding to the dipyrromethane cofactor of *Escherichia coli* porphobilinogen deaminase. Biochemistry, 27: p. 9020-9030.
74. Shoolingin-Jordan P.M., Al-Dbass A., McNeill L.A., Sarwar M., Butler D., 2003, Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. Biochem Soc Trans, 31(Pt 3): p. 731-5.
75. Hart, G.J. and A.R. Battersby, 1985, Purification and properties of uroporphyrinogen III synthase (co-synthetase) from Euglena gracilis. Biochem J, 232(1): p. 151-60.
76. Louie G.V., Brownlie P.D., Lambert R, Cooper JB, Blundell T.L., Wood S.P., Malashkevich VN, Hadener A, Warren MJ, Shoolingin-Jordan PM, 1996, The three-dimensional structure of *Escherichia coli* porphobilinogen deaminase at 1.76-A resolution. Proteins, 25(1): p. 48-78.
77. Helliwell, J. R., Nieh, Y. P., Habash, J., Fauldner, P. F., Raftery, J., Cianci, M., Wulff, M., and Hädner, A., 2003, Time-resolved and static-ensemble structural chemistry of hydroxymethylbilane synthase. Faraday Discuss., 122, 131-144
78 Bishop, D. F., and Desnick, R. J., 1982, Assays of the heme biosynthetic enzymes. Enzyme. 28, 91-93
79. El-Kabbani O., Old S.E., Ginell S.L., Carper D.A.,1999, Aldose and aldehyde reductases: structure-function studies on the coenzyme and inhibitor-binding sites. Mol Vis. Sep 3;5:20
80. Swarbrick J. D., Buyya S., Gunawardana D. , Gayler K.R. , McLennan A.G., and Gooley P.R., 2005, Structure and Substrate-binding Mechanism of Human Ap4A Hydrolase. J. boil chem, 280, 9, 8471–8481,

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

(1) Identification of the UROD gene mutation responsible for PCT in two presumably unrelated Portuguese families.

In this work, the entire UROD sequence of each proband was amplified by long-range PCR, and each exon and the adjacent intronic or flanking boundaries were amplified and sequenced. A G to A transition at cDNA nucleotide 960 was identified in both probands, in a consensus splice junction. This mutation was confirmed by RFLP analysis in several family members. RT-PCR showed that this transition was a splicing mutation, that resulted in the skipping of exon 9 and an aberrant form of the protein, which is truncated at residue 298 (291 normal residues, plus 7 abnormal residues from exon 10). This lesion presumably causes the mRNA to be degraded, so little mutant enzyme is synthesized.

These results permit pre-symptomatic molecular diagnosis and counseling of these families to enable family members to avoid disease precipitating factors. PCT typically has a very low penetrance of ~10%, as the half-normal enzyme activity is sufficient for the efficient decarboxylation of URO III, requiring further inactivation by factors such as alcohol abuse, iron overload and estrogen use to precipitate the diseases symptoms. It is intriguing that several family members show clinical symptoms in these families, and a common mechanism, still to be found, must be responsible for further contributing for the precipitation of clinical symptoms. Further studies should identify the mechanism involved in the precipitation of PCT in these families.

(2) Expression, purification and characterization of HMBS, UROS and UROD

A fast and efficient method has been developed for the expression and purification of large quantities of three of the four cytosolic heme biosynthetic enzymes. Each human recombinant enzyme was expressed in a prokaryotic expression system, which was selected and optimized individually. Purification involved three chromatographic steps: metal affinity, ion exchange, and gel filtration chromatography. Previously, small quantities of human HMBS and UROS were purified from outdated human erythrocytes [1-4] and more recently, purification of milligram quantities of URO-synthase has been reported using both a His.tag [5], and non-tag approaches [6]. The first produced an enzyme with a 21-residue N-terminal extension, and the second had a very low yield. Our method is the first to provide almost native sequence (only one terminal serine remains after the cleavage of the terminal SUMO peptide), with high yield. It was very important for us to have a high-level expression with high yield for the subsequent NMR studies to be feasible, which use very expensive labeling protocols. Furthermore, conditions were established to maintain URO-synthase stable for long periods of time, as required in the NMR experiments.

(3) Molecular interaction of UROS and HMBS

For over four decades, investigators have suggested that HMBS and UROS interact in a cytosolic complex to efficiently convert PBG to URO III, avoiding the formation of any URO I isomer, a non-physiological and toxic compound. Previous efforts to demonstrate an HMBS-UROS intermolecular association by yeast two-hybrid studies and by immunoprecipitation of *in situ* cross-linked complexes were unsuccessful [7], indicating that if an interaction exists, it must be very weak, above the detection threshold of such traditional techniques. Using NMR spectroscopy, we demonstrated that UROS does not interact directly with HMBS and/or UROD, at concentrations in the milimolar range, two orders of magnitude higher than those found in the liver. Therefore it is unlikely that these enzymes interact *in vivo*.

(4) NMR Resonance Assignments, Solution Structure Determination, and Active site mapping of URO-synthase. Future directions to investigate the role of the enzyme's dynamics and flexibility in the reaction mechanism.

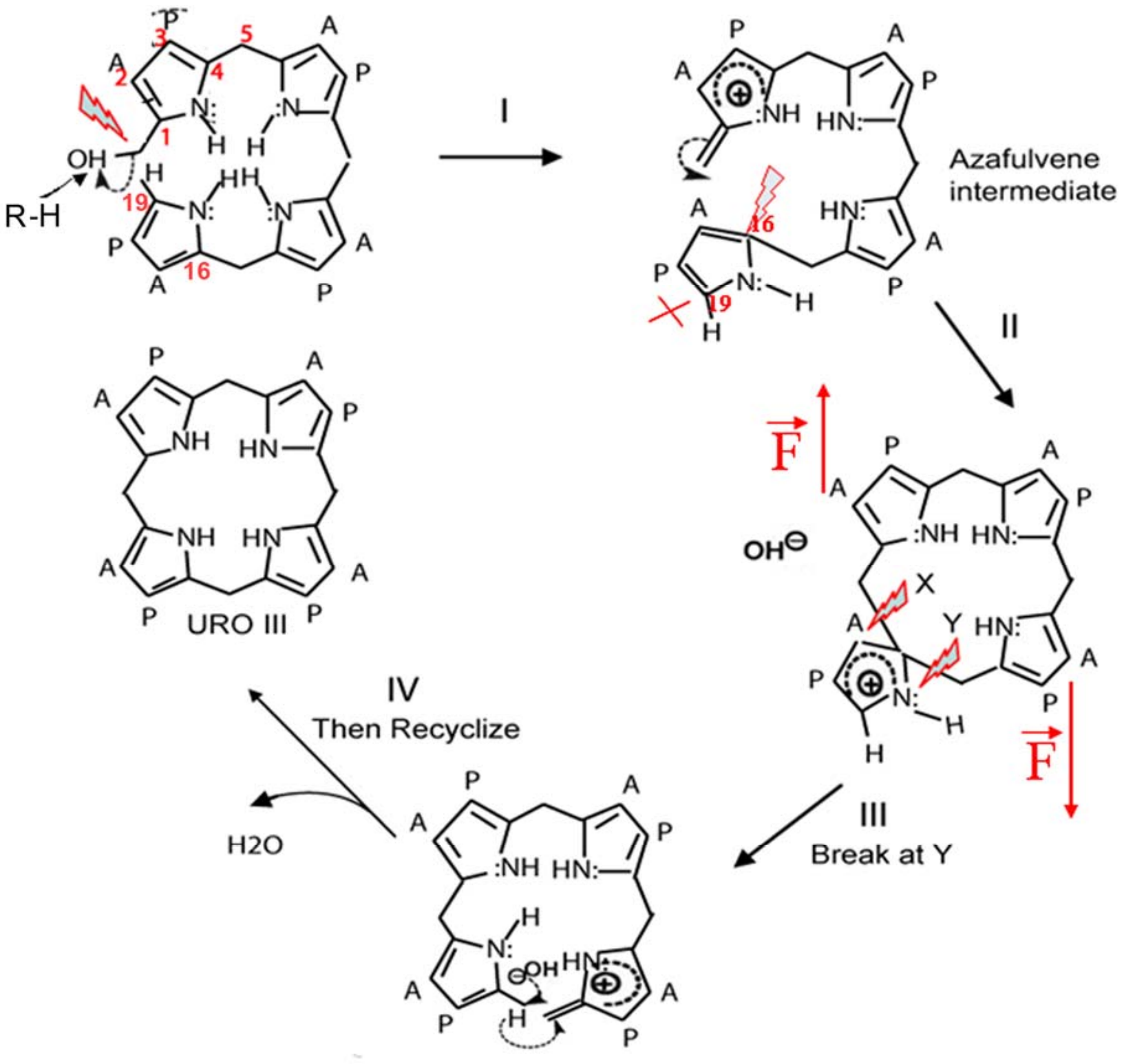
In order to use NMR techniques to study the interaction of UROS with the other cytosolic heme biosynthetic enzymes, as well as to map its active site, we determined the backbone resonance assignments. Using the chemical shift perturbation method, in which we titrated URO-synthase with two competitive inhibitors of substrate binding, we identified the cleft between the two globular domains as the binding site of these small molecule inhibitors. By identifying which resonances in the [^{15}N, ^{1}H]-HSQC spectrum are most perturbed by the presence of these inhibitors, we were able to point specific residues, which were likely binding candidates on the protein interface. Interestingly, the NMR results were congruent with the predicted results obtained by *in silico* experiments (which included both protein-ligand docking and molecular dynamics simulations). Unfortunately, carbon NOE's between the ligands and the enzyme could not be obtained, which would have provided a direct measure of the enzyme-small ligand interaction. Within the predicted active-site cleft, only one methyl side chain is within range of the enzyme-ligand model, and future experimentation might explain why a NOE was not observed for this residue. Furthermore, mutagenesis of the highly perturbed residues Asp-8 and Lys-7, as well as Tyr-19 and Arg-65 showed reduced enzymatic activity. It is significant that those perturbed residues had the lowest residual activity. Mutagenesis of Arg-65 inactivates the enzyme, contrary to the findings of Mathews *et al* [5], which retained 74% activity. NMR analysis of the mutant enzyme will be performed to confirm that our mutant enzyme retains a native structure. Further experimentation is necessary, including molecular dynamics simulations, to investigate the role of this residue in substrate binding and orientation. In the previously determined crystal structure, the cleft between the globular domains was too large to bind the ligand tightly. Contacts between the inhibitors and residues in both domains, which were predicted by experimental NMR titrations, were not possible in this “open” structure. Therefore we hypothesized that the solution structure must have a smaller cleft than that observed in the crystal structure. Supporting evidence for this came from the fact that there is a difference in size of the cleft in both structures present in the crystal cell, as well as normal mode analysis, which showed large domain movements. We further determined the resonance of methyl side chains and intermolecular NOEs, and solved the structure of the enzyme in solution. We found the solution structure to be very similar to the crystal structure, but with a smaller cleft size.

Having mapped the active site, we are now focusing our efforts in determining the role of the enzyme dynamics (flexibility) on the reaction mechanism. Several mechanistic hypothesis have been proposed for the formation of URO III by UROS [8-12]. However, these have only investigated the chemical process that HMB must undergo to form URO’gen III, without considering the role of the enzyme in catalyzing this reaction. Based on our NMR studies, we hypothesize that three enzyme features are of major importance for the enzymatic inversion of ring D, producing the URO III isomer, but not URO I: a) the positively charged cleft, contrasting with the overall negative surface charge, helps steer the substrate to the active site; b) the enzyme might accelerate the activation of the azafulvene electrophile with a catalytic residue or a conserved water molecule that protonates the substrate’s hydroxyl group; c) in the active site, binding of the substrate to specific residues orients the enzyme for stereo-specific catalysis, protecting carbon 19 from the azafulvene, and enzyme dynamics impose constraints on the substrate.

To understand the role of the enzyme dynamics in the reaction mechanism, it is important to understand the role of conserved structural and functional features. Despite the sequence similarity of the enzyme in different branches of evolution being only 33%, the available 3D structure of the *Thermus thermophilus* enzyme (PDB: 1WCW) shows a structure very similar to the human enzyme. It is apparent that the structure and functional groups are better conserved than the primary sequence and it is therefore important to understand conservation beyond the primary sequence. For instance, Tyr-19 is well conserved in eukaryotes, but not in Prokaryotes. However, bacterial species have a tyrosine or a phenylalanine (both aromatic residues) at either the position corresponding to position 19 in humans, or at position 241, located in an adjacent alpha-helix. The functional group (aromatic side chain) has been conserved in this region, as the side chain likely occupies the same space, being at the position 19 or 241. Mutagenesis of this tyrosine to a phenylalanine retains 80% residual enzymatic activity, while mutation to a cysteine has only 1%. It is therefore clear the importance of an aromatic residue in this region. Other important feature of the active site is its positive charge. The enzyme has several positively charged arginines and lysines in the cleft region including residues at the human positions 7, 65 and 148. A positive charged is almost fully conserved at position 7, but not at the other two positions. However, bacterial species that lack a positively charged residue in position corresponding to the human residue 65 have a positively charged aminoacid in the position corresponding to human residue 102, localized in an adjacent alpha-helix. Although mutagenesis of residue Arg-148 did not influence the enzymatic activity, the K7A and R65A mutant have reduced enzymatic activity. Preliminary molecular dynamics simulation data have implicated both Lys-7 and Arg-65 in substrate binding, but the role of Arg-148 in unclear. Preliminary kinetic data showed a ten fold increase in the *k*m of the K7A mutant, further suggesting this residue's importance in substrate binding. Arg-148 has a slightly increased *k*m (2-3 fold) and might be important in the initial steering the substract to the active site cleft. Understanding the conservation of structural features and of functional groups is crucial to predict the enzyme structure in different branches of evolution by homology modeling, which may help understand which features are conserved and contribute to substrate binding and catalysis, beyond simple primary sequence conservation. Several enzyme residues are absolutely conserved in the primary sequence and further experimentation/simulation is required to understand their role. Hydroxyl containing residues in the active site, such as Thr-62, Ser-63, Thr-103, Tyr 168, Ser-197, and Thr-228 may be involved in substrate binding. Other conserved residues might be important in maintaining the tertiary structure, such as conserved glycines and prolines (e.g. Gly 100, Pro-199). Therefore, understanding the conservation of functional groups, which might not be conserved in the primary sequence, will help us better understand the mode of binding and enzymatic mechanism.

The most widely accepted reaction mechanism for the conversion of HMB to URO III involves a spiropyrrolenine intermediate. Trapping experiments at low temperature and high pH have provided evidence for the formation of an azafulvene during catalysis [9]. Activation of the HMB requires the extraction of the hydroxyl group (I, Figure 4.1), resulting in the formation the azafulvene. This Azafulvene can be generated at acidic pH or enzymatically by UROS [12]. The hydroxyl is not a good leaving group and its protonation might be required to generate the azafulvene. Although this can certainly take place spontaneously, as HMB can cyclize non-enzymatically in solution, the enzyme might provide a critical residue to accelerate this protonation. The same process might be performed by a conserved water molecule, in which case it would be harder to demonstrate. Active site residues whose mutagenesis significantly reduced enzymatic activity included Asp-8 and Arg-65. Given their very low and high pKa, respectively, these residues don’t make good candidates as proton donors. However, it is possible that in the environment of the enzyme, Asp-8 could be protonated and involved in the initial extraction of the hydroxyl, as seen in other enzymes [13]. Interestingly, fungi sequences, which lack this aspartic acid, have a histidine in the position corresponding to human Phe-38. This residue is located in the same region and could act as a proton donor as well. An alternative mechanism, and more difficult to demonstrate, would involve an activated water molecule, or a proton relay system involving several water molecules/residues. It is interesting to note that several water molecules are trapped in both crystal cell structures (PDB: 1JR2), both in domain 1, under Asp-8’s side chain, and in domain 2, under Thr-62. Of note, Gly-200, a non-surface residue, highly conserved in eukaryotes, was significantly perturbed in our NMR studies. This is a buried residue that cannot make direct contact with the substrate. However, this glycine is located at the edge of a buried cavity (underneath Asp-8), which is occupied by a water molecule, seen both in the crystal structure and in our molecular dynamics simulations. Future studies are required to investigate the role of this conserved water molecule, specifically if it is involved in substrate protonation. We are currently mutagenesing this glycine into a valine, whose larger side chain will occupy the cavity and prevent the water molecule from entering this space. If the structure is not significantly altered (as evaluated by the NMR HSQC spectrum of the mutant) and the enzyme activity is significantly reduced, further experimentation will be dedicated to this issue.

Future experiments also include simulating the lowest energy binding conformation for the enzyme’s substrate (HMB), the transition state intermediate spiropyrrolenine, the product URO’gen III, and the activated substrate (an azafulvene) by molecular dynamics, using both “open” crystal structure, and the “closed” NMR structure. Furthermore, simulations will include permutations of the pyrrole side chains to study their role in substrate binding, as it as been shown experimentally that permutating the carboxylates of rings A or B prevents binding, but not ring C and D [9]. Future research might include the intermediate conformations predicted by normal mode analyses, or the use of applied forces on the domains to accelerate their closure/opening, using steered molecular dynamics (SMD). Comparison of the binding free energies of all ligands and their preferred protein conformation will help us understand the mechanical principles for ring closure in the environment of the active site. Preliminary data have shown that hydrogen bonds between the ring D carboxylate side chain and the OH groups of Ser-62, Thr-63 and Tyr-168 of are essential for a stable conformation of the spiropyrrolenine in the active site. Interestingly, these residues are highly conserved throughout evolution. One hypothesis to be tested states that binding of the Ring D acetate to these hydroxyl groups will create steric hindrance, protecting carbon 19 from an attack by the azafulvene in ring A. Further research will complement the computational simulation results with site-directed mutagenesis, including multiple site-specific mutations, for instance to remove all potential contact points of ring D's acetate (Thr-62, Ser-63, and Thr-103 and Tyr-168) or by derivatization using chemical conjugates. In addition quantum mechanics/molecular mechanics hybrid simulations might provide a new tool to study the reaction mechanism.



$A = CH_2CO_2H$ $P = CH_2CH_2CO_2H$

Figure 4.1 - Formation of URO III from HMB. Important carbons involved in the mechanism are numbered in red. The first step involves the activation of HMB into an electrophile, extracting (by protonation) of the hydroxyl group (I), which results in the formation of an azafulvene. This azafulvene attacks preferentially carbon 16 of rind D (attack on carbon 19 leads to the formation of the URO I isomer), favored by the enzyme

stereo-specificity. In II, this intermediate spiropyrrolenine undergoes a selective bond fragmentation at position Y (III), again favored by enzyme stereo-specificity and bond rearrangement (IV), leading to the formation of the URO III isomer (based on the mechanism proposed by Hawker *et al* [12]). Opposite forces acting on rings C and D (shown in red arrows), cause ring D to flip to the side, straining the bond between rings D and C, and inverting ring D. Note that an alternative mechanism for the final rearrangement as a [1, 5]-sigmatropic shift, as seen in other isomerases [14-15].

BIBLIOGRAPHY:

1. Anderson, P.M., Desnick, R.J., 1980, Purification and properties of uroporphyrinogen I synthase from human erythrocytes. Identification of stable enzyme-substrate intermediates. J. Biol. Chem., 255(5): p. 1993-9.
2. Tsai, S.F., Bishop, D.F., Desnick, R.J., 1987, Purification and properties of uroporphyrinogen III synthase from human erythrocytes. J. Biol. Chem., 262(3): p. 1268-73.
3. Alwan, A.F., Mgbeje, B.I., Jordan, P.M., 1989, Purification and properties of uroporphyrinogen III synthase (co- synthase) from an overproducing recombinant strain of Escherichia coli K-12. Biochem J, 264(2): p. 397-402.
4. Stamford, N.P., A. Capretta, and A.R. Battersby, 1995, Expression, purification and characterisation of the product from the Bacillus subtilis hemD gene, uroporphyrinogen III synthase. Eur J Biochem, 231(1): p. 236-41.
5. Mathews M.A., Schubert H.L, Whitby F.G., Alexander K.J., Schadick K., Bergonia H.A., Phillips J.D., Hill C.P., 2001, Crystal structure of human uroporphyrinogen III synthase. Embo J., 20(21): p. 5832-9.
6. Omata Y., Sakamoto H., Higashimoto Y., Hayashi S. and Noguchi M., 2004, Purification and Characterization of Human Uroporphyrinogen III Synthase Expressed in Escherichia coli J. Biochem. 136, 211–220
7. Chen, C.-H., 1994, Molecular genetics of acute intermittent porphyria and studies of the human cytosolic heme biosynthetic enzymes, Thesis dissertation in Human Genetics. Mount Sinai School of Medicine of the City University of New York: New York. p. 152.
8. Leeper, A.R, 1990, Biosynthesis of the Pigments of Life: Mechanistic Studies on the Conversion of Porphobilinogen to Uroporphyrinogen III. Chem. Rev., 90: p. 1261-1274.
9. Pichon, C., Atshaves, B., Xue, T., Stolowich, N.J., Scott, A.I., 1994, Studies of URO'GEN III synthase with Modified Bilanes. Bioorg Med Chem letters, 4(9): p. 1105-1110.
10. Pichon C., A.B., Stolowich N.J., Scott A.I., 1994, Evidence for an intermediate in the enzymatic formation of uroporphyrinogen III. Bioorg Med Chem., 2(3): p. 153-68.
11. Pichon C, A.B., Danso-Danquah R., Stolowich N.J., Scott A.I., 1994, 19-Bromo-1-hydroxymethylbilane, a novel inhibitor of uro'gen III synthase. Bioorg Med Chem, 2(4): p. 267-77.

12. Hawker C., Leeper F. and Battersby A., 1998, Biosynthesis of porphyrins and related macrocycles. Part 48.1,2 The rearrangement of 2H-pyrroles (pyrrolenines) related to the proposed spiro-intermediate for porphyrin biosynthesis. J. Chem. Soc., Perkin Trans., 1: p. 1509 -1517.
13. Ostanin, K., Van Etten, R.L., 1993, Asp304 of Escherichia coli acid phosphatase is involved in leaving group protonation. J. Biol. Chem., 268(28) p:20778-84.
14. Hur, S., Bruice, T.C., 2003, Comparison of formation of reactive conformers (NACs) for the Claisen rearrangement of chorismate to prephenate in water and in the E. coli mutase: the efficiency of the enzyme catalysis. J Am Chem Soc, 125(19): p. 5964-72.
15. Hur, S., Newby, Z.E., Bruice, T.C., 2004, Transition state stabilization by general acid catalysis, water expulsion, and enzyme reorganization in Medicago savita chalcone isomerase. Proc Natl Acad Sci U S A, 101(9): p. 2730-5.