THE ASYMMETRIC STRUCTURE OF HUMAN UDP-α-D-XYLOSE SYNTHASE SUGGESTS A MECHANISM FOR REGULATING ACTIVITY

by

SAMUEL JUSTIN POLIZZI

(Under the Direction of Zachary Arthur Wood)

ABSTRACT

Proteoglycans provide structural integrity to connective tissues and facilitate growth factor binding to receptors. The first step in the formation of most proteoglycans is the covalent linkage of xylose to the hydroxyl of a serine on acceptor proteins. UDP-xylose (UDX) is the activated substrate required for xylose incorporation into the proteoglycan linker. UDX biosynthesis proceeds via the conversion of UDP-glucose to UDP-glucuronic acid (UGA) by UDP-glucose dehydrogenase (UGDH) and subsequent decarboxylation by UDP-xylose synthase (UXS). UDX is believed to regulate UDP-sugar pools by allosterically inhibiting UGDH and UXS. In order to understand how human UXS is regulated, we solved the 2.5A crystal structure of unliganded human UXS. UXS copurifies with the NAD^+ cofactor tightly bound, unlike the only other structurally characterized UGA decarboxylase, *E. coli* ArnA. While ArnA is reported to bind NAD^+ and UGA as substrates and release NADH and UDP-4-keto-xylose (UX4O) as products, UXS retains NADH and catalyzes a second hydride transfer to regenerate NAD^+ and release UDX. We show that UXS will release the reaction intermediates UX4O and NADH in the presence of exogenous NAD^+. The release of the reaction intermediates involves a cooperative conformational change that we show is conserved in ArnA. UXS activity is also

stimulated >6-fold in the presence of a molecular crowding agent trimethylamine *N*-oxide. Solution studies show that UXS undergoes a concentration-dependent dimer-to-tetramer association (K_d = 4.68 μM), with the tetramer being the most active species. The crystal structure of UXS reveals a possible allosteric site that we predict would disrupt the tetramer. Combining these data suggests a model of regulation between high activity tetramers and low activity dimers mediated by a UDP-containing molecule.

INDEX WORDS: UDP-xylose synthase, UDP-glucuronic acid decarboxylase, nucleotide sugar metabolism

THE ASYMMETRIC STRUCTURE OF HUMAN UDP-α-D-XYLOSE SYNTHASE SUGGESTS A MECHANISM FOR REGULATING ACTIVITY

by

SAMUEL JUSTIN POLIZZI

B.S., University of West Georgia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2012

THE ASYMMETRIC STRUCTURE OF HUMAN UDP-α-D-XYLOSE SYNTHASE SUGGESTS A MECHANISM FOR REGULATING ACTIVITY

by

SAMUEL JUSTIN POLIZZI

Major Professor: Zachary Wood

Committee: William Lanzilotta
Kelley Moreman
Lance Wells

Electronic Version Approved:
Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2012

DEDICATION

For my Grandfathers Leeming and Polizzi, my heroes of head and heart.

To the first giants whose shoulders I stood upon.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

UDP-xylose synthase (UXS) utilizes an NAD^+ cofactor to decarboxylate UDP-glucuronic acid (UGA) to UDP-xylose (UDX). Evolution of the UGA carboxylate as stable CO_2 is essentially an irreversible reaction that favors UDX production. UDX formation removes glucose-based molecules from energy metabolism and represents a terminal step in UDP-hexose interconversion. For vertebrates, a primary function of UDX appears to be in initiating proteoglycan synthesis. However, a newly synthesized proteoglycan with a 100 sugar residue heparan sulfate chain has only one sugar donated by UDX, compared to 49 residues derived from UGA. In order to maintain substrate pools for energy metabolism and proteoglycan synthesis, it has been proposed that UXS is regulated at the enzymatic level.[1; 2]

Early studies of UXS (a.k.a. UDP-glucuronic acid decarboxylase or UDP-glucuronate carboxy-lyase) showed that UDX activity is regulated by UDP-sugars. Native UXS from chicken chondrocytes displayed allosteric activation by UGA and inhibition by UDX.[2] Activation by UGA and UDP-glucose (UDG) in wheat germ led to the proposal that UXS contains non-catalytic regulatory sites.[1] Exogenous NAD^+ has also been shown to stimulate UXS activity[2; 3; 4; 5; 6] despite cofactor retention and regeneration during UDX biosynthesis. Prior to our work, neither the structural basis for allosteric regulation nor the mechanism for NAD^+ sensitivity had been investigated in UXS.

In order to understand the regulation of UXS activity, a review of UDX biosynthesis and the UXS family of short-chain dehydrogenase/reductase (SDR) enzymes will be presented. In the third thesis chapter, we examine the effect of exogenous NAD^+ on recombinant, *Homo sapiens* UXS (hUXS; residues 85-336) activity. Using the high-resolution technique Capillary Zone Electrophoresis (CZE), we show that free NAD^+ induces hUXS to release the reaction intermediates NADH and UDP-4-keto-xylose (UX4O) similar to the UGA decarboxylase domain of *Escherichia coli* ArnA ($ArnA_{dc}$). We detect comparable levels of homotropic cooperativity in hUXS and $ArnA_{dc}$, suggesting NADH release occurs via a conserved conformational change. Based on dynamic loop regions in the decarboxylase active site, we propose a structure-based model to explain hUXS cooperativity.

In the fourth chapter of this work, we address hUXS regulation by changes in the quaternary structure. Using sedimentation velocity analytical ultracentrifugation, we show hUXS is a dimer-tetramer self-association in solution. Under crowding conditions, hUXS displays a saturable increase in the rate of UDX formation consistent with a binding isotherm. In the presence of increasing hUXS concentrations, NADH production also increases along a binding curve that can be modeled with the same association constant as the dimer-tetramer self-association. The only stable tetramer found in available hUXS structures is formed by the packing of dimers into a fiber along the crystallographic 3_2 axis. SAXS data are consistent with an extended conformation of the tetramer in solution. Using our structural and kinetic data, we propose a model for hUXS regulation *in vivo*.

Finally, chapter five summarizes our findings in the context of previous UXS analyses and examines future areas of UXS research. Coauthored manuscripts pertinent to this work are included as appendices. Appendix A describes the only vertebrate model system for UXS

disruption and a single amino acid substitution that abolishes UDX formation.[7] Appendix B contains a biophysical study of how UDX regulates the upstream metabolic enzyme UDP-glucose dehydrogenase.[8]

CHAPTER 2

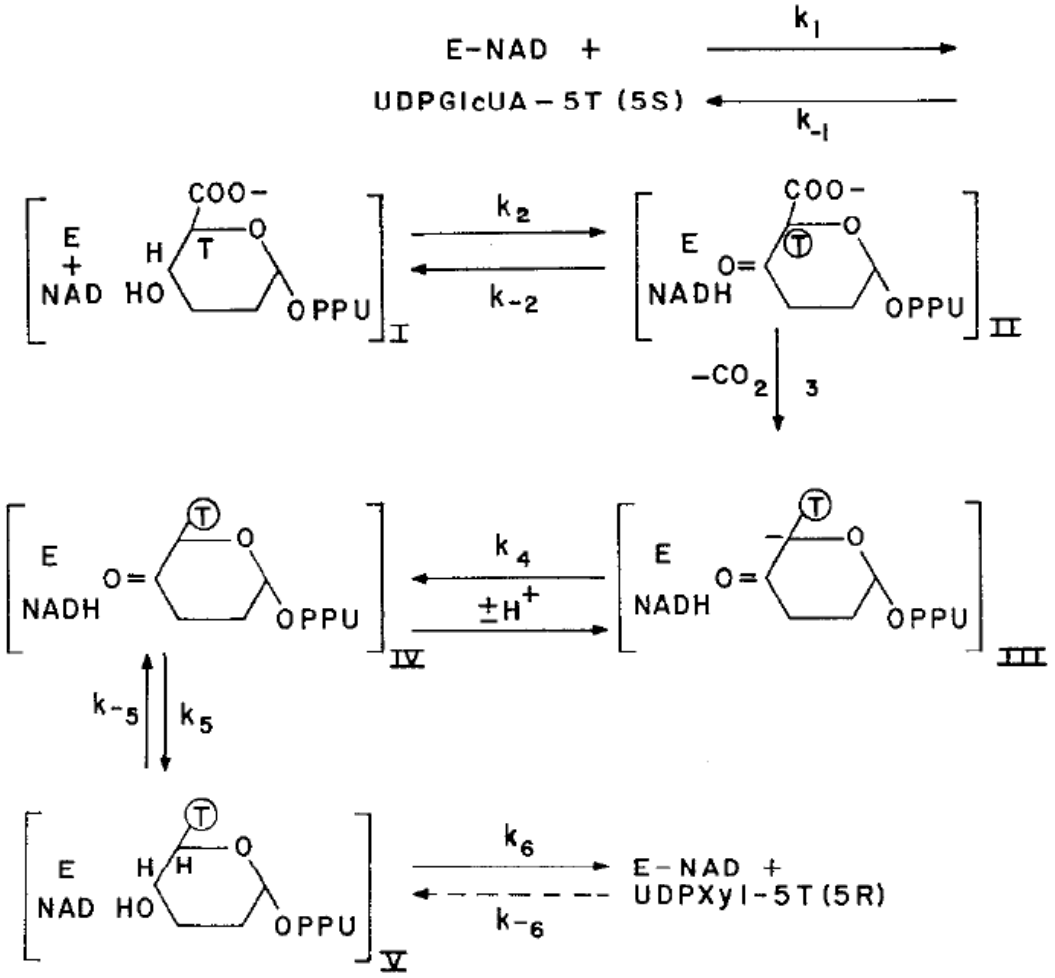
LITERATURE REVIEW

I. Mammals have hexose based UDP-xylose biosynthetic pathways.

UDP-xylose is directly synthesized from glucose via the sequential action of five enzymes.[9] First, glucose enters the cytoplasm of a cell and is phosphorylated at the C-6 hydroxyl by hexokinase (EC 2.7.1.1) as part of glycolysis.[10] Second, phosphoglucomutase (EC 5.4.2.2) removes the C-6 phosphate and adds a C-1 phosphate as part of glycogenesis.[11] Another glycogenic enzyme, glucose-1-phosphate uridylyltransferase (EC 2.7.7.9), then uses UTP to add UMP at the C-1 phosphate.[12] The resulting UDG is shunted from energy metabolism and oxidized twice at the C-6 position by UDP-glucose dehydrogenase (UGDH; EC 1.1.1.22) using two molecules of NAD^+.[13; 14] The UGA product is transported from the cytosol into the endoplasmic reticulum (ER) by UDP-galactose transporter-related isozyme 7 (UGTrel7).[15] UXS (EC 4.1.1.35) is located within the ER/golgi[16] and catalyzes the NAD^+-dependent decarboxylation of the UGA C-6 carboxyl group to form UDX.[7; 17] Mammals can also synthesize UDX by using the Leloir pathway and a UXS homolog, UDP-galactose 4-epimerase (GALE; EC 5.1.3.2), to convert galactose into the substrate for the UGDH reaction, UDG.[9] An alternative salvage pathway for xylose does not appear to exist in Chinese hamster ovary (CHO) cells[18] and is not believed to occur in mammals.

II. UDP-glucuronic acid decarboxylation occurs via elimination of a β-keto acid.

UDX formation has long been proposed to occur via β-keto elimination of the UGA carboxylate rather than by direct displacement (Figure 1.1).[19; 20] Early studies were unable to directly identify the transient UDP-keto-sugar intermediate. Instead, evidence supporting this hypothesis was provided through intensive radiolabeling experiments. Native UXS from wheat germ (*Triticum aestivum*; *Ta*UXS) and *Cryptococcus laurentii* (*Cl*UXS) display isotope effects when UGA is tritium labeled at the C-4 position, but no effect is observed for substrate that is C-3 or C-5 labeled.[20] The observed isotope effects are consistent with hydride transfer between the substrate C-4 and NAD(H). The first hydride transfer is believed to form UDP-4-keto-glucuronic acid (UGA4O) and NADH followed by elimination of the UGA4O carboxylate to create a C-5 carbanion.[20] Stereochemistry at C-5 is inverted during catalysis,[20] consistent with a planar enol that is intermediate to the *S* configuration of UGA and the *R* configuration of UDX. The inversion of absolute configuration suggests UXS retains substrate during the catalytic cycle and stereospecifically protonates the C-5 position to form UX4O. Tritium label at the C-4 position is not lost during UDX formation,[19; 21] indicating the same hydride transferred from the C-4 of UGA is returned to the C-4 of UX4O. Inversion of configuration does not occur at the C-4 position of UDX,[20] suggesting substrate adopts a similar conformation for both hydride abstraction and donation. The magnitude of the isotope effect suggests that hydride transfer is the rate-limiting step in UXS catalysis.[20] However, the isotope effect does not distinguish between the sequential hydride transfers that first generate UGA4O and then produce UDX.



Reproduced with permission from Schutzbach, J.S. and Feingold, D.S. Biosynthesis of uridine diphosphate D-xylose. IV. Mechanism of action of uridine diphosphoglucuronate carboxy-lyase. *J Biol Chem*. 1970; 245: 2476-2482.

Figure 1.1. Proposed mechanism for UDP-glucuronic acid (UDPGlcUA) decarboxylation and UDP-xylose (UDPXyl) formation. UXS contains an NAD^+ cofactor and reversibly binds substrate UDPGlcUA (k_1, k_{-1}) to form the Michaelis complex (I). Next, NAD^+ reversibly extracts a hydride (k_2, k_{-2}) from the substrate C-4 to form UDP-4-keto-glucuronic acid (II). II is then decarboxylated irreversibly leaving a C-5 carbanion (III). Enolization of III results in a planar configuration at C-5. Protonation of III at C-5 (k_4) is stereospecific and inverts the 5-S configuration of tritiated (T) substrate to form 5-R UDP-4-keto-xylose (IV). NADH reversibly returns the hydride to IV at C-4 (k_5, k_{-5}) to form UDPXyl (V). UDPXyl reversibly dissociates from UXS (k_6, k_{-6}).

III. UGA decarboxylases use NAD^+ as co-substrate or cofactor to produce species-specific UDP-sugars.

A. Bacteria, Protista and Fungi

Almost four decades after the UGA labeling experiments discussed above, UGA decarboxylase genes with 55-62% amino acid sequence identity have been annotated in archaea, bacteria and eukaryote species (www.sdr-enzymes.org, Table 1.1).[7; 22] Ironically, the first bacterial UGA decarboxylase to be characterized, *E. coli* ArnA (previously PrmI), produces NADH and the elusive UX4O using NAD^+ and UGA as co-substrates (Table 1.2).[23; 24] ArnA deviates from the UDX mechanism in order to provide UX4O (a.k.a. UDP-4-keto-arabinose) for Lipid A metabolism in pathogenic bacteria.[25; 26] In the plant pathogenic bacteria *Ralstonia solanacearum*, the Lipid A metabolic genes are present but the N- and C-terminal domains of ArnA are encoded as distinct genes separated by 10 nucleotides.[27] The analogous UGA decarboxylase domain also has an absolute requirement for NAD^+ and produces NADH and UX4O.[28] Interestingly, this UGA decarboxylase converts NADH and UX4O substrates to NAD^+ and UDX and was classified as a bifunctional UDP-4-keto-pentose/UDP-xylose synthase (*Rs*U4kpxs).[28] Additional UGA decarboxylases from the intestinal microbe *Bacteroides fragilis* (*Bf*Uxs1-2)[29] and the parasitic protozoan *Trichomonas vaginalis* (*Tv*UXS)[30] have recently been shown to form UDX in the presence of NAD^+ (Table 1.2). Although xylose has not been reported in the surface glycans of *R. solanocearum* and *B. fragilis*, UDX is hypothesized to be the donor for xylosylation of surface glycans in *T. vaginalis*[31] and bacterial pathogens[32] and endosymbiotes.[33]

Several bacteria that decarboxylate UGA have established metabolic pathways to utilize UDX. The UGA decarboxylase from *Streptomyces viridochromogenes* (AviE2) is reported to produce UDX using UGA and NAD^+ as substrate (Table 1.2).[34] The AviE2 product is required

Table 1.1. Primary accession numbers of experimentally confirmed UGA decarboxylases.

Organism	Common name	Enzyme	Accession	Reference
E. coli		ArnA	AY057445	Breazeale 2002 [23]
R. solanacearum		U4kpxs	GQ369438	Gu 2010 [28]
S. viridochromogenes		AviE2	AAK83183	Hofmann 2005 [34]
M. echinospora		CalS9	AAM70333	Simkhada 2009 [35]
B. fragilis		*Bf*Uxs1	CAH07883	Coyne 2011 [29]
		*Bf*Uxs2	CAH07260	Coyne 2011 [29]
S. meliloti		*Sm*Uxs1	GU062741	Gu 2011 [36]
T. vaginalis		*Tv*UXS	HE575670	Rosenberger 2012 [31]
C. laurentii		*Cl*UXS		Ankel 1966 [21]
C. neoformans		*Cn*UXS	AAK59981	Bar-Peled 2001 [37]
T. aestivum	Wheat germ	*Ta*UXS		Ankel 1965 [19]
A. thaliana	Thale cress	*At*UXS1	NP190920	Harper 2002 [3]
		*At*UXS2	NP191842	Pattathil 2005 [6]
		*At*UXS3	NP200737	Harper 2002 [3]
O. sativa	Rice	*Os*UXS1	AB167397	Suzuki 2004 [5]
G. hirsutum	Cotton	*Gh*UXS1	EU817581	Pan 2010 [38]
H. vulgare	Barley	*Hv*UXS1	AY677177	Zhang 2005 [39]
P. sativum	Pea	*Ps*UXS1	AB059568	Kobayashi 2002 [4]
N. tabacum	Tobacco	*Nt*UXS1	AY619950	Bindschedler 2005 [40]
C. elegans	Flatworm	SQV-1	AAN39843	Hwang 2002 [41]
G. gallus	Chicken	*Gg*UXS	XP416926	John 1977 [2]
R. norvegicus	Rat	*Rn*UXS	AAM45939	Moriarty 2002 [17]
H. sapiens	Human	*Hs*UXS	AAN39844	Hwang 2002 [41]
		*Hs*UXS splice variant	Q8NBX3	Hwang 2002 [41]

Table 1.2. Effects of exogenous NAD(P)(H) on UGA decarboxylase activity.

	Relative activity (%)					
Enzyme	−NAD^+ [a]	+NAD^+	+NADH	+$NADP^+$	+NADPH	Reference
ArnA	0	100				Breazeale 2002 [23]
U4kpxs	0	100				Gu 2010 [28]
AviE2	n/d	100				Hofmann 2005 [34]
CalS9	n/d	100				Ahlert 2002 [42]
*Sm*Uxs1	~10	100				Gu 2011[36]
*Cl*UXS	0	100				Ankel 1966 [21]
*Cn*UXS	100	204				Bar-Peled 2001 [37]
*Ta*UXS	100	100				Ankel 1965 [19]
*At*UXS1	100					Harper 2002 [3]
*At*UXS2	100	113	124	103	118	Pattathil 2005 [6]
*At*UXS3	100	110	104	100	100	Harper 2002 [3]
*Os*UXS1	100					Suzuki 2004 [5]
*Os*UXS2	100	108	107	85	126	Suzuki 2004 [5]
*Os*UXS3	100	70	119	22	196	Suzuki 2004 [5]
*Ps*UXS1	100	~170	~125			Kobayashi 2002 [4]
*Nt*UXS	100	100				Wheatley 2002 [43]
SQV-1	n/d	100				Hwang 2002 [41]
*Rn*UXS	n/d	100				Moriarty 2002 [17]
*Gg*UXS	100	180				John 1977 [2]
hUXS	100					Eames 2010 [7]
hUXS	100	138				Eixelsberger 2012 [44]

[a] Most enzymes have full activity (100) using endogenous cofactor or no activity (0) without supplied cofactor. Several enzymes were assayed only in the presence of exogenous NAD^+ (n/d) and it is unclear whether exogenous cofactor is absolutely required. Data for NADH, $NADP^+$ and NADH are given for assays in which exogenous NAD^+ is not supplied. Data originally presented in graphical form are represented here as estimates (~).

for biosynthesis of the antibiotic Avilamycin A.[45] *Micromonospora echinospora* ssp. *calichensis* also encodes an NAD^+-dependent UXS (CalS9)[42] implicated in the biosynthesis of 16-membered macrolide antibiotics.[35] Recently, a UGA decarboxylase from the plant symbiont *Sinorhizobium meliloti* (*Sm*Uxs1) was shown to have UXS activity in the presence of NAD^+ substrate.[36] *Sm*Uxs1 also displays residual UDX production in the absence of exogenous NAD^+ (Table 1.2), suggesting that unlike other bacterial UGA decarboxylases, a portion of the enzyme purifies with cofactor bound. UDX formed by *Sm*Uxs1 is converted to UDP-arabinose by a downstream epimerase and is required for proper host-symbiote interactions during nodule formation.[36]

Similar to bacterial UGA decarboxylases, fungal UXS enzymes require NAD^+ for robust activity (Table 1.2). Natively purified *Cl*UXS has an absolute requirement for exogenous NAD^+ and is inhibited by NADH (K_i = 3μM).[21] Recombinantly expressed UXS from the pathogenic fungus, *Cryptococcus neoformans* (*Cn*UXS), generates UDX using an endogenously bound NAD^+ cofactor. *Cn*UXS also displays a 2-fold increase in activity in the presence of exogenous NAD^+ and is inhibited by NADH,[37] suggesting the active site may be exposed. In *Cryptococcus* species, UDX is required for capsular polysaccharide biosynthesis.[46; 47]

B. Plants

Plant UXS enzymes appear to use a tightly bound NAD^+ cofactor to generate UDX (Table 1.2). In turn, UDX serves as the donor substrate for the xylose found in cell wall polysaccharides.[48] *Ta*UXS only releases NAD^+ upon heat denaturation and is not sensitive to NADH inhibition.[19] Also consistent with a tightly bound cofactor, UXS isoforms from *Arabidopsis thaliana* (*At*UXS1-3) synthesize UDX without exogenous NAD^+ and are not inhibited by NADH.[3; 6] UXS from rice (*Oryza sativa*; *Os*UXS1-3), pea (*Pisum sativum*;

*Ps*UXS1), tobacco (*Nicotiana tabacum*; *Nt*UXS), and barley (*Hordeum vulgare*; *Hv*UXS1) are active using endogenous cofactor, as well.[4; 5; 39; 40; 43] Further investigation into the stoichiometry of NAD^+ binding revealed approximately 90% of recombinant *Hv*UXS1 purifies with cofactor bound.[39] Although other plant UXSs are likely to purify with similar amounts of tightly bound NAD^+, some enzymes display a sensitivity to exogenous NAD(P)(H) that suggests active site accessibility. Recombinant *Ps*UXS1 displays 50% greater activity with additional NAD^+ cofactor[4] while *At*UXS2-3 and *Os*UXS2 show more moderate activation (8-13%, Table 1.2). *Os*UXS3 is reported to be inhibited 30% by NAD^+, however, and stimulated almost 200% by free NADPH (Table 1.2).[5]

Some plants contain an additional NAD^+-dependent UGA decarboxylase (UDP-apiose/UDP-xylose synthase, UAXS) that produces both UDP-apiose (UDA) and UDX. Unlike plant UXS enzymes, native UAXS from parsley does not purify with a tightly bound NAD^+,[49] and is inhibited by NADH.[50] Parsley UAXS is reported to use NAD^+ to form an enzyme bound 4-keto intermediate during catalysis.[51] UAXS from potato also requires NAD^+ and UGA as substrates, but releases stable quantities of UX4O during the reaction.[52] Apiose is derived from the activated donor molecule UDA and facilitates borate crosslinking between rhamno-galacturonan II chains in plant cell walls.[53] Although UX4O is proposed to be a suitable substrate for additional sugar modifications,[52] currently there is no accepted metabolic role for UX4O in plants.

C. Animals

In animals, UDX is required to initiate synthesis of heparin/heparan sulfate and chondroitin sulfate/dermatan sulfate proteoglycans.[9; 54] UDX is also the substrate for xylosylation

of epidermal growth factor-like (EGF) repeats found in Notch[55] and coagulation factors.[56; 57] Only two animal UXS enzymes have been reported to purify in an active form with NAD^+ bound (Table 1.2). Recombinant hUXS converts UGA to UDX without additionally supplied NAD^+.[7; 44] UXS from cultured chicken chondrocytes (*Gg*UXS) is also active as purified.[2] Washing the cellular fraction containing *Gg*UXS does not appear to impact activity and suggests NAD^+ is bound tightly in the enzyme active site. In the presence of exogenous NAD^+ however, hUXS and *Gg*UXS activity are stimulated 1.4- and 1.8-fold, respectively.[2; 44] Additional animal UGA decarboxylases have been assayed in the presence of NAD^+, possibly due to the activation observed in other animals and plants discussed previously. UXS from *Rattus norvegicus* (*Rn*UXS) and *Caenorhabditis elegans* (SQV-1) actively convert UGA to UDX in the presence of NAD^+,[17; 41] but it is not known how tightly the cofactor is bound in the active site (Table 1.2). It remains unclear whether NAD^+ stimulates activity via a common mechanism in animal UXS enzymes or simply reactivates apo-enzyme.

IV. UXS is activated or inhibited by UDP-sugars and uridine containing nucleotides.

While NAD^+ is generally associated with UXS activity (Table 1.2), uridine-containing molecules both activate and inhibit UGA decarboxylation (Table 1.3). *Gg*UXS activity is stimulated by increasing UGA substrate and displays sigmoidal steady state kinetics with a Hill coefficient, $n = 2.8$ (Figure 1.2).[2] Two natively purified *Ta*UXS isoforms are also cooperatively activated by UGA ($n = 2$) at low substrate concentrations ($\leq 20\mu M$; Figure 1.3).[1] However, at higher substrate concentrations ($\geq 70\mu M$), Hill coefficients for both *Ta*UXS isoforms change to $n = 1$ (Figure 1.4). In a similar manner, UDG activates *Ta*UXS1-2 at low UGA concentrations ($\leq 5\mu M$; Figure 1.5) but has decreased potency as substrate concentrations increase ($\geq 100\mu M$).[1]

Table 1.3. Effects of phospo-uridine based molecules on UXS activity.

	Relative activity (%) [a]							
Enzyme	UMP	UDP	UTP	UGA	UDG	UGal	UDX	Reference
*Cn*UXS		29			68		58	Bar-Peled 2001 [37]
*Ta*UXS1				(2.0) [c]	200 [d]		17 (2.5) [e]	John 1977 [1]
*Ta*UXS2				(2.0) [c]	250 [d]		72 (3.0) [e]	John 1977 [1]
*At*UXS2	79	0	0		80	91	46	Pattathil 2005 [6]
*At*UXS3	85	17	14		93	94	65	Harper 2002 [3]
*Os*UXS2	101	23 [0.16] [b]	5 [0.07] [b]		72	100	59	Suzuki 2004 [5]
*Os*UXS3	87	44 [0.16] [b]	5 [0.009] [b]		102	117	61	Suzuki 2004 [5]
*Hv*UXS1					85	95	17	Zhang 2005 [39]
*Gg*UXS				(2.8) [c]			50	John 1977 [2]

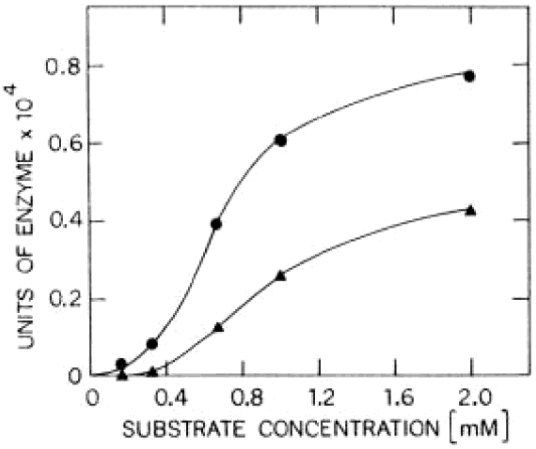
[a] UGA decarboxylase activity, relative to control UXS reaction with water or no additive, representing the greatest reported change in activity. UXS has not been shown to metabolize UDP-glucose (UDG) or UDP-galactose (UGal)

[b] Concentration of additive (mM) required for 50% inhibition

[c] Maximum Hill coefficient (*n*) reported for enzyme activation

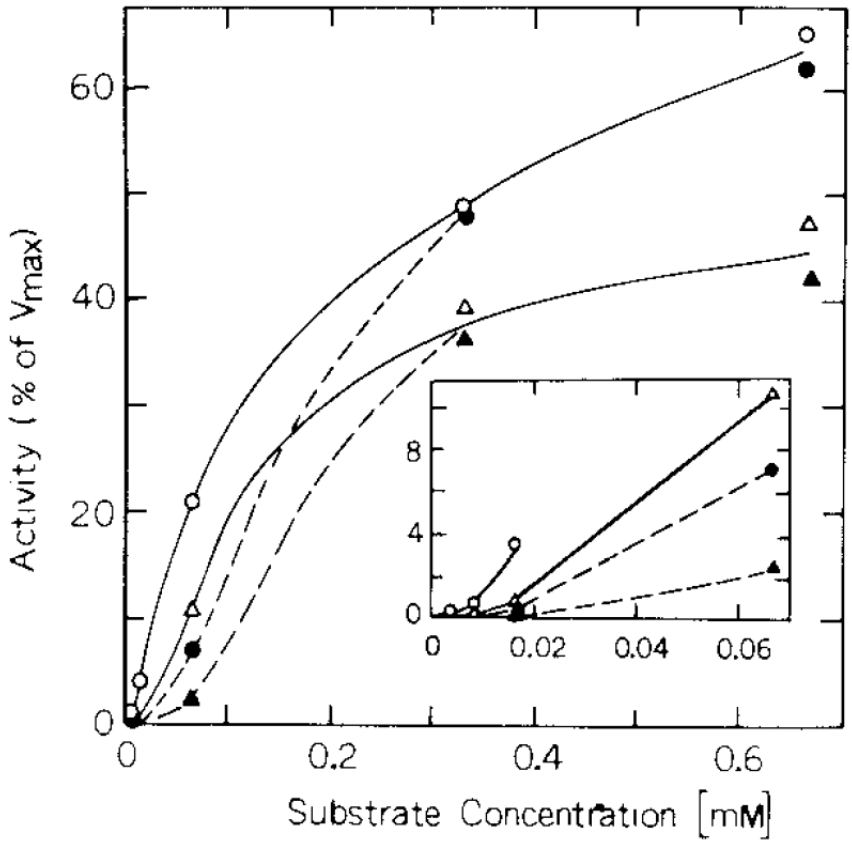
[d] Activation by UDG reported at UGA concentrations (2-5μM) below the apparent K_m for UGA (180μM or 530μM). All other non-cooperative enzymes assayed with UGA concentrations ≥ K_m

[e] Hill coefficient (*n*) for enzyme inhibition



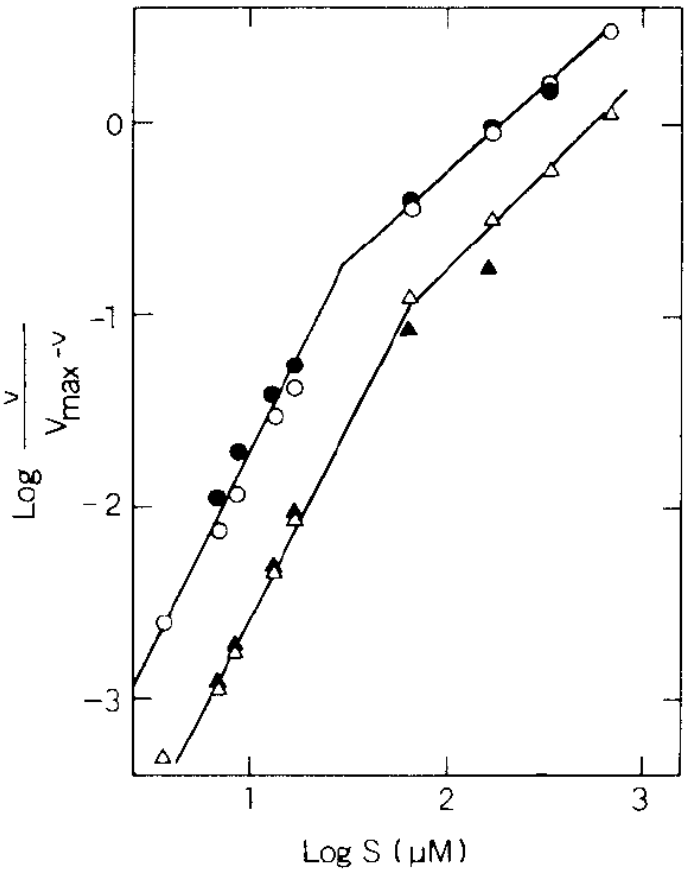
Reproduced with permission from John, K.V, Schwartz, N.B. and Ankel, H. UDP-glucuronate carboxy-lyase in cultured chondrocytes. *J Biol Chem*. 1977; 252: 6707-6710.

Figure 1.2. Dependence of reaction rate of *Gg*UXS on substrate concentration in the absence (●) or presence (▲) of 0.33 mM UDX. Natively purified enzyme from chicken chondrocytes was assayed for the release of $^{14}CO_2$ following incubation with NAD^+ (2mM) and varying concentrations of UGA uniformly labeled with ^{14}C in the glucuronosyl moiety.



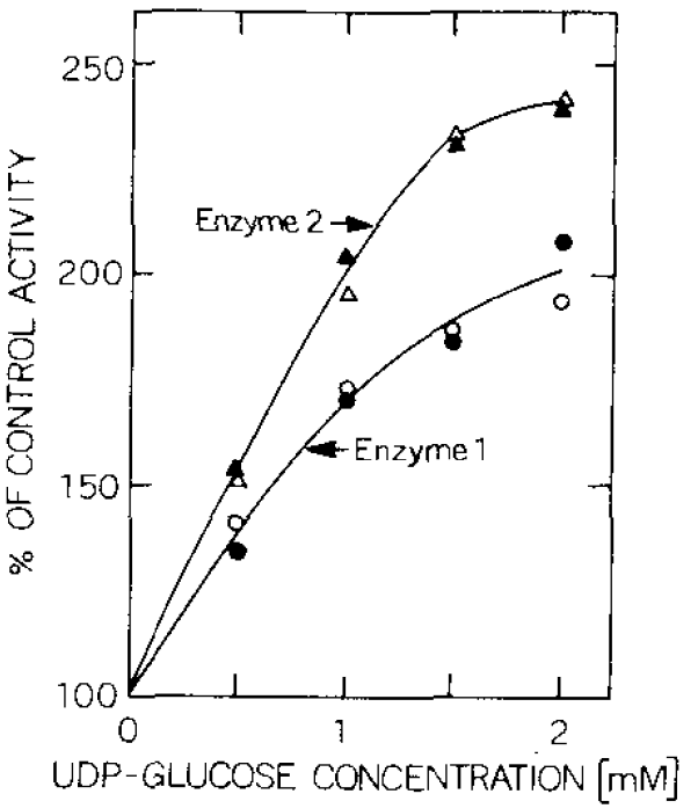
Reproduced with permission from John, K.V, Schutzbach, J.S. and Ankel, H. Separation and allosteric properties of two forms of UDP-glucuronate carboxy-lyase. *J Biol Chem*. 1977; 252: 8013-8017.

Figure 1.3. Substrate dependence of *Ta*UXS1 (o, ●) or *Ta*UXS2 (∆, ▲) in the absence (o, ∆, solid line) or presence (●, ▲, dashed line) of 0.06 mM UDX. Natively purified enzyme from wheat germ was assayed for the release of $^{14}CO_2$ following incubation with varying concentrations of UGA uniformly labeled with ^{14}C in the glucuronosyl moiety. Inset depicts the lowest substrate concentrations.



Reproduced with permission from John, K.V, Schutzbach, J.S. and Ankel, H. Separation and allosteric properties of two forms of UDP-glucuronate carboxy-lyase. *J Biol Chem*. 1977; 252: 8013-8017.

Figure 1.4. Substrate dependence of *Ta*UXS1 (o, ●) or *Ta*UXS2 (∆, ▲) plotted according to the Hill equation.[58] Assays conducted as Figure 1.3, with partially purified (o, ∆) or more highly purified enzyme (●, ▲). V_{max} values derived from Lineweaver-Burk plots.



Reproduced with permission from John, K.V, Schutzbach, J.S. and Ankel, H. Separation and allosteric properties of two forms of UDP-glucuronate carboxy-lyase. *J Biol Chem*. 1977; 252: 8013-8017.

Figure 1.5. Activation of *Ta*UXS1 (o, ●) and *Ta*UXS2 (∆, ▲) by UDG. Assays were conducted as Figure 1.3, with UGA concentrations of 5µM (o, ∆) or 2µM (●, ▲).

The observed changes in activation and *n* suggest the presence of non-catalytic regulatory sites in *Ta*UXS.[1] More recent investigations into the effect of UDG have focused on higher substrate concentrations ([UGA] ≥ K_m) and indicate inhibitory roles for non-substrate UDP-sugars in *Cn*UXS,[37] *At*UXS2-3,[3; 6] and *Os*UXS2-3.[5] Generally, UDX causes greater inhibition than any other nucleotide-sugar examined (Table 1.3), including UDP-galactose (UGal), GDP-mannose,[59] GDP-glucose,[3] UDP-galacturonic acid and TDP-glucose.[6] UDX has been shown to increase *Gg*UXS cooperativity and reduce enzyme activity by 50% (Figure 1.2).[2] UDX also increases cooperativity in *Ta*UXS1 (*n* = 2.5) and *Ta*UXS2 (*n* = 3) causing 83 or 28% inhibition, respectively (Figure 1.3).[1] The nucleotides UTP and UDP are potent inhibitors of UXS, as well (Table 1.3). In *Os*UXS2, inhibition studies with UDP and UTP yielded IC50 values of 160μM and 70μM, respectively.[5] Interestingly, the soluble *Os*UXS3 is more strongly inhibited by UTP (IC50 = 9μM),[5] which is also present in the cytosol.[60]

V. Eukaryotic UXS localizes to the ER/Golgi using an N-terminal domain.

As mentioned above, some plants contain multiple isoforms of UXS. Six UXS isoforms were originally identified in the *A. thaliana* genome (*At*UXS1-6), and divided into membrane anchored and soluble forms (Table 1.4, Figure 1.6).[61] *At*UXS1, *At*UXS2 and *At*UXS4 share an N-terminal domain predicted to contain a type II membrane-spanning region.[3] Supporting the assignment of a membrane anchor, full-length *At*UXS2-GFP and a construct consisting of the N-terminal domain both sediment in the membrane fraction of transformed *Nicotiana benthamiana* leaves.[6] *At*UXS2 is extracted from the membrane fraction by detergents, but not by denaturants or salts, indicating an integral membrane association.[6] In confocal sections of *N. benthamiana* leaves, *At*UXS2-GFP displays a punctate fluorescence pattern and co-localizes with known Golgi

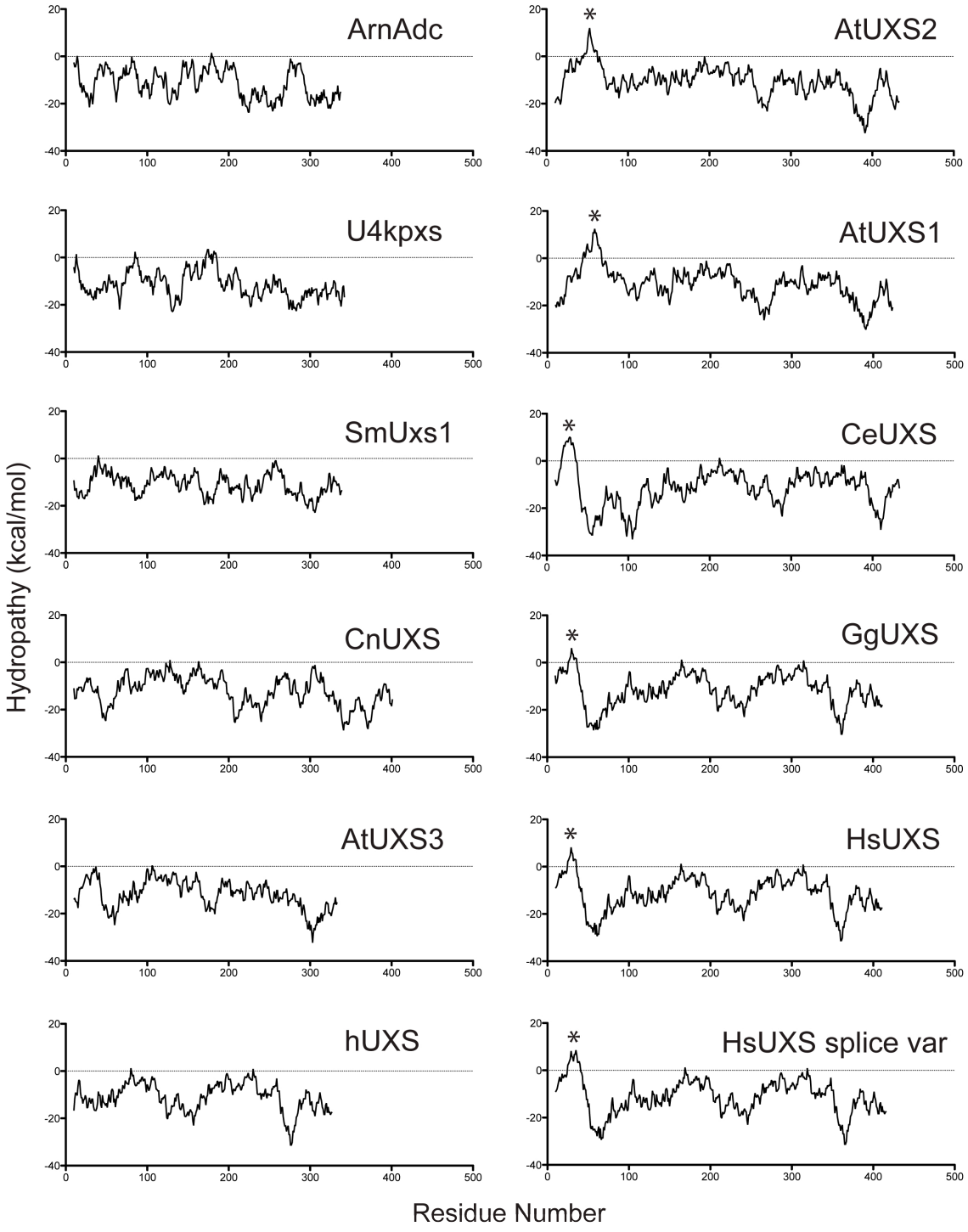
Table 1.4. Membrane association of eukaryotic UGA decarboxylases.

Enzyme	Membrane anchor	Reference
*Tv*UXS	−	Rosenberger 2012 [31]
*Cl*UXS	−	Ankel 1966 [21]
*Cn*UXS	−	Bar-Peled 2001 [37]
*Ta*UXS	−	Ankel 1965 [19]
*At*UXS1 [a]	+	Harper 2002 [3]
*At*UXS2	+	Pattathil 2005 [6]
*At*UXS3	−	Harper 2002 [3]
*At*UXS4	+	Harper 2002 [3]
*At*UXS5	−	Harper 2002 [3]
*At*UXS6	−	Suzuki 2004 [5]
*Os*UXS1 [a]	+	Suzuki 2004 [5]
*Os*UXS2	+	Suzuki 2004 [5]
*Os*UXS3	−	Suzuki 2004 [5]
*Os*UXS4	+	Suzuki 2004 [5]
*Os*UXS5	+	Suzuki 2004 [5]
*Os*UXS6	+	Suzuki 2004 [5]
*Gh*UXS1 [a]	+	Pan 2010 [38]
*Gh*UXS2	+	Pan 2010 [38]
*Gh*UXS3	−	Pan 2010 [38]
*Hv*UXS1 [a]	−	Zhang 2005 [39]
*Hv*UXS2	−	Zhang 2005 [39]
*Hv*UXS3	−	Zhang 2005 [39]
*Hv*UXS4	+	Zhang 2005 [39]
*Nt*UXS1 [a]	−	Bindschedler 2005 [40]
*Nt*UXS2	−	Bindschedler 2005 [40]
*Nt*UXS3	+	Bindschedler 2005 [40]
*Ps*UXS1	−	Kobayashi 2002 [4]
*Pv*UXS	−	Wheatley 2002 [43]
SQV-1	+	Hwang 2002 [41]
*Gg*UXS	+	John 1977 [2]
*Rn*UXS	+	Moriarty 2002 [17]
*Hs*UXS [a]	+	Hwang 2002 [41]
*Hs*UXS splice variant	+	Hwang 2002 [41]
hUXS [b]	−	Eames 2010 [7]

[a] Species with annotated isozymes or splice variants (dashed box)

[b] Construct engineered to be soluble

Figure 1.6. Hydropathy plots of representative UGA decarboxylases. UGA decarboxylases from *E. coli* (ArnAdc), *R. solanacearum* (U4kpxs), *S. meliloti* (SmUxs1) and *C. neoformans* (CnUXS) lack an N-terminal hydrophobic region. *A. thaliana* contains both soluble (AtUXS3) and membrane associated UXS isozymes (AtUXS1-2). A membrane associated UXS is present in flatworm (CeUXS), chicken (GgUXS) and rat (RnUXS). Full length human UXS (HsUXS) and a splice variant (HsUXS splice var) contain alternate transmembrane domains. Removal of the N-terminal 84 residues of HsUXS results in a soluble construct (hUXS). Transmembrane helices (*) were predicted using Wimbley-White whole-residue hydrophobicity and the octanol hydropathy scales implemented in the publicly available program Membrane Protein Explorer (http://blanco.biomol.uci.edu/mpex/).[62]


ArnAdc
AtUXS2
U4kpxs
AtUXS1
SmUxs1
CeUXS
CnUXS
GgUXS
AtUXS3
HsUXS
hUXS
HsUXS splice var
Hydropathy (kcal/mol)
Residue Number

markers.[6] Based on type II membrane topology, the N-terminus of a membrane associated UXS is expected to protrude into the cytosol with the C-terminal catalytic domain facing the lumen of the Golgi.[63] *At*UXS3, *At*UXS5 and *At*UXS6 lack the N-terminal domain conferring membrane localization.[3; 5] As a result, *At*UXS3-GFP localizes to the cytoplasm.[6] Recombinant *At*UXS1 and *At*UXS2 constructs lacking the transmembrane anchor are also soluble during purification.[3] Six UXS isoforms are predicted in *O. sativa* (*Os*UXS1-6), however only *Os*UXS3 lacks the N-terminal membrane-spanning region and is soluble.[5; 59] Similar to *At*UXS isozymes, deletion of the N-terminal membrane anchor prevents recombinant *Os*UXS accumulation in inclusion bodies and results in soluble enzyme.[5; 59] Localization of antibodies raised against *Nt*UXS supports the presence of a cytosolic form of UXS in the French bean, *Phaseolus vulgaris* (*Pv*UXS; Table 1.4), as well.[43]

In contrast to the multiple UXS isoforms encoded by plants, a single membrane anchored UXS appears to be responsible for UDX formation in animals. UGA decarboxylation in *G. gallus* has been shown to localize with xylosyltransferase activity in the ER/Golgi.[16; 64] *Gg*UXS from cultured chicken chondrocytes sediments with the particulate fraction,[2] as would be expected for an ER/Golgi-associated protein (Figure 1.6). *C. elegans* SQV-1 contains an N-terminal transmembrane signal similar to those found in membrane associated plant homologs. In *C. elegans* cells, SQV-1 fluoresces as distinct subcellular foci and colocalizes with a known Golgi-associated transporter.[65] *Rn*UXS contains a putative type II membrane topology signal and partitions in the microsomal fraction of lysed rat brain.[17] Also consistent with membrane association, *Rn*UXS localizes to the perinuclear Golgi in transfected human embryonic kidney (HEK) 293 cells.[17] Full length *Homo sapiens* UXS (*Hs*UXS) encodes an alternatively spliced transmembrane helix at the N-terminus.[41] Splicing is predicted to alter the primary structure by

five amino acids, SFLLN,[41] and shift the transmembrane region from residues 20-38 to 26-44 (Figure 1.6). Cleavage of the *Hs*UXS N-terminus is reported to produce recombinant enzyme (hUXS) that is soluble and active.[7]

VI. UXS mutant phenotypes occur in fungi and animal model systems.

In the pathogenic yeast *Cryptococcus neoformans*, xylosylation is required for proper maturation of capsular polysaccharides.[46] Disruption of the *Cn*UXS gene by cassette mutagenesis (Δ*Cn*UXS) is not lethal to *C. neoformans*, however, and has little effect on yeast doubling times.[47] Instead, Δ*Cn*UXS yeast display a loss of detectable UDX and a dysregulated UDP-sugar phenotype, characterized by a ~60% increase in cellular UGA levels.[66] External phenotypes include a ~2.5-fold decrease in capsule thickness and disrupted surface glucuronoxylomannan formation.[47] *Cn*UXS is absolutely required for virulence and mice inoculated with Δ*Cn*UXS yeast show no sign of infection after 90 days, while mice treated with *Cn*UXS yeast die within 30 days.[47]

Similar to *C. neoformans*, the budding yeast *Saccharamyces cerevisiae* does not require UXS for survival. Although *S. cerevisiae* produces UDG, the yeast lacks endogenous UXS and UDGH activity and provides a platform for studying the UDX biosynthetic pathway with UXS variants.[67] Yeast expressing only *Arabidopsis* UGDH (*At*UGDH) are effectively UXS knockouts and show a dysregulated UDP-sugar phenotype characterized by 54% lower intracellular UDG levels and UGA accumulation.[67] Coexpression of *At*UGDH and *At*UXS3 returns UDG pools to wild-type levels and converts all available UGA to UDX, consistent with a regulatory role for UXS activity. *At*UGDH is more inhibited by UDX (K_i 4.9μM) than UGA (K_i 99μM),[67] likely due to large quaternary changes observed in human UGDH:UDX structures but not the

hUGDH:UGA:NAD crystal form (Appendix B).[8] As exemplified in *S. cerevisiae* and *C. neoformans*, mutants lacking the UXS gene display phenotypic defects from upstream dys-regulation of UDG metabolism and downstream loss of xylosylation.

Partial truncations of the UXS gene have also been shown to cause mutant phenotypes. Loss of the C-terminal 125 amino acids of hamster UXS (*Cricetulus griseus*; *Cg*UXS) results in a CHO cell line deficient in glycosaminoglycan biosynthesis and xylosylation of the EFG repeats of Notch, consistent with disrupted UDX synthesis.[18] Truncation of the *Cg*UXS C-terminus also results in a 200-fold increase in endogenous UGA compared to wild-type levels. Transfection with full length *Hs*UXS compliments the mutant phenotype and restores wild-type UDP-sugar pools.[18] However, transfection with a soluble *Hs*UXS construct lacking the N-terminal 37 amino acids only partially complements the UGA phenotype, resulting in 50-fold more UGA than wild-type CHO cells.[18] Since the cytosolic *Hs*UXS was designed to localize in the cytoplasm with UGDH, it is unclear from that study why UGDH activity is not fully regulated. However, a comparison of the cytosolic *Hs*UXS construct with the hydropathy analysis of full-length *Hs*UXS (Figure 1.6) shows a portion of the unanchored linker region remains in the cytosolic construct. It is possible that the untethered linker between the transmembrane helix and the catalytic domain may lead to increased enzyme degradation or impaired UXS activity.

Single and double amino acid substitutions in UXS have been linked to developmental defects in animal model systems. In *Caenorhabditis elegans*, the squashed vulva (Sqv) phenotype is caused by six *Ce*UXS (a.k.a. SQV-1) alleles containing a total of seven missense mutations (Table 1.5).[41; 68] *Ce*UXS mutants are characterized by a partially collapsed vulval invagination, altered oocyte maturation and maternal-effect lethality.[41; 68] Similar to the UXS lesions, mutations in other *C. elegans* genes that cause the Sqv phenotype (SQV-2 to -8) are

Table 1.5. Mutations in conserved UXS residues are linked to animal phenotypes.

Scientific/ Common name	Lesion	Substitution	Human equivalent	Reference
C. elegans/ flatworm	ku246	N185I	N137I	Hwang 2002 [41]
	n2828	E264K	E216K	Hwang 2002 [41]
	n2848	C278R	C230R	Hwang 2002 [41]
	n2820	T344M	T296M	Hwang 2002 [41]
	n2824	R345Q	R297Q	Hwang 2002 [41]
	n2819	P407L	P358L	Hwang 2002 [41]
	n2819	R410G	R361G	Hwang 2002 [41]
D. rerio/ zebrafish	*mow*^{w60}	R233H	R236H	Eames 2010 [7]

proposed to be recessive and represent a loss of gene function.[68] A single amino acid substitution in *Danio rerio* UXS (Table 1.5) causes a *man o'war* (*mow*) phenotype that is also consistent with loss of *Dr*UXS function (discussed at length in Appendix A).[7] Zebrafish that are homozygous for either the *mow* lesion or a retroviral insertion into the *Dr*UXS gene show reduced proteoglycan levels and disrupted craniofacial development. Mutation of the equivalent *mow* amino acid in hUXS abolishes UDX production *in vitro* (Appendix A)[7] and supports the assignment of *mow* as a null allele *in vivo*. Collectively, the missense mutations that disrupt decarboxylase activity also highlight the functional significance of single amino acids conserved during UXS evolution.

VII. UGA decarboxylases are members of the 'extended' short-chain dehydrogenase/reductase family of enzymes.

Based on amino acid sequence, UGA decarboxylases are classified as members of the short-chain dehydrogenase/reductase (SDR) superfamily of enzymes.[22; 24] SDR enzymes are typically ~250 residues in length and have a common pattern of motifs.[69] The amino-termini of SDR enzymes contain a Gly-rich motif (Figure 1.7) for binding the dinucleotide cofactor, NAD(P)(H).[70] NAD^+ binding is usually facilitated by an Asp located ~20 amino acids downstream of the Gly-rich motif. In SDR enzymes that prefer NADP(H), the Asp interaction is often replaced by one or two basic residues that can interact with the adenine ribose 2'-phosphate of cofactor.[71; 72] The Gly-rich motif is followed by two conserved, extended α-helices (α3 and α4) (Figure 1.8).[69; 73] The first member of the SDR catalytic triad, Ser (or Thr), is located at the C-terminus of a β-strand (β5) separating α3 and α4. The remaining catalytic residues, Tyr-XXX-Lys, are located ~12 amino acid downstream in α4 (Figure 1.8).[24; 69] SDR enzymes form homodimers or homotetramers using the conserved helices as a packing surface. Enzymes in a

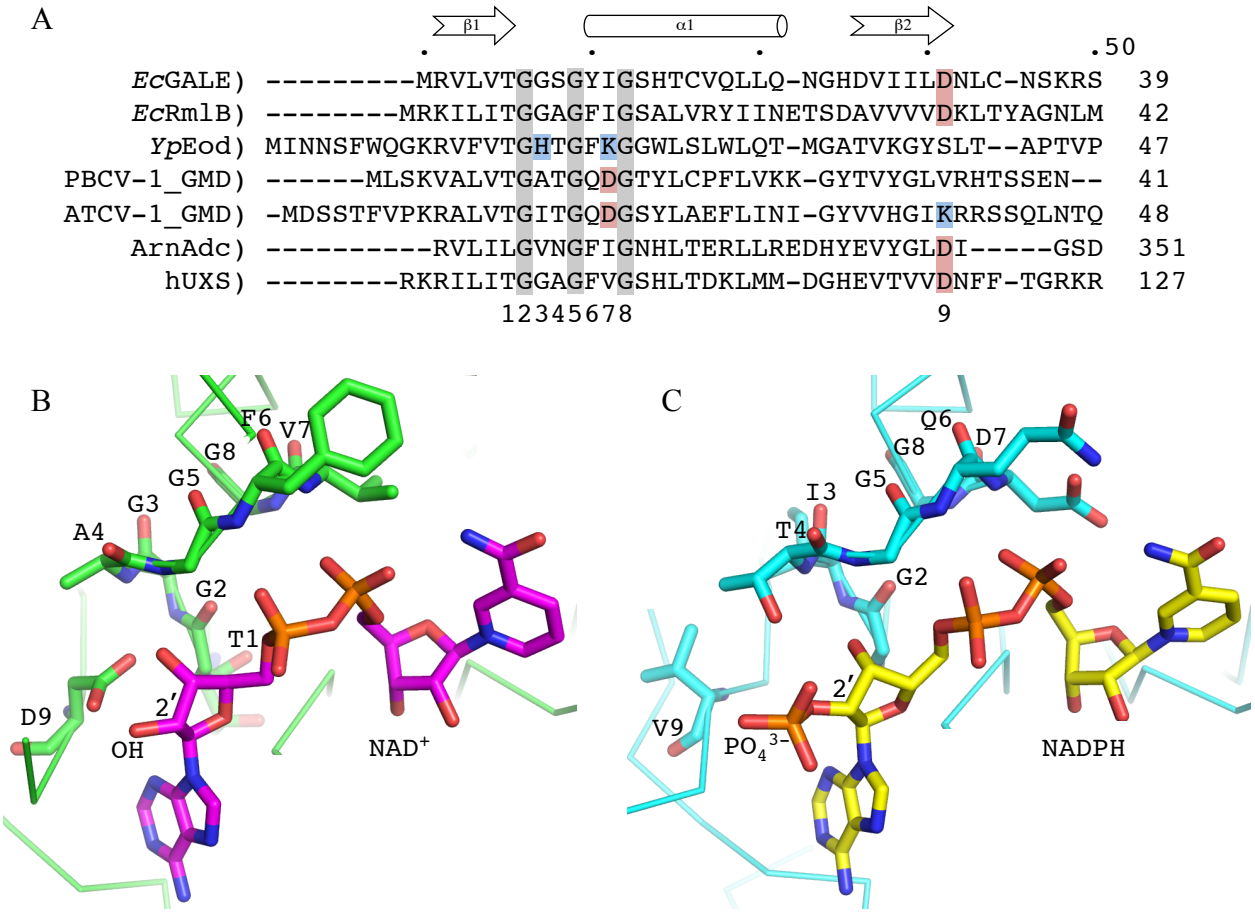


Figure 1.7. Short-chain dehydrogenase/reductase (SDR) enzymes conserve an amino terminal NAD(P)(H) binding motif. (A) Structure based sequence alignment of the SDR enzymes *E. coli* UDP-galactose 4-epimerase (*Ec*GALE, PDB entry 2UDP), *E. coli* dTDP-glucose 4,6-dehydratase (*Ec*RmlB, 1BXK), *Yersinia pseudotuberculosis* CDP-glucose 4,6-dehydratase (*Yp*Eod, 1RKX), chlorella virus PBCV-1 GDP-mannose 4,6-dehydratase (PBCV-1_GMD), chlorella virus ATCV-1 GDP-mannose 4,6-dehydratase (ATCV-1_GMD), *E. coli* ArnA UDP-glucuronic acid decarboxylase (ArnAdc, 2BLL) and human UDP-xylose synthase (hUXS, 2B69) performed using SEQUOIA.[74] A Ser or Thr residue (labeled 1) at the end of β-strand 1 often precedes the conserved GXXGXXG motif (labeled 2-8, shaded grey). An Asp residue at the end of β-strand 2 (labeled 9) is often conserved in enzymes that bind NAD^+ (e.g. *Ec*GALE, *Ec*RmlB, ArnAdc and hUXS) but is not found in the NAD^+-dependent *Yp*Eod. The Asp is replaced in enzymes that bind $NADP^+$ (e.g. PBCV-1_GMD and ATCV-1_GMD). Charged amino acids within the motif are shaded blue (positive) or red (negative). Secondary structures presented as α-helix (cylinder) or β-strand (arrows) correspond to hUXS. (B) Nicotinamide adenine dinucleotide binding motifs in the NAD^+ (purple sticks) bound hUXS (green sticks) and (C) NADPH (yellow sticks) bound *Pseudomonas aeruginosa* GDP-mannose 4,6-dehydratase (*Ps*GMD, cyan sticks, PDB entry 1RPN). Amino acid positions 1-9 in (A) are labeled by the carbonyl oxygen to show *Ps*GMD conserves residues 2-8 of ATCV-1_GMD and the 9 position Val of PBCV-1_GMD. Structures of the chlorella virus GMD enzymes are currently unavailable.

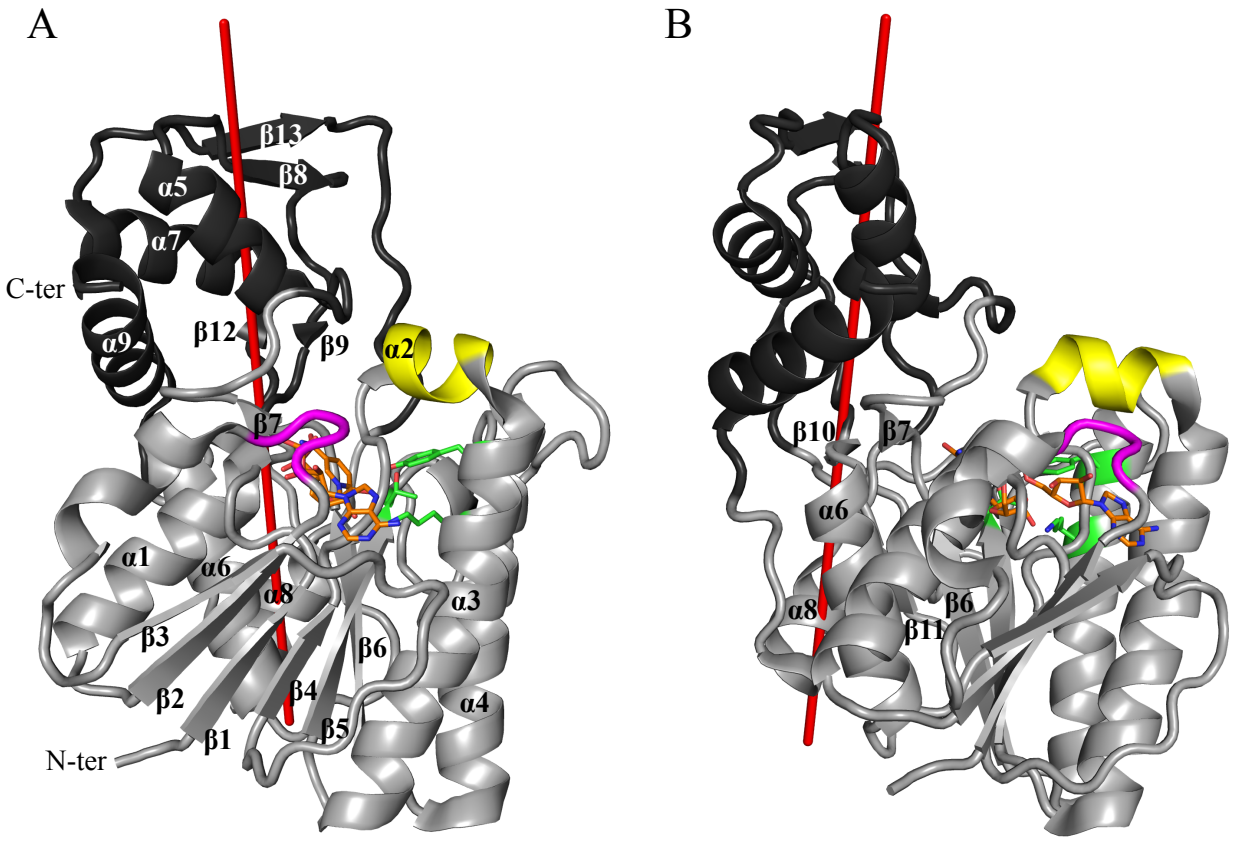


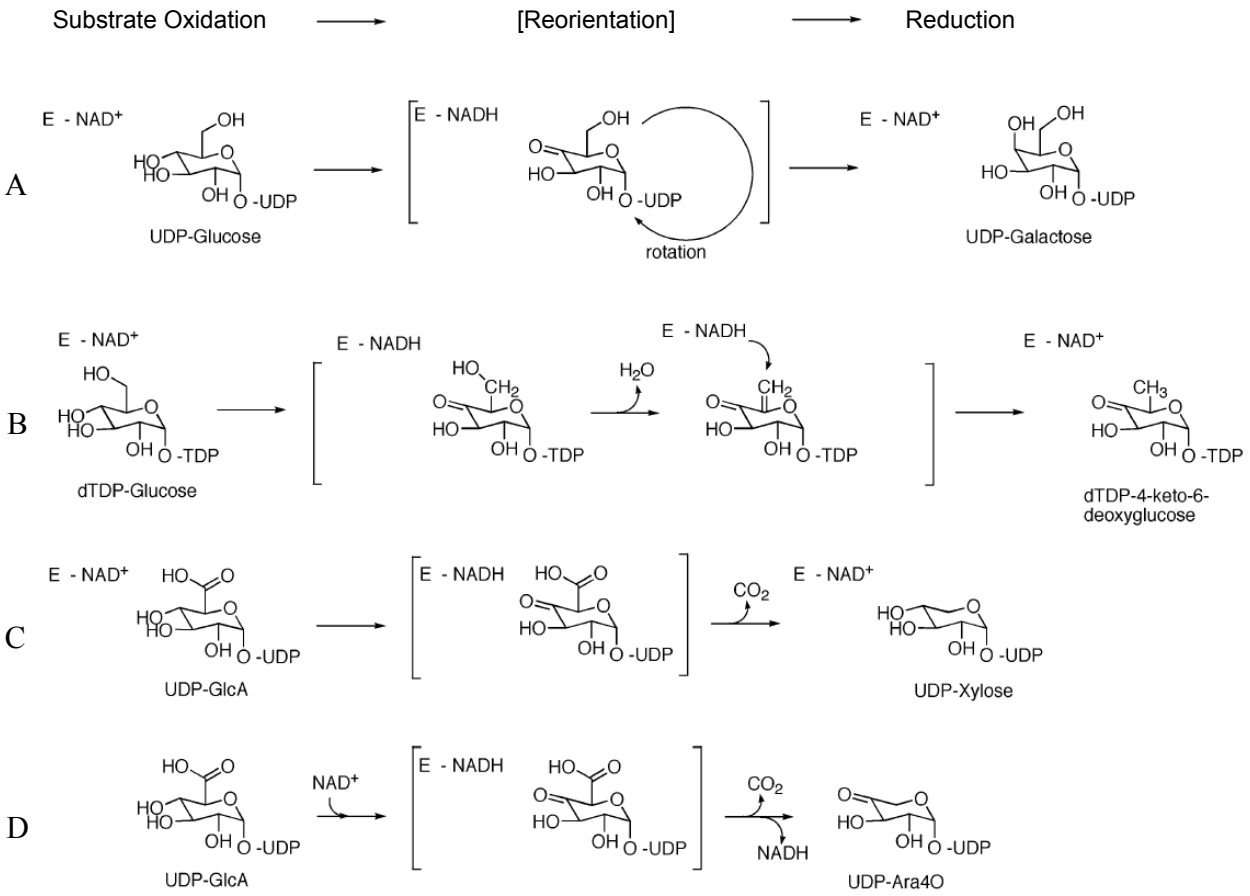
Figure 1.8. The extended SDR enzyme fold. (A) Cartoon representation of hUXS (PDB entry 3IAT, structure to be published by our group) showing the core catalytic domain (grey) of classical SDR enzymes is extended by ~100 amino acids organized into a nucleotide sugar binding domain (NSBD, black). Secondary structure topology is indicated for the alpha helices (α1-9) and beta strands (β1-13) identified using DSSP.[75] NAD^+ (orange sticks) is oriented in the catalytic domain by i) a divergent loop (magenta) between β2 and β3, ii) the α2 'thumb' helix (yellow) and N-terminal loop, and iii) the catalytic residues (green sticks) spanning β5 (Thr) and α4 (Tyr-XXX-Lys). DYNDOM [76] analysis of non-crystallographically related hUXS monomers reveals dynamics about an axis (red) that roughly bisects the NSBD and catalytic domains. α3 and α4 provide an extensive helical surface for catalytic domain dimerization, allowing the NSBD to rotate toward α2 and NAD^+. (B) 45° rotation of hUXS showing an additional view of structural elements buried in (A).

canonical SDR dimer are related by a two-fold axis perpendicular to the resulting four-helix bundle.[73] SDR tetramer structures build upon the four-helix bundle and form a symmetrical sandwich of canonical dimers.[77]

UGA decarboxylase contains ~100 additional amino acids organized into a nucleotide sugar binding domain (NSBD) and is further classified as an 'extended' SDR enzyme.[72] Unlike classical SDR enzymes, most extended SDRs prefer NAD(H) over NADP(H).[72] Cofactor binding in the extended family is facilitated by the Gly-rich motif, GXXGXXG, often directly preceded by a non-catalytic Ser or Thr (Figure 1.7).[73] Most extended SDRs use the conserved catalytic residues to catalyze the epimerization, dehydration or decarboxylation of nucleotide sugars.[73] All three reactions are believed to proceed through common steps (Figure 1.9), summarized as: i) proton abstraction from the substrate C-4 hydroxyl by the catalytic tyrosine and concomitant hydride transfer from the C-4 of substrate to NAD^+, ii) repositioning of the resulting C-4-keto-sugar within the active site, and iii) hydride transfer back to the substrate to regenerate NAD^+.[78] Extended SDR enzymes are present in organisms that lack classical SDRs[79] and are therefore hypothesized to resemble the ancestral SDR progenitor.[80] In order to understand the evolution and mechanistic diversity of SDR enzymes, the extended SDR family has been well-characterized structurally and biochemically.[78]

VIII. UDP-galactose 4-epimerase is the prototypical extended SDR enzyme.

UDP-galactose 4-epimerase (GALE) has been extensively studied (reviewed in references[81; 82; 83; 84]) and much of our structural understanding of the extended SDR subfamily comes from analyses of the ~30 sets of GALE atomic coordinates deposited in the Protein Data Bank (PDB, www.rcsb.org). In *E. coli* (*Ec*GALE) and *H. sapiens* (*Hs*GALE), GALE is a 39-40



Adapted with permission from Gatzeva-Topalova, P. Z., May, A. P., and Sousa, M. C. Crystal structure of Escherichia coli ArnA (PmrI) decarboxylase domain. A key enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance, *Biochemistry 43*, 13370-13379. Copyright (2004) American Chemical Society.

Figure 1.9. Schematic representation of extended SDR enzymes catalyzing the interconversion of nucleotide sugars. (A) UDP-galactose 4-epimerase (GALE), (B) dTDP-glucose 4,6-dehydratase (RmlB) and (C) UDP-xylose synthase (UXS) first use NAD^+ to oxidize substrate to an intermediate NDP-4-keto-sugar and NADH, then reorient the stable intermediate within the active site and finally transfer the hydride from NADH back to substrate. Reorientation of the stable intermediates (bracketed) varies by reaction. GALE rotates UDP-4-keto-glucose within the active site and presents the opposite face of substrate to NADH for C-4 reduction. RmlB dehydrates dTDP-4-keto-glucose and aligns the resulting dTDP-4-keto-glucose 5,6-ene with NADH for C-6 reduction. UXS decarboxylates UDP-4-keto-glucuronic acid and presents UDP-4-keto-xylose for C-4 reduction. UDP-4-keto-glucuronic acid has not been isolated as a stable intermediate and is therefore not included in the schematic. (D) ArnA is a bacterial UDP-glucuronic acid decarboxylase that uses NAD^+ substrate as a hydride acceptor, similar to classical SDR enzymes. ArnA releases NADH and UDP-4-keto-xylose (a.k.a. UDP-Ara4O) as product, rather than reorienting substrate for a second hydride transfer.

kDa, two lobed protein with an amino-terminal catalytic domain and a smaller carboxy-terminal NSBD.[85; 86] In the yeast *S. cerevisiae* and *Kluyveromyces marxianus* (a.k.a. *K. fraglis*), however, GALE comprises the amino-terminal half of the 65-75 kDa bifunctional galactose mutarotase/UDP- galactose 4-epimerase (Gal10p).[87; 88] Whether occurring as a mono- or bifunctional enzyme, GALE adopts superimposable tertiary structures (Figure 1.10).[87] The *Hs*GALE catalytic domain contains a dinucleotide binding fold consisting of six parallel β-strands organized into a pleated sheet and a total of six α-helices. The NSBD contributes a seventh parallel strand to the β-sheet to complete the modified Rossmann fold.[86] The Gly-rich motif found in extended SDR enzymes is vital to the function of the Rossmann fold[89] and places the first Gly at the end of the β-strand 1. The tight turn facilitated by the first Gly positions the second Gly at the amino-terminus of α-helix 1 and allows for a helix dipole interaction with the diphosphate of NAD^+. The third Gly permits close packing of the first α-helix and β-strand and is proposed to favor the binding of NAD^+ over $NADP^+$.[89] In *Hs*GALE and *Ec*GALE, NAD(H) binding is facilitated by 13-19 hydrogen bonds, respectively, between the cofactor and enzyme (Figure 1.11).[86] The catalytic YxxxK residues are found in the catalytic domain and make direct side chains interactions with NAD^+. NAD^+ also participates in an intramolecular hydrogen bond between the β-phosphate and the amide of the nicotinamide ring (Figure 1.11). As a result, NAD^+ adopts a syn-conformation relative to the ribose.[90] The syn-conformation is characteristic of class B dehydrogenases and presents the *si* face of the nicotinamide ring for pro-S hydride removal from substrate.[91]

The NSBD of *Hs*GALE is structurally simpler than the catalytic domain and contains six shorter β-strands and five α-helices organized into a UDP binding site.[86] The NSBD of *Ec*GALE also maintains an extensive network of interactions for binding UDP during sugar

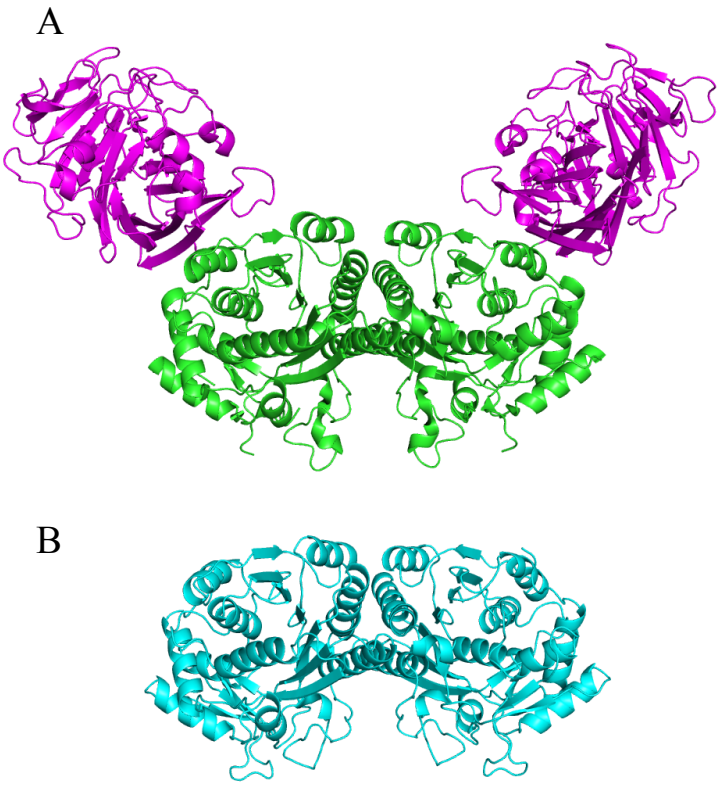


Figure 1.10. UDP-galactose epimerization occurs in mono- and bifunctional GALE enzymes. (A) Crystal structure of the bi-functional galactose mutarotase /UDP-galactose 4-epimerase (Gal10p) from *S. cerevisiae* (PDB entry 1Z45). The epimerase domain of Gal10p (green, residues 1-356) makes minimal contacts with the mutarotase domain (magenta, residues 357-699) and is homologous with (B) *Ec*GALE (cyan, 1XEL). Enzymes are oriented to look down the four helix bundle of the SDR dimer formed through crystal symmetry.

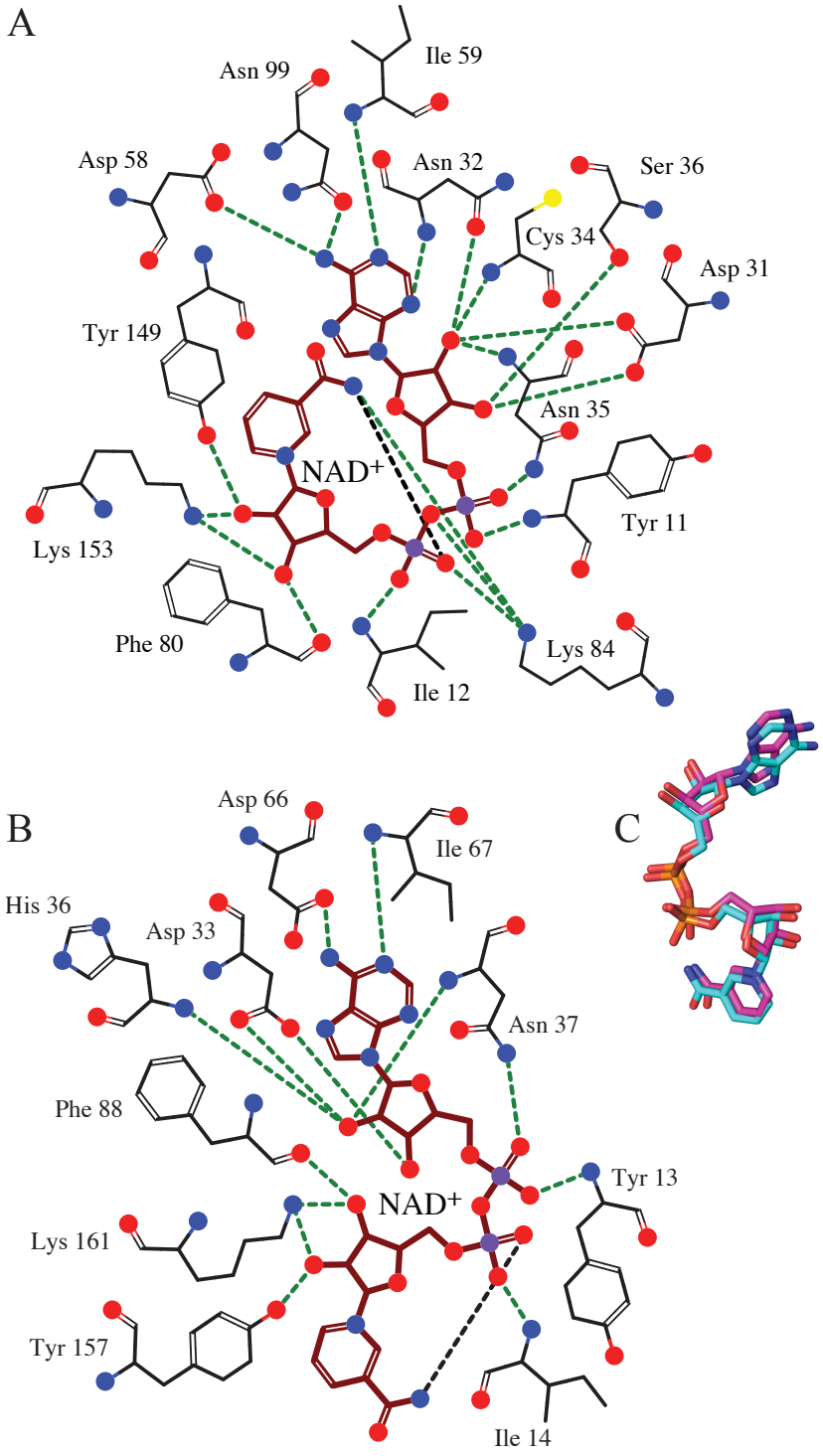
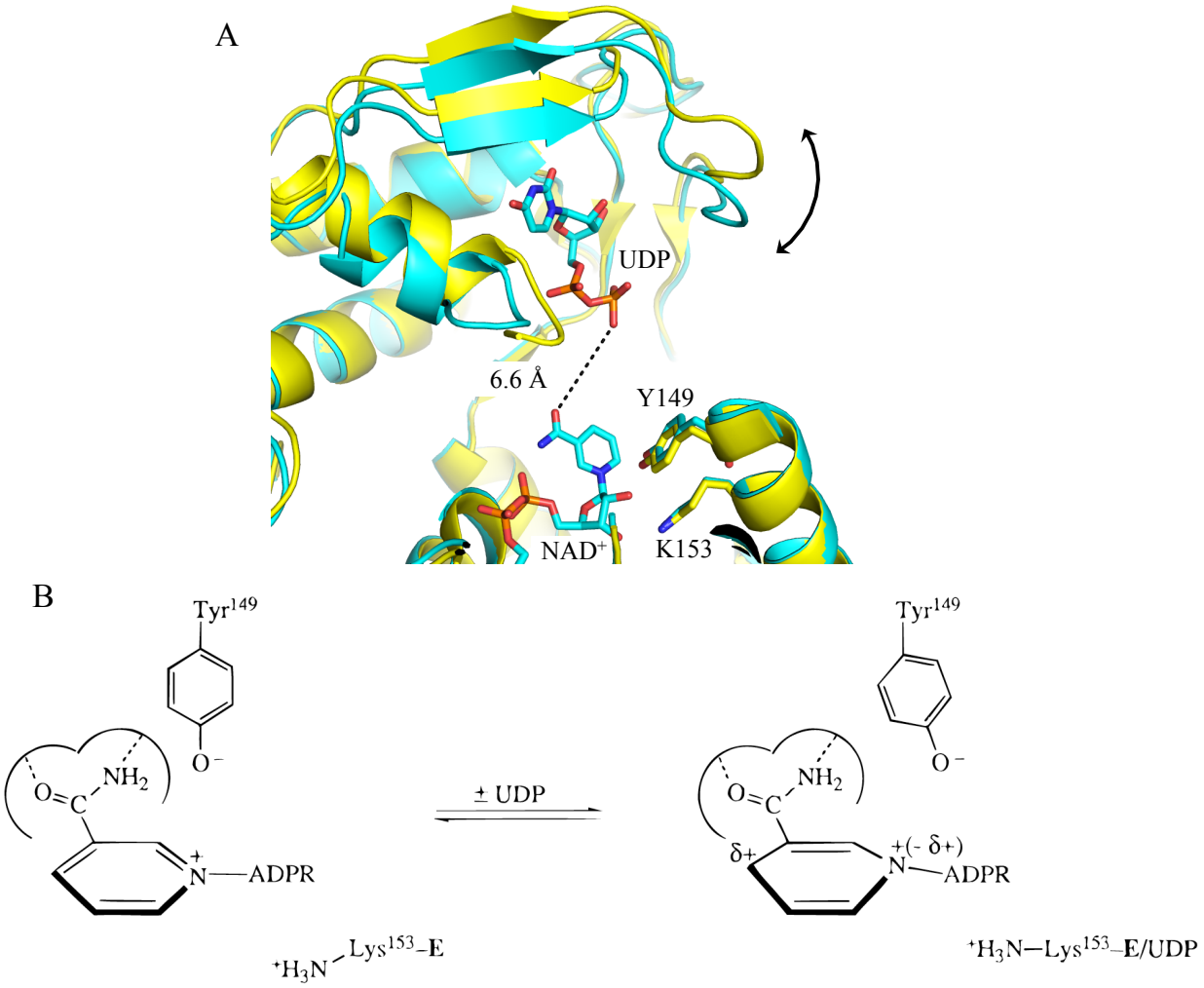


Figure 1.11. GALE binds the NAD^+ cofactor through extensive hydrogen bond interactions. (A) Hydrogen bonding network (green dashes) of amino acids (black sticks) coordinating NAD^+ (magenta sticks) in *Ec*GALE (PDB entry 1NAI) or (B) *Hs*GALE (1EK5). Oxygen (red), nitrogen (blue), sulfur (yellow) and phosphorus (purple) atoms are shown as spheres. A conserved intramolecular hydrogen bond is depicted by black dashes. Interaction networks were generated using the LIGPLOT[93] tool of PDBsum[94] and do not reflect (C) the similar conformations of NAD^+ in the *Ec*GALE (cyan) and *Hs*GALE (magenta) crystal structures.

modification.[81; 92] Fluorescence studies of *Ec*GALE support disproportionate binding of the substrate moieties and estimate that an average of 80% of the binding free energy is contributed by the nucleotide portion of ligands.[95] As such, the NSBD can accommodate several UDP-based molecules, such as UDP-phenol,[85; 96] UDP-4-deoxy-4-fluoro-α-D-galactose, UDP-4-deoxy-4-fluoro-α-D-glucoseor UDP-mannose by displacing ordered waters in the active site.[97] By additionally altering the conformation of an Asn side chain, *Hs*GALE can bind and epimerize the larger substrate UDP-*N*-acetyl-galactosamine.[98]

The GALE catalytic cycle (Figure 1.9 A) has been elucidated through a combination of structural and biochemical approaches. A comparison of the unliganded *Hs*GALE:NAD^+ resting state and the *Ec*GALE:NAD^+:UDP complex indicates that nucleotide alone is sufficient to close the *Ec*GALE NSBD onto the catalytic domain.[86] The resulting rotation brings the substrate diphosphate to within 6.6 Å of the nicotinamide of NAD^+ (Figure 1.12 A).[92] Despite the distance between the nucleotide and NAD^+, ^{31}P-NMR studies indicate that UMP binding changes the microenvironment surrounding the *Ec*GALE cofactor.[99] UDP also induces a long-range NMR effect that appears to distort the cofactor and polarize the π-electrons toward the nicotinamide ring N-1.[100] Consequently, nucleotide binding increases the positive charge at C-4 and is estimated to enhance NAD^+ reactivity toward native substrates and nonspecific reducing agents by 3 orders of magnitude.[100] Increased reactivity appears to be caused by the catalytic Lys, which bears a positive charge on the ε-ammonium group[101] and is positioned within 5.3 Å of the nicotinamide ring N-1.[85] Mutation of the Lys to Ala or Met does not affect substrate binding, but instead decreases GALE activity by greater than 3 orders of magnitude, abolishes UMP-dependent reduction by sugars and negates the effect of UMP binding on nonspecific reducing agents.[101] Subsequent analyses of Tyr mutations and ^{13}C-NMR chemical shifts indicate the



Adapted with permission from Liu, Y., Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Ruzicka, F. J., Holden, H. M., and Frey, P. A. Mechanistic roles of tyrosine 149 and serine 124 in UDP-galactose 4-epimerase from Escherichia coli, *Biochemistry 36*, 10675-10684. Copyright (1997) American Chemical Society.

Figure 1.12. UDP binding activates NAD^+ via a long-range, unseen conformational change. (A) Overlay of the *Hs*GALE resting state (yellow, PDB entry 1EK5) and *Ec*GALE:UDP:NADH abortive complex (cyan, 1NAI) showing that rotation of the NSBD (curved arrow) brings UDP no closer than 6.6 Å (black dashes) to the catalytic domain NAD^+ and does not facilitate direct interactions with the nicotinamide ring. The catalytic YxxxK pairs (given *Ec*GALE numbering) are represented as sticks to show UDP binding has little effect on the Tyr and Lys side chain conformations. (B) The current model of NAD^+ activation invokes a conformational change that is not present in current GALE crystal structures to explain biochemical data. In the resting state, the Y149 phenolate anion shields the N-1 of NAD^+ from electrostatic repulsion with the K153 side chain. UDP binding causes Y149 to change conformation and allows K153 to perturb the electron distribution across the nicotinamide ring. As a result, positive charge from the nicotinamide N-1 shifts toward C-4 and increase the reactivity of NAD^+ as a hydride acceptor.

conserved active site Tyr also participates in the nucleotide induced conformational change.[102; 103] In the current model of NAD^+ activation (Figure 1.12 B), the resting state Tyr phenolate anion shields the positive charge of Lys from the positive charge of the nicotinamide N-1. Nucleotide-induced conformational changes increase the Tyr-Lys and/or the Tyr-N-1 distance(s), restoring electrostatic repulsion between the Lys guanidinium and the nicotinamide N-1 of NAD^+. Repulsive forces distort the cofactor, transfer positive charge from the nicotinamide N-1 to C-4 and increase the reactivity of NAD^+ as a hydride acceptor.[103] Large Tyr-Lys motions invoked in the current model are not present in the available crystal structures of *Ec*GALE abortive complexes.

According to the extended SDR mechanism, catalysis begins with alignment and deprotonation of the substrate sugar C-4 hydroxyl by an active site base. Analysis of the *Hs*GALE:NADH:UDG abortive complex indicates that rotation of the NSBD aligns the substrate sugar C-4 with the cofactor C-4 (Figure 1.13).[86] In place of a traditional catalytic base, such as Asp, the conserved SDR Tyr is positioned at the junction of substrate C-4 and cofactor C-4.[86] Partial closure of the NSBD brings the C-3 hydroxyl of the substrate sugar near the catalytic Tyr but does not appear to shift the tyrosyl towards the substrate C-4. Only a larger conformation closing of the NSBD observed in the *Hs*GALE:UDG abortive complex places the sugar C-4 hydroxyl within direct hydrogen bonding distance of the conserved Tyr.[86] Tyr is not a traditional choice for the GALE general acid-base catalyst since the accepted tyrosyl pK_a (10.1) is higher than the Bronsted base pK_a originally calculated from *Ec*GALE kinetics (6.1).[104] However, the pK_a of the conserved Tyr in *Ec*GALE was later shown to be 4 pH units lower than free Tyr and is consistent with the kinetically determined pK_a (6.08 vs 6.1).[102] The positive electrostatic environment created by the catalytic Lys and charged cofactor is postulated to stabilize the Tyr

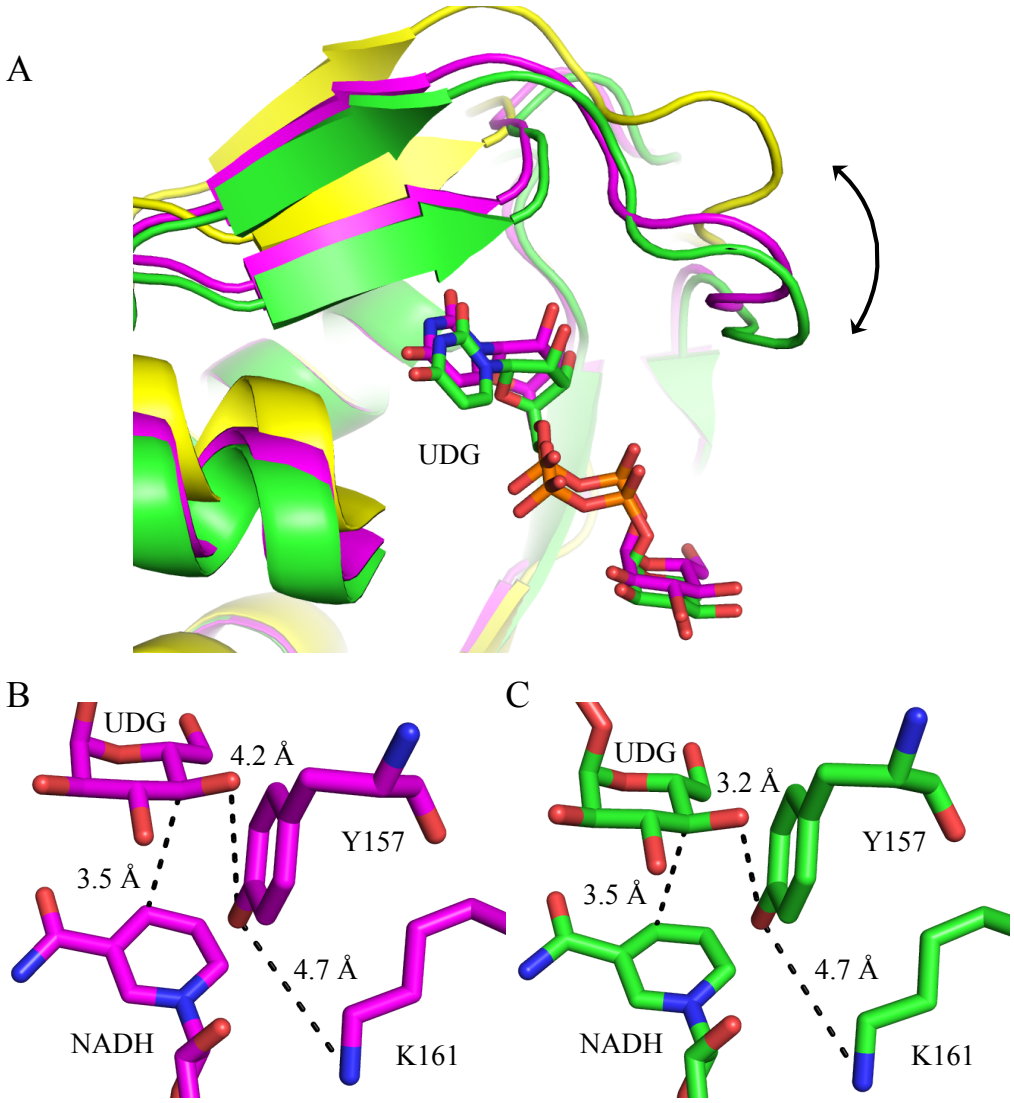


Figure 1.13. NSBD rotation positions the GALE substrate for catalysis. (A) Overlay of the *Hs*GALE resting state (yellow, PDB entry 1EK5) with chain A (magenta) and chain B (green) of the *Hs*GALE:NADH:UDG abortive complex (1EK6). The NSBD is cut away to show that chain B is more closed than chain A and positions UDG further into the active site. (B) Partial rotation of the NSBD in chain A aligns the C-4 of UDG with the C-4 of NADH (3.5 Å) but (C) full rotation of chain B is required to orient the C-4 hydroxyl of UDG for direct interaction with the catalytic Tyr 157 (3.2 Å). The pK_a of Tyr 157 is proposed to decrease by proximity to the positive charges on the oxidized nicotinamide N-1 and catalytic Lys.

phenolate, thereby lowering the pK_a. The catalytic Ser, as well as its conservative substitute Thr, also appears to stabilize the phenolate anion.[102]

After the catalytic Tyr nonstereospecifically abstracts a proton from the C-4 hydroxyl of UGal or UDG, a hydride is stereospecifically transferred to the nicotinamide C-4 of NAD^+ to form NADH and a UDP-4-keto-hexose intermediate.[101] *Ec*GALE does not display a significant deuterium kinetic isotope effect during epimerization at physiological pH, indicating hydride transfer is not the rate-limiting step in catalysis.[105] Instead, the rate-limiting catalytic step appears to be a conformational change, such as the aforementioned conformational activation of NAD^+ or rotation of the UDP-4-keto-hexose within the active site.[105] A comparison of *Ec*GALE:NADH:UGal and *Ec*GALE:NADH:UDG abortive complexes indicates a level of active site plasticity that allows the UDP-4-keto-hexose intermediate to change conformations without dissociating from the enzyme.[106] Substrate rotation occurs about two dihedral angles spanning the β-phosphate and glucosyl linkage. An additional translation of the substrate diphosphate backbone enables the catalytic Tyr and Ser residues to form analogous interactions with both UGal and UDG.[106] The identification of similar structural contacts supports the use of core catalytic residues during the nonstereospecific dehydrogenation of substrate C-4. The second hydride transfer regenerates NAD^+ and (re)forms UDG or UGal, depending upon the orientation of the glucosyl moiety within the active site cleft.[106]

The GALE catalytic cycle is believed to occur in the context of a dimeric quaternary structure for both mono- and bifunctional enzymes (Figure 1.10).[81] In addition to an active dimer, *Ec*GALE is reported to form a monomeric species when diluted. The *Ec*GALE monomer retains 90% of dimer activity and multiple physico-chemical properties, such as pH stability, secondary structure, and rate of inactivation by acid or heat.[107] On the other hand, the bifunctional *K.*

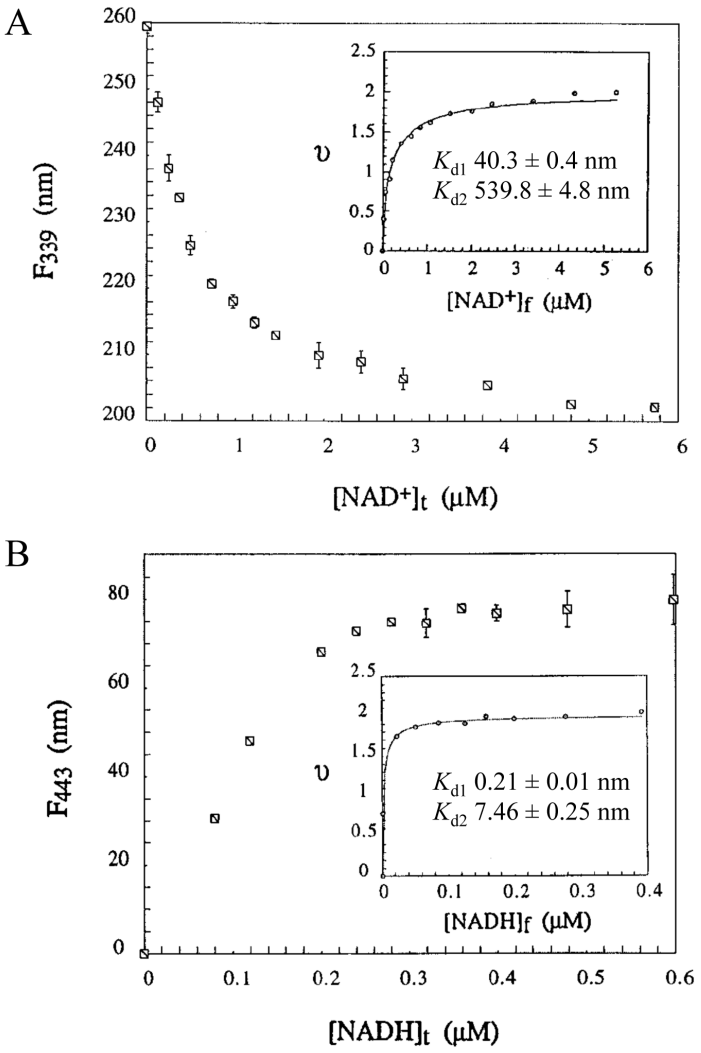
marxianus Gal10p (*Km*Gal10p) appears to be an obligate dimer[108] capable of forming a tetrameric species in the presence of cations and protein concentrations >1 mg/ml.[109] There is no available crystal structure of the Gal10p enzyme from *K. marxianus* and structural analysis of the homologous *S. cerevisiae* Gal10p:ligand complex indicates only an SDR dimer interface (Figure 1.10).[87] It is currently unclear which protein surface might stabilize the tetramer reported in *K. marxianus* and whether the higher order oligomer serves a biological role.

IX. Nucleotide sugar 4,6-dehydratases utilize two sets of catalytic residues to activate substrate for syn elimination of water.

Crystal structures of dTDP-glucose 4,6-dehydratase (RmlB; EC 4.2.1.46), CDP-glucose 4,6-dehydratase (E_{od}; EC 4.2.1.45) and GDP-mannose 4,6-dehydratase (GMD; EC 4.2.1.47) confirm the presence of SDR sequence motifs and tertiary structures.[110; 111; 112] Variations in the cofactor binding site (discussed previously for SDR enzymes) reflect a preference for either NAD(H) in RmlB[110] and E_{od}[111] or NADP(H) in GMD.[112] Additional modifications in the SDR GXXGXXG motif contribute to the binding mode for NAD(P)(H). In the GALE and RmlB motifs (Figure 1.7), 'X' represents an amino acid with a small and/or uncharged side chain (e.g. GGSGYIG vs GGAGFIG in *Ec*GALE[113] and *E. coli* RmlB (*Ec*RmlB),[114] respectively). GALE and RmlB bind NAD^+ tightly as a cofactor,[81; 115] while E_{od} from *Yersinia pseudotuberculosis* (*Yp*E_{od}) reversibly binds NAD(H) cosubstrate using an altered GHTGFKG motif (Figure 1.7).[116] Mutating the charged amino acids from the *Yp*E_{od} motif to the structural equivalents in *Ec*RmlB (i.e. HΔG or KΔI) results in decreased K_m and K_d values for NAD^+. In addition, the HΔG mutant displays a 3-fold increase in k_{cat}/K_m, suggesting the first residues of the GXXGXXG motif may contribute significantly to NAD(H) positioning during catalysis.[116] Interestingly, fluorescence

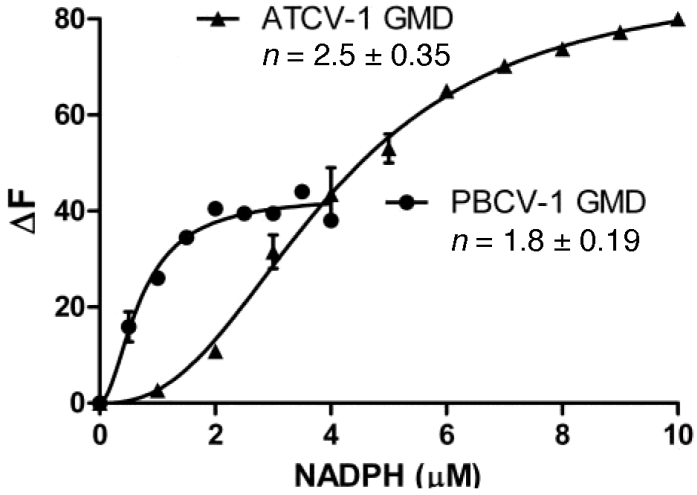
studies indicate that NAD^+ and NADH binding are anticooperative in *Yp*E_{od} (Figure 1.14). Either of the HΔG or KΔI mutations decreases the anticooperativity displayed during NAD(H) binding, further supporting a cooperative role for the modified GXXGXXG motif. NAD(H) binding does not appear to be the rate limiting step in the CDP-glucose dehydration, however, since *Yp*E_{od} shows classical Michaelis-Menten saturation kinetics with nucleotide sugar and NAD^+ substrates.[116] Similar to E_{od}, GMD enzymes encoded by the chlorella viruses PBCV-1 (open reading frame A118R of GenBank™ accession number NC000852)[117] and ATCV-1 (open reading frame Z804L of GenBank™ accession number EF101928)[118] reversibly bind NADPH using modified GATGQDG and GITGQDG motifs, respectively (Figure 1.7).[118] Although in contrast to *Yp*E_{od}, chlorella virus GMD enzymes show positive cooperativity with respect to NADPH binding (Figure 1.15).[118] Detailed kinetic and structural studies will be necessary to determine whether basic or acidic amino acids in the Gly-rich motif contribute directly to negative or positive cooperativity, respectively, during NAD(P)(H) binding.

Since NAD(P)(H)-dependent redox chemistry occurs at the C-4 and C-6 of substrate sugars, differences in C-2 chirality (i.e. mannose is the C-2 epimer of glucose) and nucleoside moiety (e.g. deoxythymidine vs guanosine) do not preclude the generalization of a single mechanism for the NDP-sugar 4,6-dehydratases. Similar to the GALE mechanism, the nucleotide portion of substrate is sufficient to cause rotation of the NSBD toward the catalytic domain in RmlB from *Salmonella enterica* serovar Typhimurium (*Se*RmlB) and *Streptococcus suis* sertotype 2 (*Ss*RmlB) (Figure 1.16 A).[119] A comparison of the *Se*RmlB:NAD^+ resting state with *Se*RmlB:NAD^+:dTDP and *Se*RmlB:NAD^+:dTDP-glucose complexes indicates two additional conformational changes close the active site upon ligand binding.[119] In the resting state, an unstructured loop at the carboxy-terminus of the first extended α-helix is oriented across the



Adapted with permission from He, X., Thorson, J.S., and Liu, H.W. Probing the coenzyme and substrate binding events of CDP-D-glucose 4,6-dehydratase: mechanistic implications, *Biochemistry 35*, 4721-4731. Copyright (1996) American Chemical Society.

Figure 1.14. NAD^+ and NADH binding are anticooperative in *Yersinia pseudotuberculosis* E_{od}. (A) E_{od} tryptophan fluorescence emission at 339 nm is quenched by increasing concentrations of NAD^+. The inset displays the plot of $v = n(I - I_a)/(I_h - I_a)$, where n represents the number of coenzyme binding sites; I_a ,the fluorescence of the apoenzyme; and I_h , the fluorescence of the holoenzyme with its n coenzyme binding sites saturated with NAD^+. The free NAD^+ concentration was estimated using the formula $[NAD^+]_f = [NAD^+]_t - n[E]_t$ in which $[NAD^+]_t$ and $[E]_t$ represent the total NAD^+ and enzyme concentration, respectively. A nonlinear least-squares procedure was used to fit v and $[NAD^+]_f$ to the Adair-Klotz equation[120] to determine the dissociation constants K_{d1} and K_{d2}. Alternatively, fitting the inset data to the Hill equation[58] gives a Hill coefficient $n = 0.6$, consistent with negative cooperative binding (i.e. $n < 1$). (B) NADH fluorescence emission at 443 nm is enhanced when bound to E_{od}. The inset data were treated as in (A). Fitting the inset data to the Hill equation gives a Hill coefficient $n = 0.4$.



Modified with permission from Fruscione, F., Sturla, L., Duncan, G., Van Etten, J.L., Valbuzzi, P., De Flora, A., Di Zanni, E., and Tonetti, M. Differential role of $NADP^+$ and NADPH in the activity and structure of GDP-D-mannose 4,6-dehydratase from two chlorella viruses. *J Biol Chem*. 1977; 283: 184-193.

Figure 1.15. NADPH binding is cooperative in GMD enzymes from the chlorella viruses ATCV-1 (▲) and PBCV-1 (●). The change in NADPH fluorescence emission intensity at 350 nm upon binding to GMD (ΔF) is plotted as a function of increasing NADPH concentration. ΔF was calculated by subtracting the fluorescence intensity of NADPH solutions from the intensity of samples containing protein and an equivalent concentration of NADPH. Each data set is best fit by non-linear regression to a one binding site Hill equation and yields a Hill coefficient *n* that is consistent with positive cooperativity (i.e. $n > 1$).

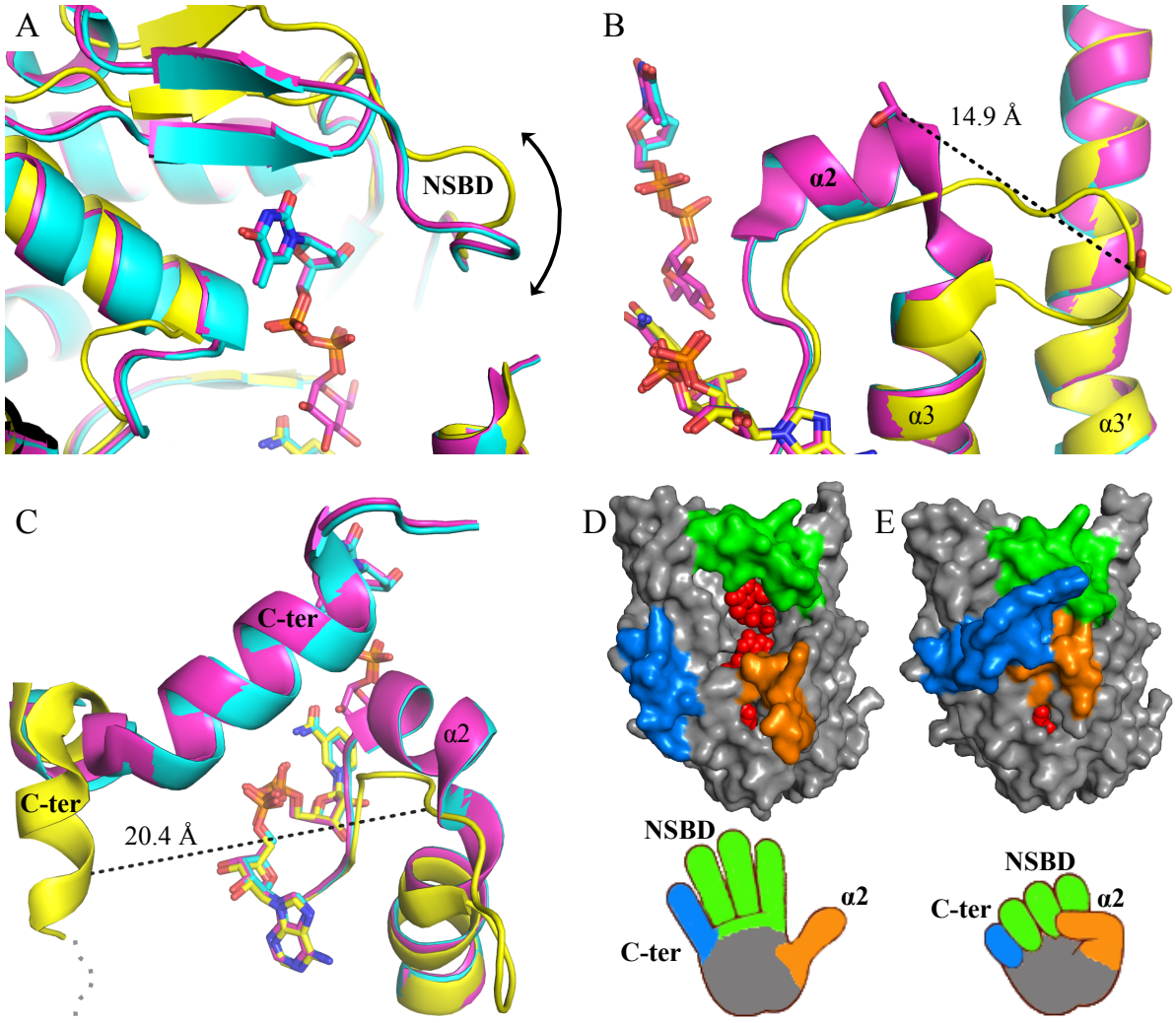


Figure 1.16. RmlB from *Salmonella enterica* serovar Typhimurium (*Se*RmlB) closes like a hand around substrate. (A) Overlay of the *Se*RmlB:NAD^+ resting state (yellow, PDB entry 1G1A) with *Se*RmlB:NAD^+:dTDP (cyan, 1KEW) and *Se*RmlB:NAD^+:dTDP-glucose (magenta, 1KEU) complexes. Monomers are oriented to show that dTDP (cyan sticks) and dTDP-glucose (magenta sticks) cause the same degree of NSBD rotation (curved arrow) relative to the resting state. (B) In the resting state, Thr92 (sidechain sticks) is part of a loop stretching from the first extended helix (α3, numbered as hUXS in Fig 1.8) of one monomer, across the dimer interface and toward the analogous helix (α3') of the second monomer. Ligand binding causes the Cα of Thr92 to translate 14.9 Å (dashed line) and form part of a helix (α2) packed over the active site. (C) The extreme C-terminus of *Se*RmlB forms an α-helix that is partially disordered (grey dots) in the resting state and over 20 Å away from the α2 loop. Ligand binding causes the helix to rotate ~130° toward the newly formed α2 helix and results in the ordering of the terminal nine amino acids over the active site. (D) Surface rendering of the *Se*RmlB resting state (grey) with NAD^+ cofactor and the dTDP (red spheres) of 1KEW modeled into the active site. The structural elements involved in NSBD rotation (green), α2 helix formation (orange) and C-terminal movement (blue) form a hand around the active site cleft. (E) Surface rendering of the *Se*RmlB:NAD:dTDP complex showing the structural elements (colored as in D) pack over the active site in the presence of substrate and resemble a closed fist.

dimer interface. Ligand binding causes the loop to assume α-helical properties and translates mainchain atoms away from the dimer interface by as much as 15 Å to pack against the active site (Figure 1.16 B).[119] The motion resembles the closing of a hand, with the catalytic domain acting as the stationary palm, the NSBD representing three fingers and the loop forming an opposable 'thumb' helix. Expanding upon the hand analogy, the second conformational change seen in *Se*RmlB:ligand complexes resembles the pinching of the little finger to the thumb. In the resting state, the extreme carboxy-terminal residues form an α-helix that is partially disordered and oriented approximately 20 Å away from the thumb helix loop. Substrate binding causes the terminal α-helix to rotate more than 130° and pack against the thumb helix (Figure 1.16 C),[119] thereby closing the hand (Figure 1.16 D,E).

In GALE, nucleotide binding is believed to activate the NAD^+ cofactor via long-range conformational changes in the catalytic YxxxK pair (discussed above in detail).[103] In *Se*RmlB, ligands cause the catalytic Tyr to move away from the conserved Lys and toward the position of substrate C-4 (Figure 1.17 A), consistent with the GALE model for NAD^+ activation (Figure 1.12 B). Despite structural evidence that the RmlB Tyr moves ~1 Å further than the GALE Tyr (Figure 1.17 B),[86; 119] RmlB does not appear to activate NAD^+ towards the nonspecific reductant dimethylamine-borane complex.[121] While the RmlB YxxxK may not serve all the same roles as the GALE catalytic pair,[121] the Tyr is generally accepted to deprotonate the substrate sugar C-4 hydroxyl in accordance with the SDR mechanism.[110] Concomitantly, hydride is transferred from the C-4 of substrate to the nicotinamide ring of NAD^+ to form an NDP-4-keto-hexose and NADH (Figure 1.9 B).

According to the extended SDR mechanism, hydride transfer is followed by repositioning of the 4-keto-sugar within the active site. In RmlB, repositioning dTDP-4-keto-glucose involves

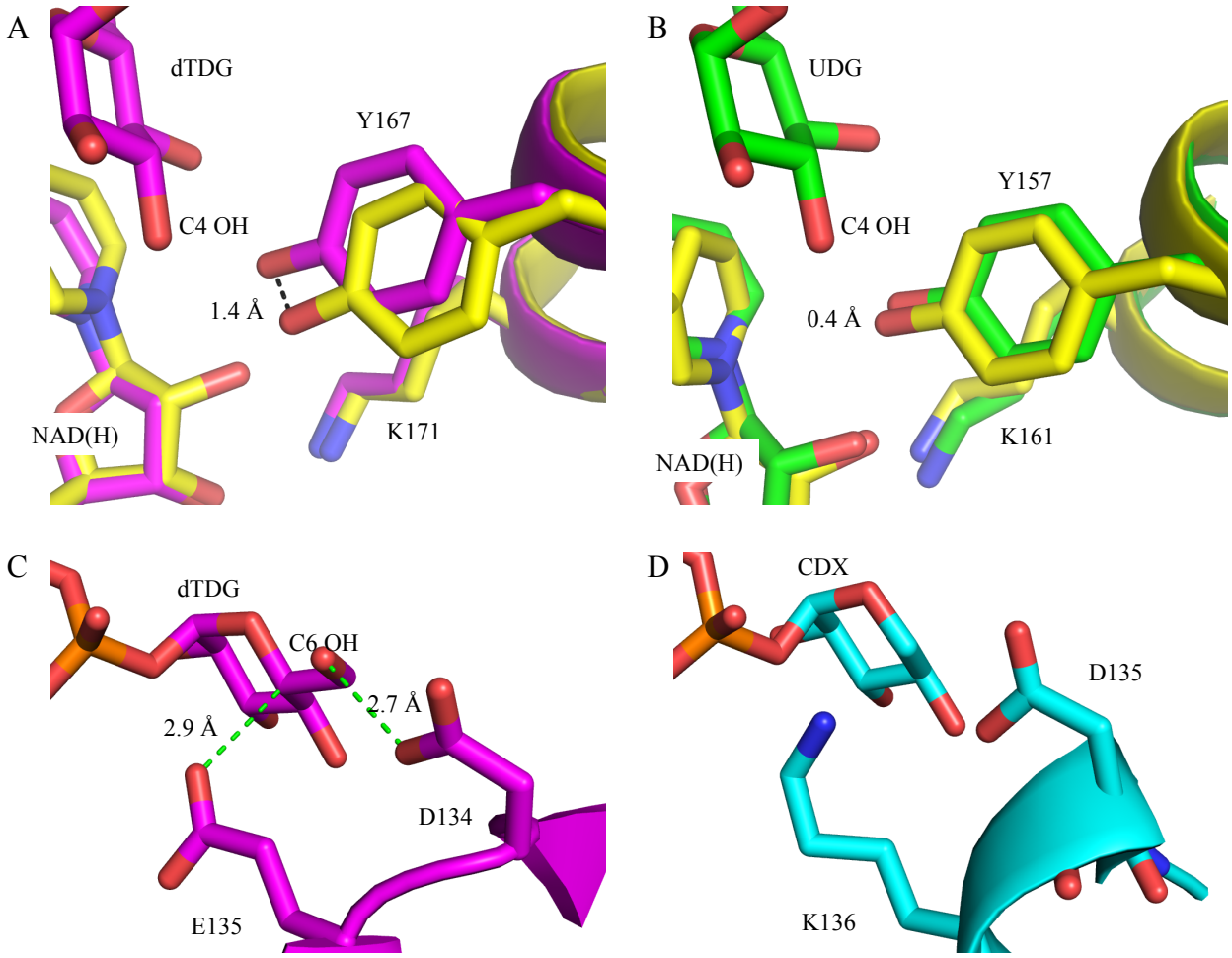


Figure 1.17. Ligand binding positions two sets of catalytic residues in NDP-sugar 4,6-dehydratases. (A) Superposition of the *Se*RmlB:NAD^+ resting state (yellow, PDB entry 1G1A) and *Se*RmlB:NADH:dTDP-glucose complex (magenta, 1KEU) showing substrate dTDP-glucose (dTDG) causes the catalytic Tyr 167 phenolate to move 1.4 Å, decreasing the distance to substrate (C4 OH) and increasing the distance from the positive charge on catalytic Lys 171. (B) Superposition of the *Hs*GALE:NAD^+ resting state (yellow, 1EK5) and *Hs*GALE:NADH:UDG complex (green, 1EK6) showing substrate UDG causes minimal movement of the catalytic Tyr 157 hydroxyl (0.4 Å) compared to *Se*RmlB (in A). (C) Cut away of the *Se*RmlB:NADH:dTDG complex (colored as in A) showing the Asp134-Glu135 catalytic pair is in position to interact with the C6 hydroxyl (2.7 Å) and C5 proton (2.9 Å), respectively. (D) Active site cut away of the *Salmonella enterica* serovar Typhi E_{od}:CDP-xylose complex (cyan, 1WVG). The Asp135-Lys136 pair orients the substrate mimic CDP-xylose (CDX) and occupies the equivalent position to the Asp-Glu pair from *Se*RmlB (in C).

enzyme-catalyzed dehydration across the sugar C-5−C-6 bond. Elimination of water is proposed to occur *syn* relative to the C-5 proton and C-6 hydroxyl,[122] and necessitates control of C-5−C-6 rotational freedom within the active site.[121] In RmlB, an active site Asp-Glu pair is positioned to serve as catalytic acid and base for C-6(OH) protonation and C-5 deprotonation, respectively (Figure 1.17 C).[123] Mechanistic evidence indicates that in addition to serving as a catalytic acid, the *Escherichia coli* RmlB (*Ec*RmlB) Asp restricts rotational freedom about the glucosyl C-5−C-6 bond.[124] Interestingly, kinetic studies suggest the *Ec*RmlB Glu can partially replace the catalytic Tyr as base in the dehydrogenation of substrate C-4(OH).[123] In this model, proton exchange may occur between the adjacent Asp-Glu side chains to ensure the correct protonation state for the dehydration reaction. The active site Asp-Glu pair is replaced with an Asp-Lys couple in E_{od} (Figure 1.17 D) [111; 125] highlighting the importance of the conserved aspartate−C-6(OH) interaction during the dehydration reaction and subsequent NDP-4-keto-hexose-5,6-ene formation.[124]

An analysis of the RmlB structure finds it unlikely that both the C-4 and C-6 positions of dTDP-4-keto-glucose-5,6-ene simultaneously adopt productive trajectories for hydride transfer from NADH.[110] Based on homology modeling with the RmlB:UDG complex, dTDP-4-keto-glucose-5,6-ene rotation about the β-phosphate-anameric oxygen bond may cause the glucosyl C-6 to move ~2.5 Å into an optimal position for hydride acceptance.[119] Allard et al., further propose that van der Waals repulsion between the newly eliminated water molecule and substrate could drive a ~2.2 Å movement of the glucosyl C-6 towards a productive hydride trajectory.[119] Supporting this model, the structure of *Ss*RmlB bound to dTDP-xylose (which lacks a C-6) contains an ordered water molecule that occupies the position of a substrate C-6 hydroxyl and is anticipated to sterically clash with substrate C-6.[119] After the C-6 of NDP-4-keto-hexose-5,6-ene

is properly aligned for hydride transfer from the C-4 of NADH, substrate is reduced to form the final product NDP-4-keto-6-deoxy-hexose.

While the 4,6-dehydratase catalytic mechanism is common to RmlB, E_{od} and GMD, the enzyme quaternary state appears to vary. RmlB enzymes from *S. enterica*,[126] *S. suis*,[119] *Streptomyces venezuelae*[127] and *E. coli*[114] form a dimer in crystal structures. Unlike other RmlB crystal forms, the *Se*RmlB:NAD^+ asymmetric unit contains two dimers (Figure 1.18 A). However, structural analysis indicates the dimer-dimer interactions are limited to crystal contacts and do not appear to stabilize a higher order (>2n) oligomer.[110] *Yp*E_{od}, on the other hand, has been reported in dimeric[128] and tetrameric states by gel filtration (although experimental data is not shown) and crystalizes in an NAD^+ bound tetramer with near 222 symmetry (Figure 1.18 B).[111] E_{od} from *S. enterica* serovar Typhi (*Se*E_{od}) is also reported to form a tetramer by sedimentation equilibrium analysis (although solution data is not presented) and crystalizes in a mixed CDP-xylose:NAD^+ complex with 222 tetramer symmetry.[125] The reported E_{od} tetramer arranges the parallel four helix bundles of each dimer on the outside of the complex, creating a central cavity in the quaternary structure (Figure 1.18 B).

The quaternary structure of GMD displays even greater variety. Gel filtration shows the GMD enzymes encoded by PBCV-1 and ATCV-1 are monomeric in the absence of cofactor, and form catalytically competent dimers driven by NADPH binding.[118] In contrast to the solution studies, PBCV-1 GMD has also been presented as a homotetramer formed between crystal symmetry mates.[129] A superimposable tetramer complex with near 222 symmetry is found in structures of GMD from *A. thaliana* (*At*GMD, a.k.a. MUR1, Figure 1.18 C),[130] *Pseudomonas aeruginosa*,[131] and *Aquifex aeolicus* (unpublished, PDB entries 2Z1M and 2Z95), as well as the extended SDR enzyme CDP-tyvelose 2-epimerase from *S. enterica* serovar Typhi.[132] Gel

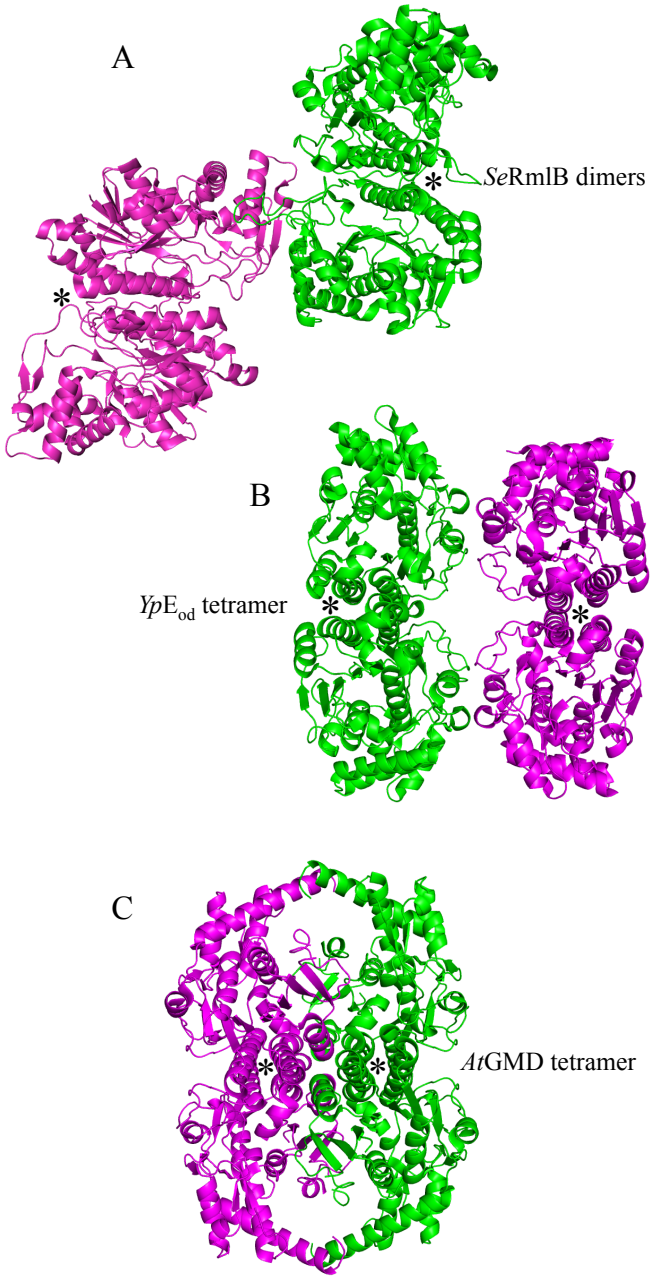


Figure 1.18. Quaternary structures formed by NDP-sugar 4,6-dehydratases. (A) Contents of the *Se*RmlB asymmetric unit (PBD entry 1G1A) showing four polypeptides organized into two dimers (magenta or green) with minimal crystal contacts and perpendicular four helix bundles (*). (B) Contents of the *Yp*E_{od} asymmetric unit (1RKX) showing four polypeptides with near-222 noncrystallographic symmetry assemble as a tetramer. Dimers (colored as in A) are oriented to show the dimer four helix bundles are parallel and positioned to the outside of the tetramer interface. (C) Crystal structure of the *At*GMD tetramer (1N7H) formed between dimer crystal symmetry mates (colored as in A) showing the dimer four helix bundles are parallel (as in B) but positioned to the outside of the tetramer interface. The tetramer is staggered with (B) to illustrate that distinct tetramers are observed from packing on opposite faces of a superimposable dimer (green in B and C).

filtration results are reported for the tetrameric state of *A. thaliana* GMD[130] and the dimeric states of human[133] and *E. coli* GMD (*Ec*GMD)[112] (although supporting experimental data are not shown in these reports). Contrasting size exclusion data show that *Ec*GMD forms a hexamer in solution.[134] A crystallographic analysis of *Ec*GMD characterizes the SDR dimer structure[112] rather than the hexamer formed by crystal symmetry.

While the GMD, RmlB and E_{od} monomers are similar in fold and serve as a scaffold for higher order oligomers (Figure 1.18), the resulting quaternary structures are variable. In contrast to E_{od}, the four helix bundles from each GMD dimer are arranged at the center of the oligomer building interface. Interestingly, the GMD tetramer results from dimer packing on the opposite face of an E_{od} dimer (Figure 1.18). As a consequence, the GMD tetramer interface is more extensive than the corresponding E_{od} interface (e.g. ~6200 $Å^2$ total buried in *At*GMD vs ~2800 $Å^2$ in *Se*E_{od}).[125; 130] Beyond the cooperativity discussed above for cofactor binding in GMD[118] and E_{od}[116] dimers, no additional functions have been ascribed to higher order self-assemblies of NDP-sugar 4,6-dehydratases. However, an additional layer of oligomeric complexity exists if one considers the similar overall fold of extended SDR dimers and their potential to form hetero-complexes. In support of a biological role for hetero-SDR assemblies, *At*GMD immuno-precipitates with the extended SDR enzyme GDP-mannose 3,5-epimerase (EC 5.1.3.18) from *A. thaliana* (*At*GME) or *E. coli* (*Ec*GME).[135; 136] Further, yeast cells coexpressing *At*GMD and either GME show increased *At*GMD protein levels and higher dehydratase activity, compared to cells expressing *At*GMD alone.[135] Western blot densitometry indicates the stoichiometry of precipitated *At*GMD:*At*GME and *At*GMD:*Ec*GME complexes is 3:1 and 1:1, respectively.[135] Given the propensity for GMD to form a sandwich of SDR dimers in crystal structures,[130] it is tempting to speculate that an *At*GME monomer, or an *Ec*GME dimer, incorporates into an

*At*GMD heterotetramer and stabilizes dehydratase activity. Currently there is no available crystal structure of an extended SDR hetero-complex to support this model.

X. UDP-glucuronic acid decarboxylases are underrepresented among structurally characterized extended SDR enzymes.

When we began to study hUXS, *E. coli* ArnA was the only UGA decarboxylase that had been deposited in the PDB and structurally analyzed in the literature. ArnA has been characterized in the context of a full-length enzyme (residues 1-660 in PDB entry 1Z7E) or as a functional C-terminal decarboxylase domain ($ArnA_{dc}$, residues 314-660 in PDB entries 1U9J and 2BLL) that is separable from the N-terminal formyltransferase domain (residues 1-304). During the preparation of this thesis, a structural description of hUXS in complex with NAD^+ and the substrate analog UDP was published (PDB entry 2B69).[44] Analyses of the ArnA structures reveal two ligand-induced conformational changes (Figure 1.19). The first conformational change occurs in a dynamic NSBD loop (ArnA residues 605-616) that is disordered and unmodeled in both structures of apo-$ArnA_{dc}$.[24; 26] A comparison with the ArnA:UGA:ATP complex (containing ATP as an NAD^+ mimic) indicates that substrate binding induces the dynamic NSBD loop to become better ordered and pack over the nucleotide portion of UGA (Figure 1.19 A).[25] The nucleotide alone appears sufficient to order the equivalent NSBD loop in the hUXS structure (data not shown). The second ArnA conformational change involves another NSBD loop (ArnA residues 500-509) that extends into the apo-ArnA catalytic domain and occludes the NAD^+ binding site. Upon ligand binding, the loop is displaced from the catalytic domain, assumes α-helical secondary structure and caps part of the active site cleft (Figure 1.19 A).[25] In the hUXS:UDP complex, the analogous loop adopts a similar capping conformation (data not

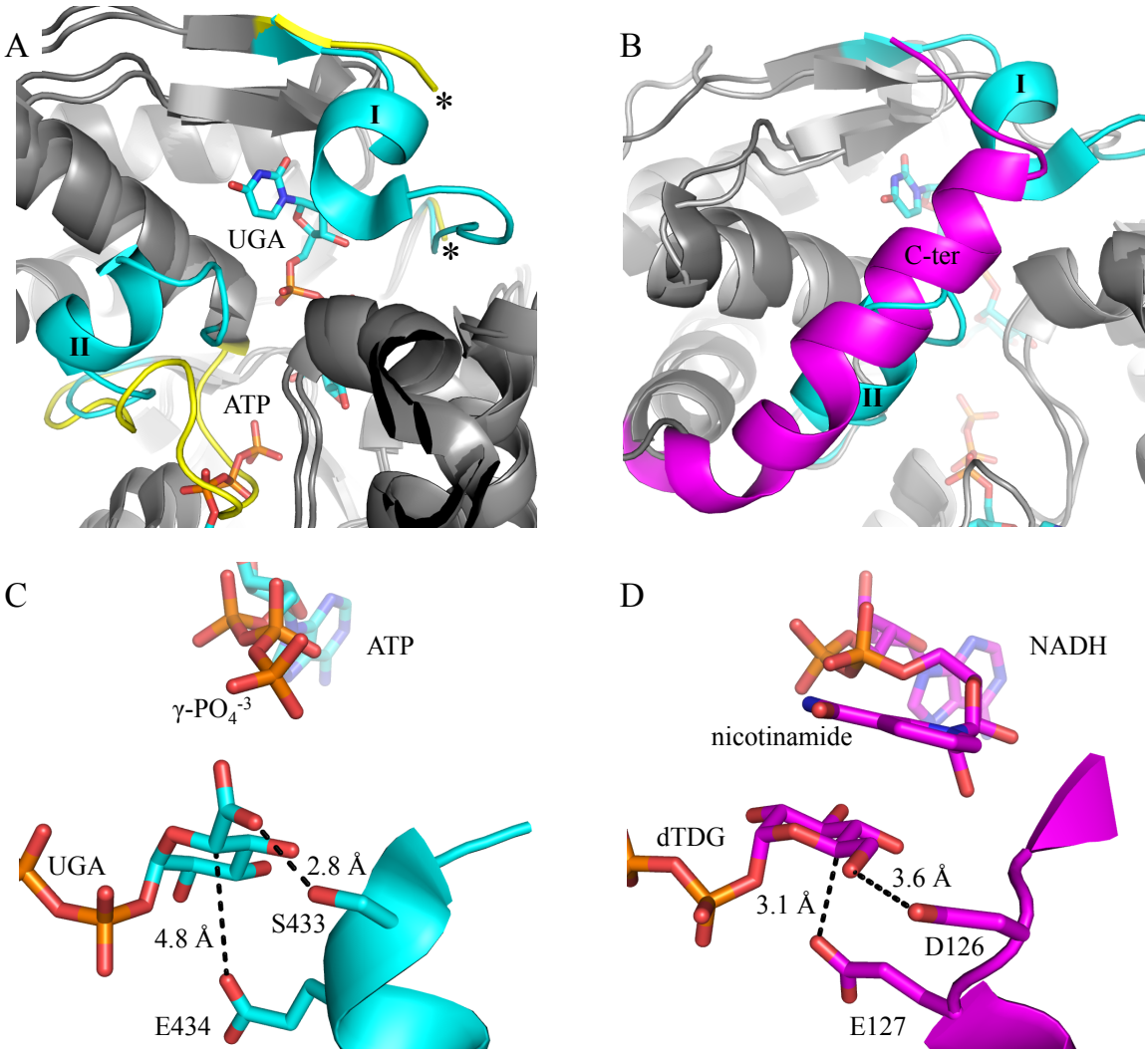


Figure 1.19. Ligand binding restructures the ArnA decarboxylase active site cleft. (A) Overlay of the $ArnA_{dc}$ resting state (yellow and grey, PDB entry 1U9J) and chain A of the $ArnA_{dc}$:UGA:ATP mixed substrate:analog complex (cyan and grey, 1Z7E). Enzymes are oriented to show that a disordered loop (*) reorders as a helix (I) upon UGA binding and an extended loop occupying the NAD^+ binding site is restructured (II) in order to accommodate the substrate analog, ATP. (B) Overlay of $ArnA_{dc}$:UGA:ATP (colored as in A) and the *Ss*RmlB:NAD^+:dTDP-glucose complex (magenta and grey, 1KER) showing the ordered ArnA helices (I, II) serve to close the active site similar to the RmlB C-terminal helix. (C) Close up of the $ArnA_{dc}$:UGA:ATP complex (cyan) showing the conserved Ser-Glu pair is positioned in the active site near the glucosyl C-6 carboxylate (2.8 Å) and C-4 (4.8 Å), respectively. (D) Close up of the *Ss*RmlB:NAD^+:dTDP-glucose complex (magenta) showing the positions equivalent to the $ArnA_{dc}$ Ser-Glu pair are occupied by the catalytic Asp-Glu pair of 4,6-dehydratases. The active site is superimposable with (C) to illustrate that the γ-phosphate of the mimic ATP is nonisoteric with the nicotinamide of authentic NADH. In the presence of NADH cofactor, the glucosyl moiety of dTDP-glucose (dTDG) adopts an alternate conformation compared to UGA, resulting in Asp126−C-6(OH) and Glu127−C-4 distances of 3.1 and 3.6 Å, respectively.

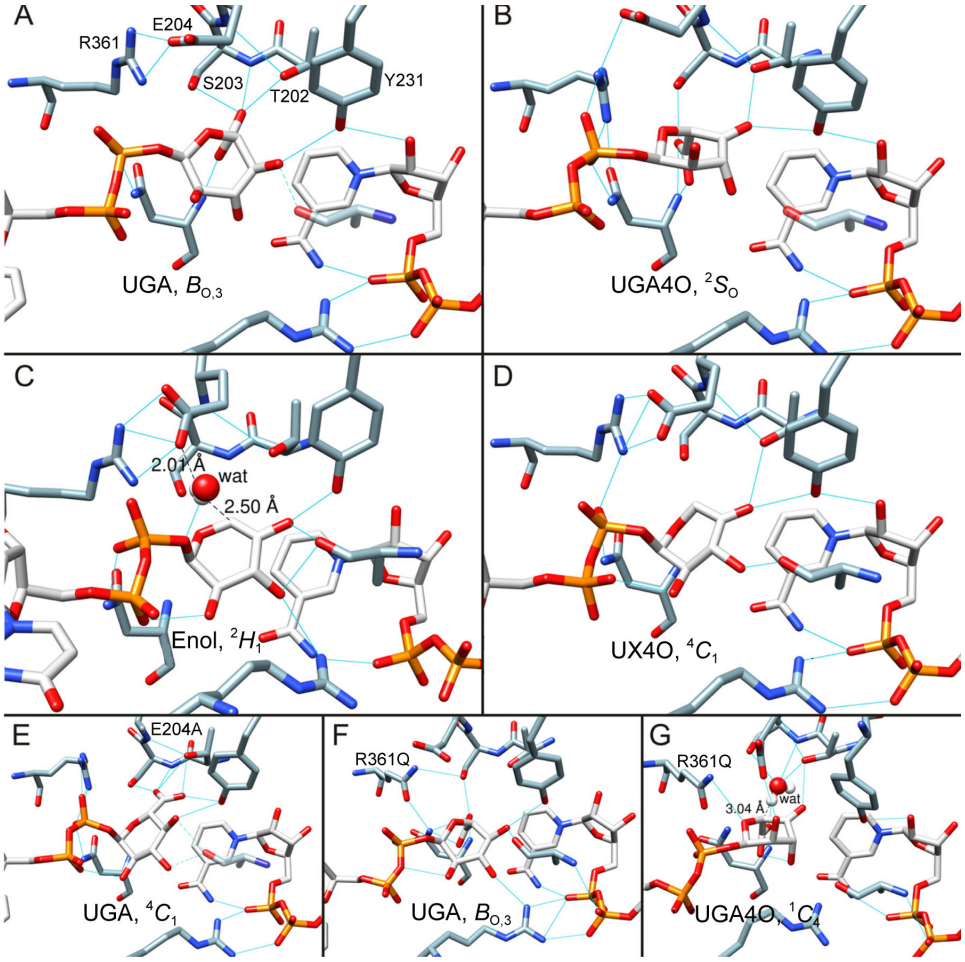
shown) although it is unclear from the ternary complexes whether it is in response to the cofactor and/or substrate. It should be noted that both flexible loops occur in variable regions of the extended SDR fold and are not strictly conserved across epimerase, dehydratase and decarboxylase primary structures. The capping interactions appear to be important for closing the active site, however, and in *Ss*RmlB, a 130° rotation of the extreme C-terminal helix (discussed previously) partially replaces both sets of loop interactions (Figure 1.19 B).

In accordance with the extended SDR mechanism, nucleotide sugar binding places the glucosyl C-4 hydroxyl within hydrogen bonding distance to the catalytic tyrosine. Similar to the dehydratases, UGA decarboxylases are additionally expected to make conserved interactions with the C-6 functional group of substrate. Analysis of the ArnA:UGA:ATP structure indicates that the UGA C-4 hydoxyl is positioned ~5 Å away from the catalytic tyrosine.[25] Nonproductive UGA binding is likely caused by interactions with the γ-phosphate of ATP, which is nonisosteric with the ribose and nicotinamide of authentic NADH (Figure 1.19 C,D).[26] In order to assess native ArnA:UGA interactions, Williams and coworkers[26] created a conceptual model by superimposing the substrates from a RmlB:NAD:dTDP-glucose complex[119] onto apo-ArnA. The modeling exercise suggests that proper alignment of UGA orients the C-6 carboxylate towards a Ser433-Glu434 pair in the active site (Figure 1.19 C,D). The Ser-Glu pair is conserved among UGA decarboxylases (Ser203-Glu204 in hUXS) but not across the extended SDR family,[24] consistent with a role for C-6, rather than C-4, stabilization.[26] The Ser-Glu pair also superimposes with the RmlB catalytic Asp-Glu pair shown previously to interact with glucosyl C-5 and C-6 positions (Figure 1.19 C,D).[26] Consistent with the modeled substrate orientation, single alanine mutations in the ArnA Ser-Glu pair cause ~3-fold changes in the K_m for UGA but

have no effect on the K_m for NAD^+. Based on the RmlB model, the Ser-Glu pair was proposed to serve a catalytic role.[26]

Similar to the RmlB model, docking simulations between UGA and hUXS show that authentic NAD^+ cofactor alters the substrate orientation found in the ArnA:UGA:ATP complex (Figure 1.20).[44] Additional refinement of the hUXS:UGA model using molecular dynamic simulations (MDS) indicates the conserved Ser203 and Glu204 residues do not function like the RmlB catalytic pair. Instead, Glu204 and Arg361 form a salt bridge that closes the active site and positions Ser203 for substrate binding (Figure 1.20 A). Interestingly, the glucuronyl moiety of UGA binds in a distorted $B_{O,3}$ *boat* conformation in simulations containing the salt bridge. *In silico* hUXS E204A and R361Q mutations prevent salt bridge formation and the associated Ser203 repositioning during MDS. As a consequence, the glucuronyl moiety can relaxe to a *chair* conformation that moves the C-4 hydroxyl away from the catalytic tyrosine (Figure 1.20 E,G).[44] Experimental evidence supports the MDS correlation between active site closure, sugar ring distortion and productive substrate alignment. A hUXS E204A mutant displays reduced UDX formation, consistent with a reduced ability to close the active site via the Glu204 and Arg361 loops.[44] The alanine substitution is not expected to sterically prevent the loop-loop interaction. In contrast, an ArnA E434Q mutant displays a loss of detectable activity,[26] as might be expected if the equivalent Glu434 and Arg619 loops are sterically unable to close the active site from solvent and distort the unreactive *chair* conformation.

Proper binding of substrate in the decarboxylase active site allows tyrosine mediated proton abstraction from the C-4 hydroxyl of UGA and concomitant C-4 hydride transfer to the nicotinamide ring of NAD^+. NADH formation (representing the oxidation of UGA) appears to be the rate-limiting step in hUXS catalysis.[44] The resulting UGA4O is a labile β-keto acid that has



Reproduced with permission from Eixelsberger, T., Sykora, S., Egger, S., Brunsteiner, M., Kavanagh, K.L., Oppermann, U., Brecker, L., and Nidetzky, B. Structure and mechanism of human UDP-xylose synthase: evidence for a promoting role of sugar ring distortion in a three-step catalytic conversion of UDP-glucuronic acid, *J Biol Chem*. 2012; 287: 31349-31358.

Figure 1.20. Trajectory of sugar ring conformation along the hUXS reaction coordinate determined by molecular dynamics simulations. (A) UGA binding induces the conserved E204 and R361 residues to form a salt bridge that contributes to closing the active site. UGA adopts a distorted $B_{O,3}$ *boat* conformation within hydrogen bonding distance (blue lines) to the catalytic Y231 residue. The C-6 carboxylate adopts an axial position that is supported by hydrogen bonds to the conserved S203 and catalytic T202 residues. (B) After hydride transfer to NAD^+, the labile UGA4O intermediate adopts a ^2S_O *skew-boat* conformation that alters the coordination of the C-6 carboxylate and C-4 ketone. (C) Decarboxylation produces an enol that adopts a $^2H_1$ *half-chair* conformation. The space created by the leaving carboxylate allows a water molecule (red and white spheres) to stereospecifically protonate the C-5 enolate. (D) The resulting UX4O is no longer coordinated through the carboxylate via S203 and relaxes to a $^4C_1$ *chair* conformation. (E) *In silico* E204A mutation breaks the conserved salt bridge and allows UGA to adopt a low energy $^4C_1$ *chair* conformation. (F) *In silico* R361Q mutation breaks the conserved salt bridge but induces UGA to adopt the distorted $B_{O,3}$ *boat* conformation. (G) *In silico* R321Q mutant binds UGA4O in the $^1C_4$ *chair* conformation. Disruption of the conserved salt bridge contributes to active site solvent accessibility.

not been observed in ArnA[23; 26] or hUXS reactions.[44] MDS of the hUXS catalytic cycle indicate the transient UGA4O adopts a $^{2}S_{O}$ *skew boat* conformation (Figure 1.20 B). Together, the *skew boat* conformation and conserved Ser are proposed to stabilize the UGA4O carboxylate in an axial position that is ideal for β-keto acid decarboxylation.[44; 137] After decarboxylation, the transient enol is predicted to adopt a $^{2}H_{1}$ *half-chair* conformation (Figure 1.20 C). Consistent with the reported stereospecific protonation of the enolate C-5,[20] the cavity vacated by the carboxylate can accommodate a water molecule coordinated by Glu204 and the α-phosphate of substrate. Without the conserved Ser-carboxylate interaction distorting the sugar ring, protonation of the enolate yields UX4O in the low energy $^{4}C_{1}$ *chair* conformation in MDS (Figure 1.20 D).[44]

The ArnA and hUXS mechanisms diverge following UX4O formation. ArnA is reported to release UX4O and NADH from the active site prior to returning a hydride to the glucosyl C-4 (Figure 1.9 D).[23] Similar to E_{od} and GME homologs that reversibly bind NAD(P)(H), ArnA contains a modified GVNGFIG motif with extended residue side chains at the second and third amino acid positions (Figure 1.7).[24] A comparison of the ArnA active site with *Ec*GALE reveals another structural feature that may explain the use of NAD^{+} as cosubstrate or cofactor, respectively. *Ec*GALE makes seven hydrogen bonds to the adenine base, ribose and α-phosphate of NAD^{+} via an eight amino acid loop that excludes the cofactor from solvent (Figure 1.21). In contrast, the analogous loop in ArnA is four residues shorter and unable to make extensive contacts with the NAD^{+} substrate.[24] The truncated ArnA loop exposes NAD^{+} to solvent and is similar to the structures found in NAD(P)(H)-exchanging members of the classical SDR subfamily (Figure 1.21). Recently, the equivalent loop in *Pyrobaculum calidifontis* GALE was mutated and truncated, resulting in reversible NAD binding.[138] hUXS contains an extended loop

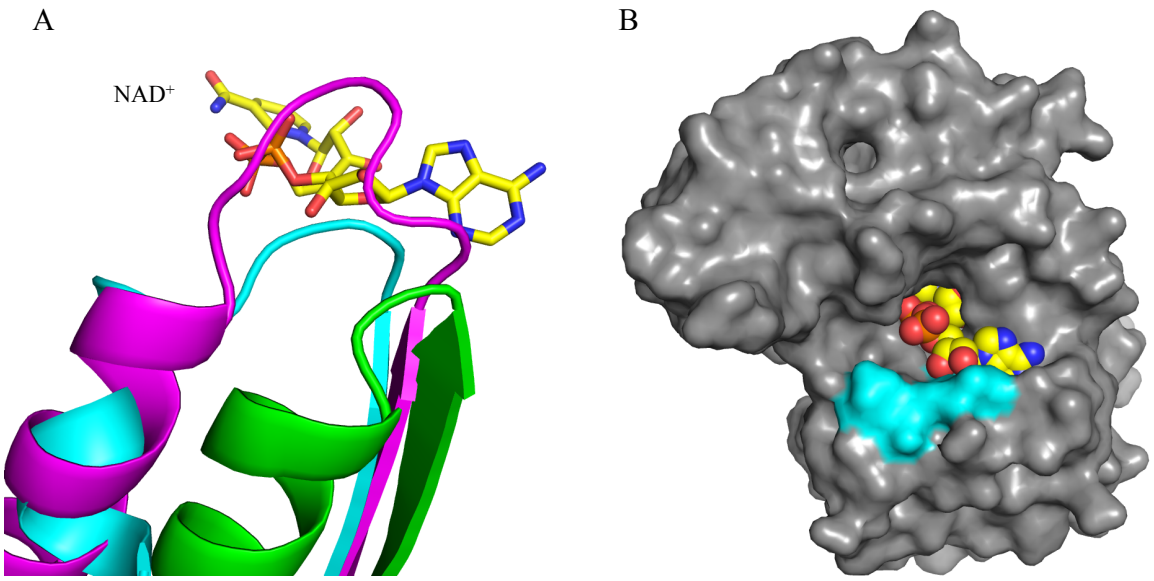


Figure 1.21. SDR enzymes contain a divergent loop at the NAD^+ binding site. (A) Overlay of the extended SDR enzymes *Ec*GALE (magenta, PDB entry 1XEL) and $ArnA_{dc}$ (cyan, 2BLL) with the classical SDR enzyme 3α/30β hydroxysteriod dehydrogenase (green, 2HSD). Enzymes are oriented to show that the ArnA loop is intermediate in length and ability to pack over NAD^+ (yellow sticks, 1XEL), compared to extended and classical SDR homologs. (B) Surface rendering of $ArnA_{dc}$ showing the ArnA loop (cyan) exposes the modeled NAD^+ (yellow spheres, 1XEL) to solvent.

that is similar to *Ec*GALE[24] and may account for the NAD^+ recycling observed during UDX biosynthesis (discussed in detail in Chapter 3).[7]

The decarboxylase quaternary structure also appears to diverge between ArnA and hUXS and has even been reported to vary between constructs of the same enzyme. Full length ArnA is reported to form a hexamer in the presence and absence of substrates (although gel filtration data are not shown).[25] A crystal structure of the full length ArnA:UGA:ATP complex reveals that a hexamer is built upon three SDR dimer interfaces formed by $ArnA_{dc}$ and less extensive interactions between the formyltransferase domains (Figure 1.22). Surprisingly, solution studies report (but do not provide data to the effect) that separately purified $ArnA_{dc}$ and formyl-transferase domains are monomeric during gel filtration[24; 139] and higher order oligomers only occur in the context of full length ArnA.[25; 26] However, $ArnA_{dc}$ crystallizes as a typical SDR dimer in apo-structures,[26] indicating the potential for oligomer formation in solution. A superposition of ArnA, *Ec*GALE, and *Ss*RmlB reveals that ArnA adopts one of two possible dimer orientations supported by the 'canonical' SDR dimerization surface (Figure 1.23). *Ss*RmlB monomers pack along a four helix bundle constructed by extended helices that are nearly parallel. In contrast, ArnA and GALE monomers adopt an orientation that rotates the extended helices of one monomer by as much as 40° relative to the opposite monomer. Although the dimer rotation appeared to be specific to extended SDR subfamilies, divergent orientations were recently reported within GALE enzymes. GALE from hyperthermophilic archaea, such as *Archaeoglobus fulgidus* and *Pc*GALE, display a ~37° rotation relative to the dimers found in *Ec*GALE, *Hs*GALE and *Trypanosoma brucei* GALE crystal structures.[138] The orientations of the hyperthermophilic GALE dimers appear to maximize intersubunit hydrophobic interactions[138] and are more consistent with the parallel interfaces of RmlB enzymes. It is tempting to speculate

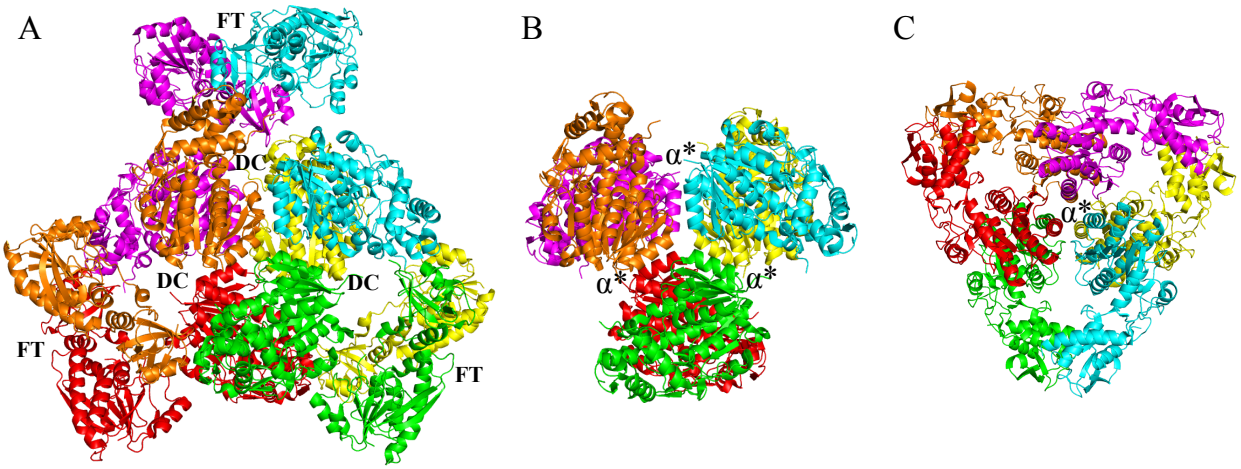


Figure 1.22. Hexameric extended SDR crystal forms. (A) Quaternary assembly in the asymmetric unit of the full-length ArnA:ATP:UGA complex (PDB entry 1Z7E). Full-length ArnA dimers (red and green, cyan and yellow or magenta and orange) contain a central core of decarboxylase (DC) domains and an outer ring of formlytransferase (FT) domain interactions. (B) The central DC core of ArnA is represented without the FT domains to show the ArnA hexamer packs three canonical SDR dimers using alpha helices (α*) from symmetry related molecules. α* helices form three bundles of two helices. (C) A hexamer found in the *Ec*GMD crystal lattice (PDB entry 1DB3) packs using two bundles of three α* helices, shown front to back in this view. *Ec*GMD is colored by dimer (as in A), and may resemble the hexamer reported by gel filtration.[134]

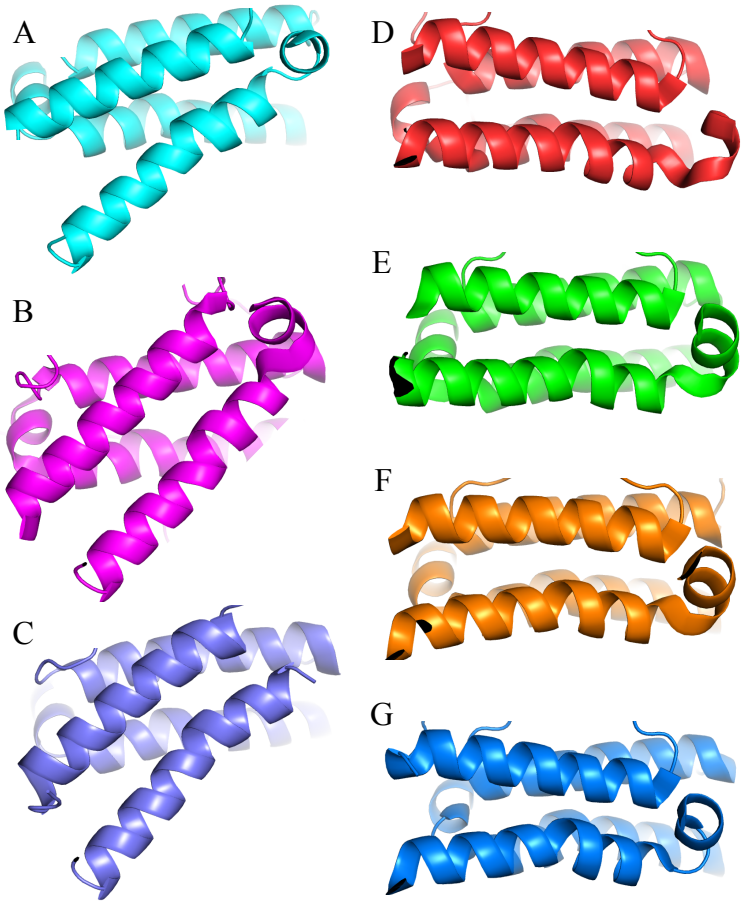


Figure 1.23. The canonical SDR dimerization surface facilitates two monomer packing arrangements. Cutaway of the dimer interface showing the four helix bundles of (A) mesophilic *Ec*GALE (PDB entry 1XEL), (B) $ArnA_{dc}$ (1Z7E) and (C) hUXS (4GLL) are distorted by as much as 40° relative to the parallel bundles found in (D) hyperthermophilic *Pc*GALE (3KO8), (E) *Ss*RmlB (1KER), (F) *Se*E_{od} (1WVG) and (G) *Pa*GMD (1RPN). Each monomer contributes two extended α-helices to the dimer interface, depicted as either roughly horizontal and faded for depth (monomer one) or in full color (monomer two). All dimer structures were superimposed onto the same $ArnA_{dc}$ monomer, in order to illustrate rotation of the second monomer.

that reorientation of the dimer interface may facilitate alternate quaternary states or contribute to cooperative ligand binding. However, there is currently no description of an extended SDR enzyme rotating about its dimer axis, despite crystallization of multiple enzymes in different space groups with various ligands.

To briefly summarize, most extended SDR enzymes structurally conserve an NAD(P)(H) binding motif and three catalytic residues used to first oxidize and then reduce nucleotide sugars. Despite conserving the extended SDR catalytic machinery, members of the UGA decarboxylase subfamily either catalyze a single hydride transfer or a full oxidation-reduction cycle. ArnA uses NAD^+ as a true substrate and releases NADH and UX4O as products, while UXS regenerates a tightly bound NAD^+ cofactor during UDX biosynthesis. Interestingly, UXS is sensitive to exogenous NAD(H), as might be expected for the exposed active site of ArnA. ArnA appears to be active as a soluble hexamer, although variable monomer, dimer, tetramer and hexamer quaternary structures are found in the extended SDR family. UXS naturally occurs as either a soluble or membrane anchored protein but lacks the formyltransferase domain reported to facilitate higher order ArnA oligomers. In order to address the differences between UGA decarboxylases and gain an understanding of UXS regulation, we set out to characterize the UGA decarboxylase catalytic cycle. Further, we sought to determine a structural and oligomeric context for human UXS activity. Our kinetic and biophysical findings are presented in the following two chapters.

CHAPTER 3

HUMAN UDP-α-D-XYLOSE SYNTHASE AND *ESCHERICHIA COLI* ARNA CONSERVE A CONFORMATIONAL SHUNT THAT CONTROLS WHETHER XYLOSE OR 4-KETO-XYLOSE IS PRODUCED[1]

[1] Reprinted with permission from Polizzi, S.J., Walsh, R.M., Jr., Peeples, W.B., Lim, J-M., Wells, L. and Wood, Z.A. *Biochemistry 51,* 4721-4731. Copyright (2012) American Chemical Society.

Abstract

Human UDP-α-D-xylose synthase (hUXS) is a member of the short-chain dehydrogenase/reductase family of nucleotide-sugar modifying enzymes. hUXS contains a bound NAD^+ cofactor that it recycles by first oxidizing UDP-α-D-glucuronic acid (UGA), and then reducing the UDP-α-D-4-keto-xylose (UX4O) to produce UDP-α-D-xylose (UDX). Despite the observation that purified hUXS contains a bound cofactor, it has been reported that exogenous NAD^+ will stimulate enzyme activity. Here we show that a small fraction of hUXS releases the NADH and UX4O intermediates as products during turnover. The resulting apoenzyme can be rescued by exogenous NAD^+, explaining the apparent stimulatory effect of added cofactor. The slow release of NADH and UX4O as side products by hUXS is reminiscent of the *Escherichia coli* UGA decarboxylase (ArnA), a related enzyme that produces NADH and UX4O as products. We report that ArnA can rebind NADH and UX4O to slowly make UDX. This means that both enzymes share the same catalytic machinery, but differ in the preferred final product. We present a bifurcated rate equation that explains how the substrate is shunted to the distinct final products. Using a new crystal structure of hUXS, we identify the structural elements of the shunt and propose that the local unfolding of the active site directs reactants toward the preferred products. Finally, we present evidence that the release of NADH and UX4O involves a cooperative conformational change that is conserved in both enzymes.

Introduction

Proteoglycans act as receptors for growth factors and are essential for cell proliferation, migration and adhesion.[140; 141; 142] Disrupting proteoglycan biosynthesis can attenuate tumor growth and progression, thus controlling proteoglycan biosynthesis is a promising strategy for treating cancer.[141; 142; 143; 144] The biosynthesis of most proteoglycans begins with the covalent attachment of xylose to the hydroxyl of a serine.[9; 54; 145] In mammals, xylose is also used in non-mucin O-glycosylation of key regulatory proteins such as Notch and α-dystroglycan that play important roles in cancer progression and metastasis.[146; 147; 148; 149] The nucleotide sugar donor that initiates xylose transfer is produced by UDP-α-D-xylose synthase (UXS; E.C. 4.1.1.35), a member of the short-chain dehydrogenase/reductase (SDR) family. Understanding the mechanism of UXS may contribute to the design of strategies that will slow or prevent metastasis.

SDR enzymes share a common catalytic domain containing a Rossmann fold for binding the NAD^+ cofactor.[150] UXS is further classified as an 'extended' SDR due to an inserted nucleotide sugar binding domain (NSBD).[44; 72] The extended SDR subfamily includes nucleotide sugar epimerases, dehydratases and decarboxylases, all of which use a similar NAD^+-dependent mechanism. The recent crystal structure of human UXS (hUXS) in complex with NAD^+ and UDP has shed light on the mechanism (Figure 2.1).[44] Briefly, hUXS uses a bound NAD^+ cofactor to oxidize of the C4′ atom of UDP-α-D-glucuronic acid (UGA) to produce UDP-α-D-4-keto-glucuronic acid. The unstable β-keto-acid intermediate decarboxylates to form the more stable UDP-α-D-4-keto-xylose (UX4O). Finally, hUXS uses the NADH cofactor produced in the first step to reduce the UX4O intermediate to UDP-α-D-xylose (UDX). hUXS purifies with a bound cofactor and therefore has no requirement for additional NAD^+.[7] Despite this observation,

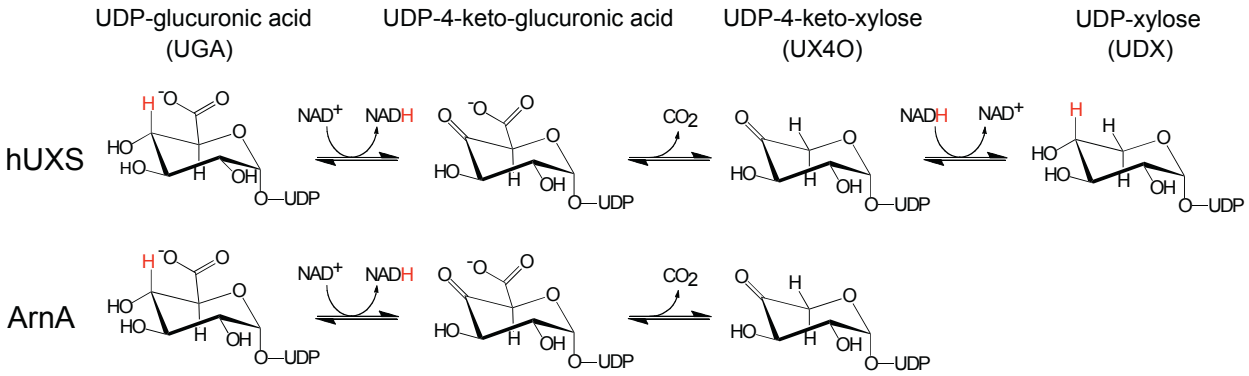


Figure 2.1. UXS and ArnA have similar catalytic mechanisms. UDP-glucuronic acid (UGA) decarboxylases catalyze NAD^+-dependent oxidation of the C4′ position of substrate to form UDP-4-keto-xylose (UX4O) and NADH. ArnA releases UX4O and NADH as products. UXS retains UX4O and NADH as intermediates to form UDP-xylose (UDX). The C4′ proton involved in hydride transfer is depicted in red.

several reports have demonstrated that adding exogenous NAD^+ can stimulate UXS activity by 10-104%.[2; 3; 4; 6; 37; 44] This stimulation has been attributed to a significant contamination of apo-enzyme in purified UXS.[2; 44]

We present evidence that the stimulatory effect of exogenous NAD^+ is due to the accumulation of apo-hUXS during catalytic turnover. We show that the hUXS mechanism shunts UGA into two different paths: i) the slow production of NADH and UX4O, or ii) the preferred conversion to UDX. The first path results in an inactive apo-enzyme that can be rescued by exogenous NAD^+. In contrast to hUXS, the homologous UGA decarboxylase domain of *E. coli* ArnA uses NAD^+ as a cosubstrate and releases NADH and UX4O as products (Figure 2.1).[23; 24] We show that cofactor binding in UXS and ArnA invokes a conserved, cooperative conformational change. Interestingly, ArnA conserves the bifurcated mechanism of hUXS and can produce UDX, albeit rather inefficiently. Using a new crystal structure of hUXS, we identify the flexible active site elements that contribute to the UX4O and NADH shunt. We present a model for this bifurcated mechanism that explains the stimulatory effect of exogenous NAD^+ on hUXS and the production of UDX by ArnA.

Materials and Methods

Protein Purification

A hUXS construct lacking the 84 residue N-terminal transmembrane domain was designed based on the crystal structure of hUXS in complex with UDP (PDB entry 2B69),[44] then expressed and purified as previously reported.[7] It has been suggested that heterologously expressed hUXS contains a significant fraction of apoenzyme. We used the following assay to determine the stoichiometry of hUXS:NAD^+ in our expressed protein. The absorbance spectrum

of purified, total hUXS (hUXS + NAD^+) was recorded on an Agilent 8453 UV-Vis spectrophotometer (Figure 2.2A). To quantitate the amount of NAD^+, the sample was recovered from the cuvette, boiled for 5 min and centrifuged to remove protein. The absorbance spectrum of the supernatant containing the cofactor was recorded and quantitated using the molar absorptivity of NAD^+ (ε_{260} = 17400 M^{-1} cm^{-1}).[151] To quantitate the amount of pure protein, the NAD^+ supernatant spectrum was subtracted from that of the total protein (hUXS + NAD^+) and the resulting spectrum (protein alone) was quantitated using the molar absorptivity of hUXS calculated with ProtParam (ε_{280} = 37360 M^{-1} cm^{-1}) (Figure 2.2A).[152] An expression construct (pMS159) containing the decarboxylase domain of *E. coli* ArnA (residues 306-660) was kindly provided by M. Sousa. This fragment has previously been shown to have the same decarboxylase activity as full length ArnA[24] and is simply referred to as ArnA throughout this work. ArnA was expressed in BL21 cells grown to an OD_{600} ~1.3 at 37 ºC in Terrific Broth containing 50 μg/ml Kanamycin (Research Products International). Expression was induced after cooling the cultures to 15 °C and adding IPTG (Research Products International) to a final concentration of 2 mM. Cells were harvested after 24 hours. Purification and cleavage of His-tagged ArnA was performed as previously described for hUXS.[7]

Capillary Electrophoresis

The products of hUXS and ArnA were resolved using a G1600 3D capillary electrophoresis system (Agilent Technologies). Reactions were quenched by plunging in liquid nitrogen and extracted with 1:1 v/v chloroform while thawing. Capillary zone electrophoresis (CZE) of the aqueous layer was performed in a fused-silica capillary (56 cm x 50 μm) with an extended light path using 20 mM borate pH 9.3 as the background electrolyte at 22kV and 18 °C.

Absorbance spectra were collected between 200-400nm and products were identified by co-migration with known standards. The UX4O standard was a kind gift of M. Bar-Peled. Product peaks were manually integrated using Agilent CE ChemStation software Rev. B.03.01, and scaled using an internal caffeine standard (5 mM; Sigma-Aldrich) for progress curves. UX4O and NADH peaks were converted to molar concentrations using the 260nm molar absorption coefficients 9820 M^{-1} cm^{-1} and 14100 M^{-1} cm^{-1},[151] respectively.

HPLC Purification and Collision Induced Dissociation (CID) Mass Spectrometry

UX4O and UDX were purified using a Dionex HPLC equipped with an AD20 absorbance detector and a GP40 gradient pump. The hUXS reaction was resolved on a CarboPac analytical column (4 mm x 250 mm) using a linear ammonium formate gradient flowing at 1 ml/min over 25 minutes. Fractions were collected and products confirmed using CZE as described above. Purified UX4O and UDX were identified using a hybrid linear ion trap-Fourier transform mass spectrometer (LTQ-Orbitrap XL ETD, ThermoFisher, San Jose, CA). UDP-sugars were dissolved in 50 μL of sample solution: 15 μL of 100% methanol followed by the addition of 35 μL of 1 mM NaOH in 50% methanol and injected directly into the mass spectrometer using a nanospray ion source with a fused-silica emitter (360 × 75 × 30 μm, SilicaTip, New Objective) at 2.0 kV capillary voltage, 240 °C capillary temperature, and a syringe flow rate of 0.5 μL/min. The $FTMS^2$ (Fourier transform mass spectrometry) spectra at 60,000 resolution was collected at 36% collision-induced dissociation (CID) from the sodiated mass of UDP-xylose (602.9992 m/z) with 2.0 m/z isolation width. The $FTMS^3$ was performed on a b3-H (358.926 m/z) fragment ion with 36% CID and 2.0 m/z isolation width. For analysis of UDP-4-keto-xylose, a MS/MS spectrum of ITMS (Ion trap mass spectrometry) was acquired at

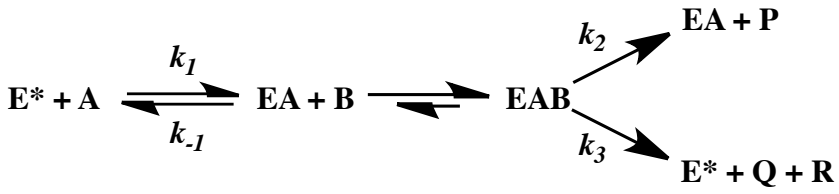
36% CID and 0.8 m/z isolation width from the sodiated mass of UDP-4-keto-xylose (600.9835 m/z).

Enzyme Assays

hUXS (1mg/ml) was assayed in standard reaction conditions (50 mM Tris pH 8.0, 10 mM DTT and 1 mM EDTA) at 25 °C. Initial velocities with variable NAD^+ (0.05-5 mM; Sigma-Aldrich) and saturating UGA (1 mM), or variable UGA (0.005-5 mM) with saturating NAD^+ (5 mM) were measured by monitoring NADH absorbance at 340 nm with an Agilent 8453 UV-Vis spectrophotometer equipped with a Peltier temperature controller set at 25 °C. The NADH concentration was calculated using the 340 nm molar absorptivity coefficient 6220 M^{-1} cm^{-1}.[151] Initial velocities for ArnA (0.5 mg/ml) were measured under the same buffer and temperature conditions as hUXS, with variable NAD^+ (0.05-15 mM) and saturating UGA (5 mM), or variable UGA (0.05-5 mM) with saturating NAD^+ (5 mM). Initial velocities were fit to the bifurcated steady state rate equation (derived below) using nonlinear regression in Prism (GraphPad Software, Inc).

The following scheme illustrates the bifurcated steady state kinetics of hUXS with variable cofactor NAD^+ (A) and saturating substrate UGA (B):

Scheme 1



E* and EA represent the apo and cofactor-bound forms of the enzyme, respectively. We assume that the concentration of EA is negligible in the presence of saturating [B] and EAB is the

relevant Michaelis complex. The divergent pathways that yield the reduced product UDX (P), or release the reaction intermediates NADH (Q) and UX4O (R) are governed by the distinct rate constants, k_2 and k_3, respectively. The fraction of [EAB] that produces and releases the intermediates during each catalytic cycle is:

(1) $$f_{release} = \frac{k_3}{k_2 + k_3}[EAB]$$

The rate equation for the release of the intermediate products Q and R is:

(2) $$V_0(A) = k_3 f_{release} = \frac{k_3 k_3}{k_2 + k_3}[EAB]$$

To solve this equation, we assume the steady state condition (3) and express [E*] in terms of total enzyme (4):

(3) $$k_1[E^*][A] = k_{-1}[EAB] + k_2[EAB] + k_3[EAB]$$

(4) $$[E^*] = [E]_{Total} - [EAB]$$

Substitute (4) into (3) and simplify:

(5) $$k_1([E]_{Total} - [EAB])[A] = (k_{-1} + k_2 + k_3)[EAB]$$

Solve for [EAB] and rearrange,

(6) $$[EAB] = \frac{k_1[E]_{Total}[A]}{(k_{-1} + k_2 + k_3) + k_1[A]} = \frac{[E]_{Total}[A]}{\frac{(k_{-1} + k_2 + k_3)}{k_1} + [A]} \quad ; \quad \text{let } \frac{(k_{-1} + k_2 + k_3)}{k_1} = K_M^{app}$$

(7) $$[EAB] = \frac{[E]_{Total}[A]}{K_M^{app} + [A]}$$

Substitute (7) into (2):

(8) $$V_0(A) = \frac{\frac{k_3 k_3}{k_2 + k_3}[E]_{Total}[A]}{K_M^{app} + [A]} = \frac{V_{\max}^{app}[A]}{K_M^{app} + [A]}$$

Finally, we express the rate equation in the more general form to accommodate cooperativity by assuming that the Hill coefficient (h) may not be unity:

(9) $$V_0(A) = \frac{V_{\max}^{app}[A]^h}{(K_M^{app})^h + [A]^h}$$

Crystallization, Data Collection and Refinement

hUXS crystals were obtained with precipitant containing 1.3 M ammonium sulfate, 0.1 M magnesium formate and 0.15% 1.5K PEG at 26°C using the hanging-drop vapor-diffusion method. Rod-shaped crystals took 7 to 21 days to appear and grew to approximately 0.15 mm in length. hUXS crystallized in space group $P3_221$ with unit cell dimensions a = 125.6 Å, c = 98.9 Å and two molecules per asymmetric unit. For data collection, crystals were cryoprotected by passing through a mixture of paratone:paraffin (50:50 ratio) and then plunged into liquid nitrogen. Data were collected on beamline 8.2.1 at the Advanced Light Source (ALS, Berkeley, CA) using an ADSC Q210 detector. Data sets were processed with both the HKL suite of programs[153] and XDS.[154] Data collection statistics are reported in Table 2.1.

The structure of hUXS was solved by molecular replacement using the CNS suite of programs.[155] The crystal structure of a hUXS monomer (PDB entry 2B69) was used as a search model. For molecular replacement, the ligands and waters were removed from the search model, and 5% of the data were set aside for cross validation.[156] Two monomers were placed in the asymmetric unit to form the complete biological dimer. Iterative cycles of model refinement and manual rebuilding were carried out using REFMAC[157] and Coot,[158] respectively. Model refinement statistics are reported in Table 2.1.

Table 2.1: Crystallographic Data and Refinement Statistics for Open hUXS

Data Collection Statistics		Model Refinement Statistics	
Wavelength (Å)	1.00	No. of amino acids	760
Resolution (Å)	50 - 2.5	No. of NAD^+	2
Total observations	343,332	No. of waters	146
Unique reflections	30,682	Total non-H atoms	5084
Completeness (%)	97.3 (90) [a]	Average B ($Å^2$)	39.7
Multiplicity	11.2 (11.0)	R-factor	0.178 (0.227)
<I/s>	26.2 (4.7)	Free R-factor	0.223 (0.303)
R_{meas} [b]	8.3 (57.5)	Stereochemical ideality	
R_{mrgd-F} [c]	6.4 (31.1)	bond length rmsd (Å)	0.014
		bond angle rmsd (°)	1.44
		f,y most favored (%)	97.4
		f,y allowed (%)	2.6

[a] Numbers in parenthesis represent the highest resolution shell (2.57 to 2.5 Å)

[b] R_{meas} is the multiplicity weighted merging R-factor[159]

[c] R_{mrgd-F} is a measure of data accuracy[159]

Structural Analyses

Structural comparisons were performed using the atomic coordinates of unliganded hUXS (PDB entry 4GLL), the hUXS:UDP complex (2B69),[44] apo-ArnA (2BLL),[26] and the decarboxylase domain of the full length ArnA:ATP:UGA complex (1Z7E).[25] Main chain B-factors were averaged with BAVERAGE.[157] Solvent accessible surface area was calculated using AREAIMOL[157] with a probe radius of 1.4 Å. Sequence identities were calculated using NCBI-BLAST[160] and/or DaliLite.[161] LIGPLOT[93] was used to analyze hydrogen bonding and packing interactions between hUXS and the bound ligands. Substrate induced conformational changes were analyzed using difference distance matrix analysis.[162] The NSBD rotation axis was generated using DYNDOM.[76] Structural figures were generated using Pymol (http://www.pymol.org).

Results

Human UXS releases the reaction intermediates NADH and UDP-4-keto-xylose.

We expressed human UXS lacking the N-terminal membrane-spanning domain (hUXS; residues 85-420) as described elsewhere.[7] In our hands, recombinant hUXS purifies with the NAD^+ bound in stoichiometric amounts (see Methods). The average of three different protein preps yielded 1.07 ± 0.05 moles of NAD^+ per mole of protein (Figure 2.2A). Using CZE, we show that recombinant hUXS converts UGA to UDX without need of additional cofactor (Figure 2.2B). In the presence of exogenous NAD^+, two additional products are formed (Figure 2.2B). The first product was identified as NADH based on co-migration with known standards and a characteristic absorbance spectrum (Figure 2.2C). The second product has a UV absorbance spectrum matching a UDP-sugar (Figure 2.2D). Based on molar absorptivity of uridine, the

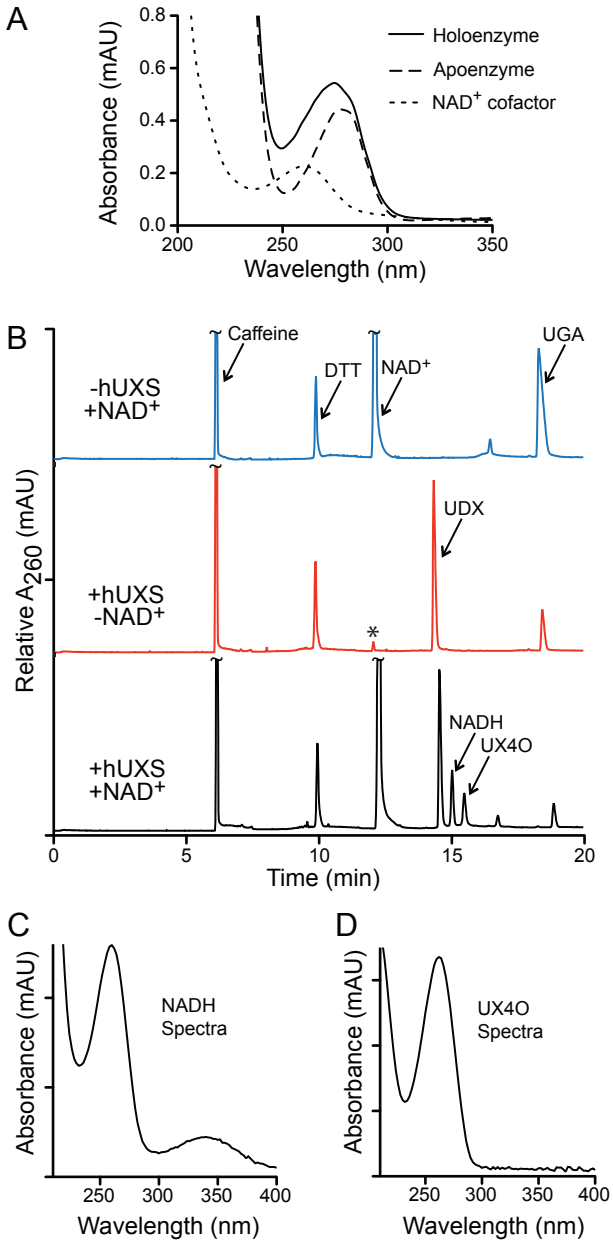


Figure 2.2. hUXS releases UX4O and NADH in the presence of NAD^+. (A) Absorbance spectra of purified recombinant hUXS holoenzyme (solid line). Boiling releases NAD^+ cofactor (dotted line) in stoichiometric amounts with apoenzyme (dashed line). The apoenzyme spectrum was calculated as the difference between the total protein and NAD^+ spectra (see Methods). (B) Capillary zone electrophoresis (CZE) chromatogram of hUXS reaction with NAD^+ only (blue, upper trace), hUXS only (red, middle trace), or both NAD^+ and hUXS (black, lower trace). The bound NAD^+ cofactor in recombinant hUXS is depicted with an asterisk (*). Absorbance spectra measured at 260 nm (A_{260}) are offset vertically and the caffeine standard and NAD^+ peaks have been truncated (~) for ease of viewing. (C) Absorbance spectra of NADH showing characteristic peaks at 260 nm and 340 nm and (D) UX4O showing a single peak at 262 nm consistent with the uridine moiety of UDP-sugars. Standard reactions with or without 2.5 mM NAD^+ and 1 mg/ml hUXS were incubated with 1 mM UGA at 37 °C and quenched with chloroform after 1 hour.

additional product is formed in stoichiometric amounts with NADH (1.1:1) and is consistent with a UDP containing reaction intermediate from the hUXS catalytic cycle (Figure 2.1). We postulated that the product was UX4O, since UDP-4-keto-glucuronic acid is a labile β-keto acid and would be expected to undergo decarboxylation. In support of our hypothesis, the UDP-containing product co-migrates with a UX4O standard during CZE (data not shown).

We used mass spectrometry to confirm the identity of the putative UX4O. The UDP-containing products were purified using HPLC. We identified fractions corresponding to UDX and UGA, but we did not detect a separate peak for UX4O (Figure 2.3A). We used the higher resolution of CZE to show UX4O comigrates with UDX during HPLC purification (Figure 2.3B). The HPLC purified UDX/UX4O fraction was analyzed with a hybrid linear ion trap-Fourier transform mass spectrometer (see Methods). Collision-induced dissociation mass spectroscopy ($CID\text{-}MS^n$) confirms that the HPLC fraction contains two species: UDX and UX4O. The major species is consistent with sodiated UDX cleaving along the glycosidic bond of uridine with hydrogen rearrangement (b4-H, 490.967 m/z) and the phosphoester bond of phosphate with rearrangement (b3-H, 358.926 m/z and b2-H, 256.978 m/z respectively) (Figure 2.3C). The minor species shows the same cleavage pattern as UDX, but with a mass reduced by 2 Da (Figure 2.3D), consistent with the loss of two hydrogen atoms to form the C4′ ketone of UX4O.

hUXS and ArnA conserve NAD^+ substrate cooperativity.

Our observation that hUXS produces both UDX and the intermediates NADH and UX4O as terminal products indicates that the reaction mechanism is bifurcated (Figure 2.4A). The release of the reaction intermediates from hUXS is ~2 orders of magnitude slower than the

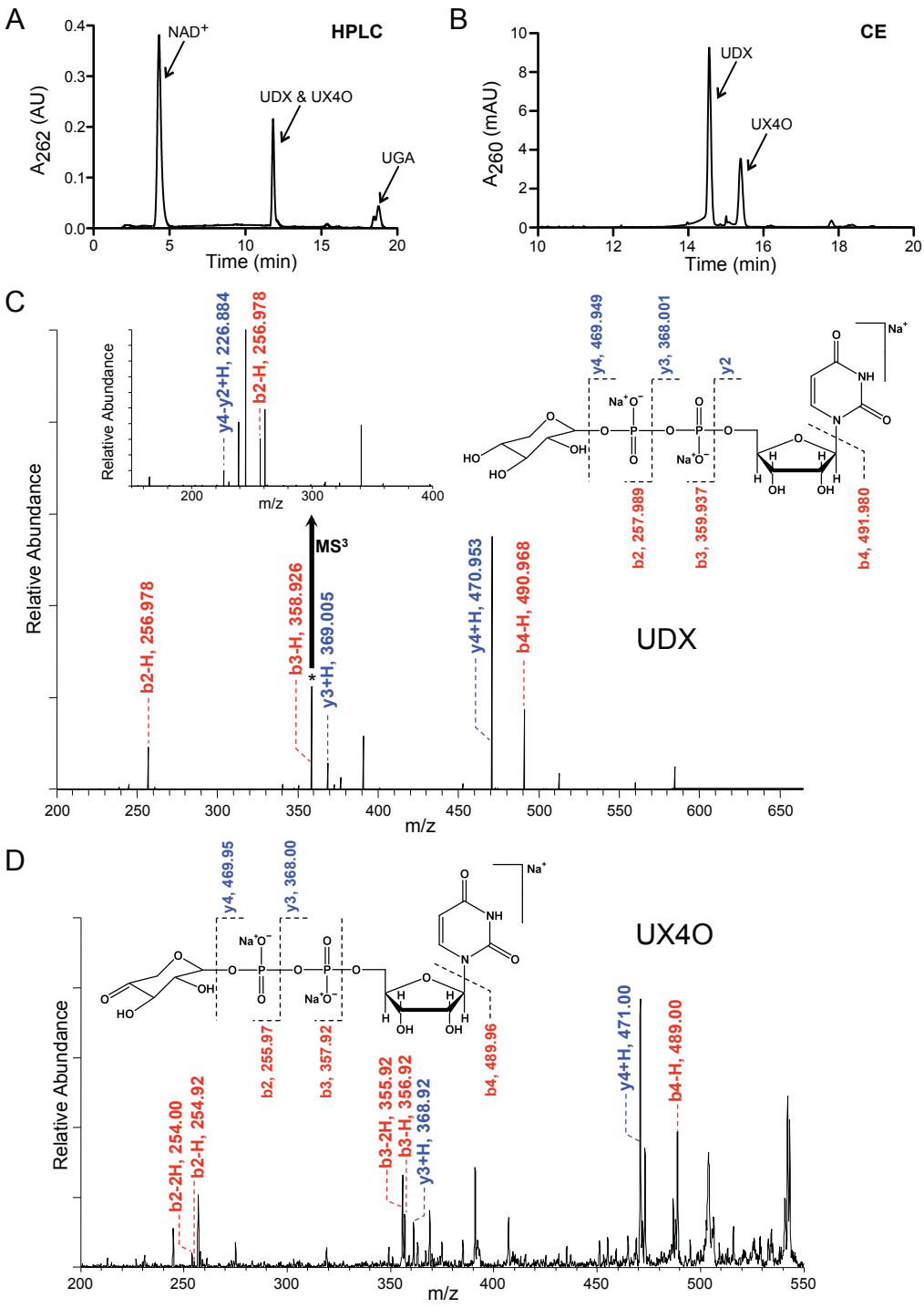


Figure 2.3. Identification of UX4O. (A) HPLC chromatogram of the hUXS reaction showing UX4O and UDX co-migrate as a single product peak between NAD^+ and UGA. (B) CZE chromatogram of the HPLC purified UDX fraction, confirming UX4O and UDX co-migrate by HPLC but can be resolved with CZE. (C) Collision induced dissociation mass spectra ($CID\text{-}MS^n$) of the HPLC product peak, showing the major hUXS product is consistent with sodiated UDX and (D) the minor product corresponds to sodiated UX4O. Sodiated UDP-sugars are diagrammed to show the expected masses of the fragment ions before proton rearrangement.

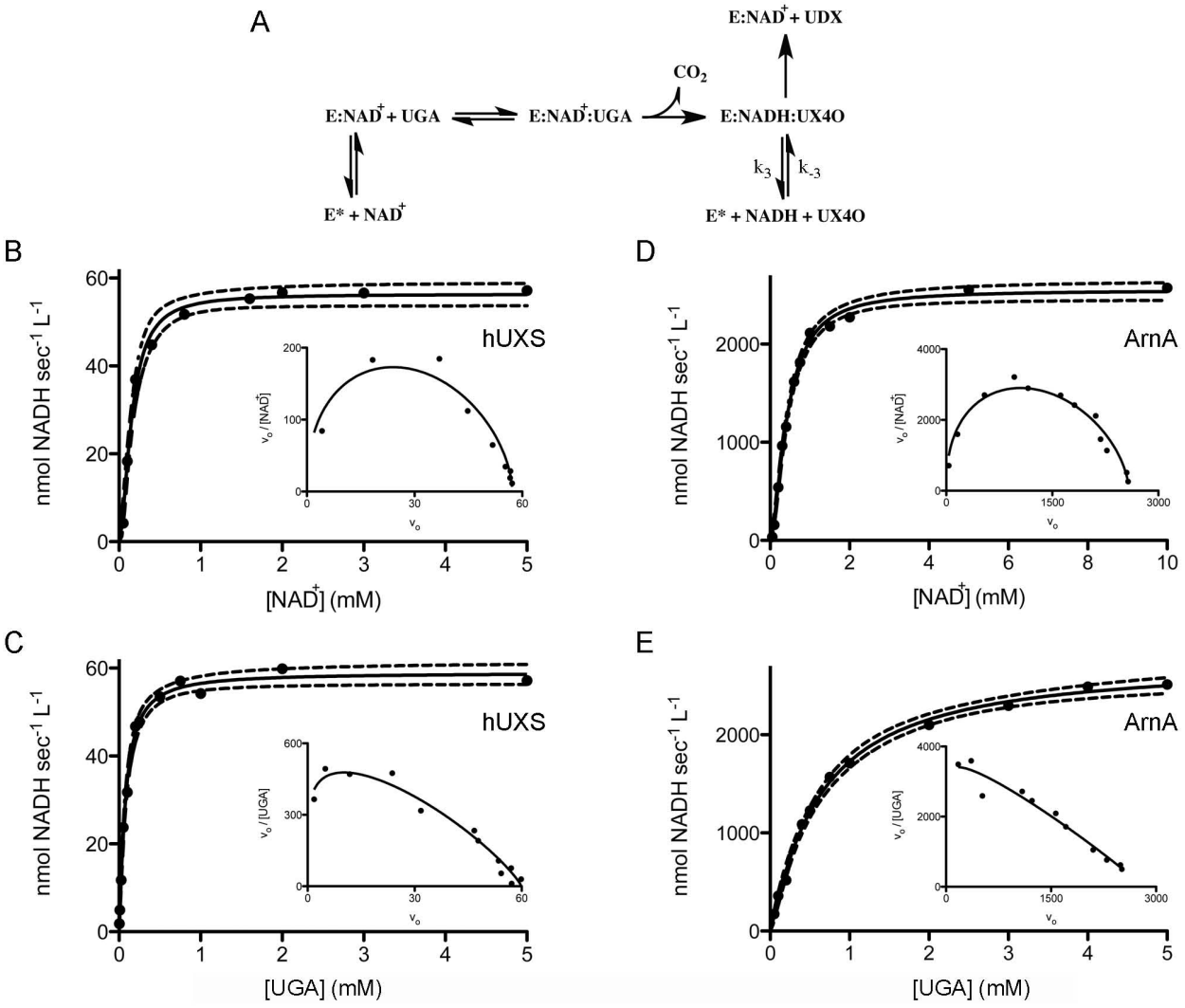


Figure 2.4. Decarboxylase activities of hUXS and ArnA. (A) Bifurcated reaction scheme for UGA decarboxylases. hUXS purifies with NAD^+ bound ($E:NAD^+$). UGA binding results in formation of the Michaelis complex ($E:NAD^+:UGA$). UGA decarboxylation evolves CO_2 and leads to formation of the intermediate complex (E:NADH:UX4O). hUXS favors a second hydride transfer to form UDX, but can release the NADH and UX4O intermediates using a catalytic shunt (k_3). (B,C) Substrate saturation curves for hUXS release of NADH and UX4O. Initial velocity data were fit using nonlinear regression to the bifurcated rate equation (see Methods for derivation) and 95% confidence intervals (dashed lines) are depicted. Eadie-Scatchard plots (insets) are presented for illustrative purposes only. (D,E) Substrate saturation curves for ArnA release of NADH and UX4O depicted as above.

production of UDX (specific activities of 3.45 ± 0.08 and 534 ± 29 nmol mg^{-1} min^{-1}, respectively). This means that the release of the NADH and UX4O intermediates only weakly competes with the synthesis of UDX. Still, to accurately describe the steady state kinetics of the slower NADH and UX4O release, we derived a bifurcated rate equation for fitting the substrate saturation curves (see Methods). The NAD^+-dependent substrate saturation curve for hUXS intermediate release is sigmoidal (Hill coefficient = 1.8) (Table 2.2), showing that cofactor binding to apo-hUXS (E*) is cooperative (Figure 2.4B). The UGA substrate saturation curve is also cooperative, albeit with a significantly lower Hill of 1.2 (Figure 2.4C). Next, we examined the *E. coli* UGA decarboxylase ArnA. ArnA conserves the first catalytic step with hUXS to produce NADH and UX4O as end products (Figure 2.1). Like hUXS, the steady state kinetics of ArnA also displays a strong NAD^+ dependent cooperativity (Hill = 1.7) (Figure 2.4D). The UGA substrate saturation curves show very little, if any, cooperativity in ArnA (Hill = 1.1) (Figure 2.4E).

ArnA catalyzes UDP-xylose synthesis.

Based on the similarities in structure and chemistry, we asked if ArnA has a bifurcated reaction mechanism (Figure 2.4A). Using CZE, we show that ArnA produces both UX4O and UDX (Figure 2.5A). The ArnA progress curves for UDX production display a distinctive lag in activity (Figure 2.5B). A lag in progress curves usually indicates the presence of a transient in the reaction and can be modeled using the equation:[163]

$$P = V_{ss} - (V_{ss} - V_i)(1 - e^{-t/\tau})\tau$$

where P is the product concentration at time, t, and τ is the half-life for the transition between the initial (V_i) and final steady state velocities (V_{ss}). The τ calculated from the UDX progress curve is

Table 2.2. Steady state kinetic parameters for UGA decarboxylation reactions.

Enzyme [a]	Substrate dependence	Kinetic property [b]	Nonlinear regression best-fit value [c]
hUXS	NAD^+	V_{max}^{app}	56.4 ± 1.1
		h	1.76 ± 0.18
		K_M^{app}	155 ± 10
	UGA	V_{max}^{app}	58.9 ± 1.1
		h	1.22 ± 0.09
		K_M^{app}	75.8 ± 11.2
ArnA	NAD^+	V_{max}^{app}	2550 ± 43
		h	1.65 ± 0.09
		K_M^{app}	433 ± 16
	UGA	V_{max}^{app}	2755 ± 85
		h	1.10 ± 0.07
		K_M^{app}	623 ± 50

[a] Enzyme assays monitoring NADH absorbance at 340 nm were performed under standard conditions. hUXS (1 mg/ml) was incubated with varying NAD^+ (0.05-5 mM) and saturating UGA (1 mM) or varying UGA (0.005-5 mM) and saturating NAD^+ (5 mM). ArnA (0.5 mg/ml) was incubated with varying NAD^+ (0.05-15 mM) and saturating UGA (5 mM) or varying UGA (0.05-5 mM) and saturating NAD^+ (5 mM).

[b] Apparent maximum reaction velocity (V_{max}^{app}) is reported in nmol L^{-1} sec^{-1}, the Hill coefficient (h) is unitless, and the apparent Michaelis constant (K_M^{app}) is reported in μM.

[c] Values were fit by nonlinear regression to the bifurcated rate equation (see Methods for derivation) as implemented in Prism.

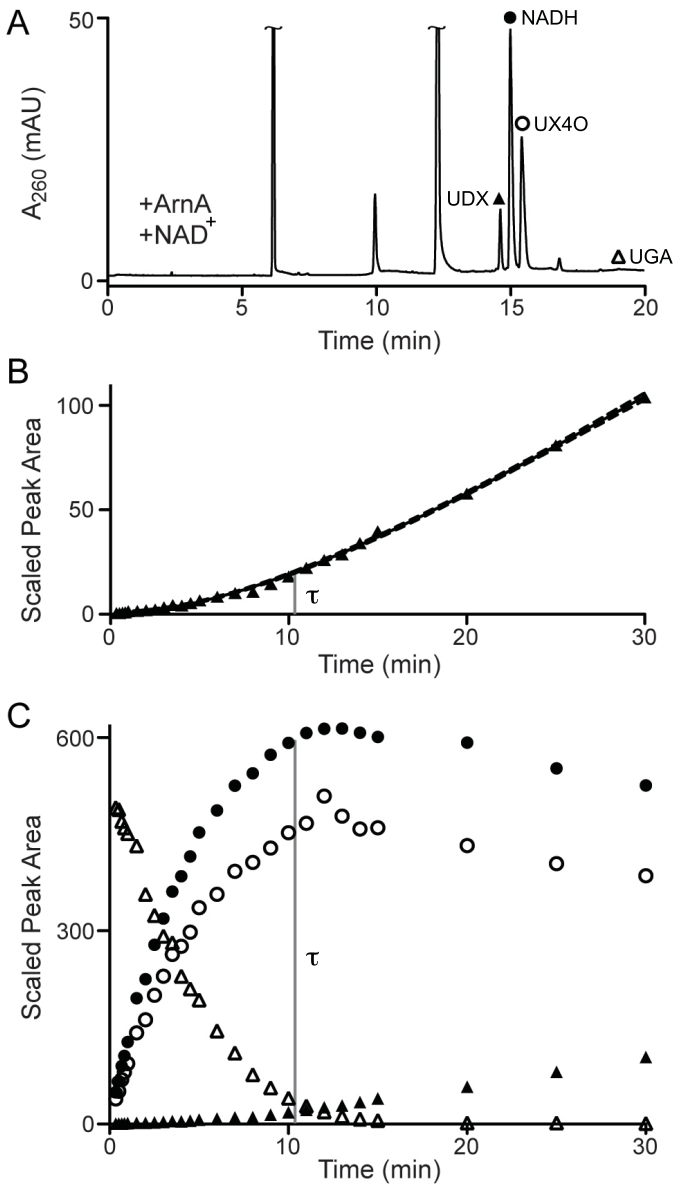


Figure 2.5. ArnA catalyzes UDP-xylose synthesis. (A) CZE chromatogram of ArnA reaction showing UDX (▲), NADH (●) and UX4O (o) are produced from NAD^+ and UGA (Δ). Peaks were identified by co-migration with known standards and caffeine and NAD^+ peaks are truncated (~) for ease of viewing. (B) Progress curve showing ArnA displays a lag in UDX production. UDX was resolved by CZE and product peaks were scaled using the internal caffeine standard. Data was fit by nonlinear regression to the equation of Frieden[163] and the half-life for the transition (τ) is denoted by a grey line. The 95% confidence interval (dashed line) is narrow and superimposes with the curve and data points. (C) Progress curves of the ArnA reaction following all species labeled above. τ (grey line) is extrapolated to show production of UDX increases after UGA is depleted and UX4O and NADH accumulate.

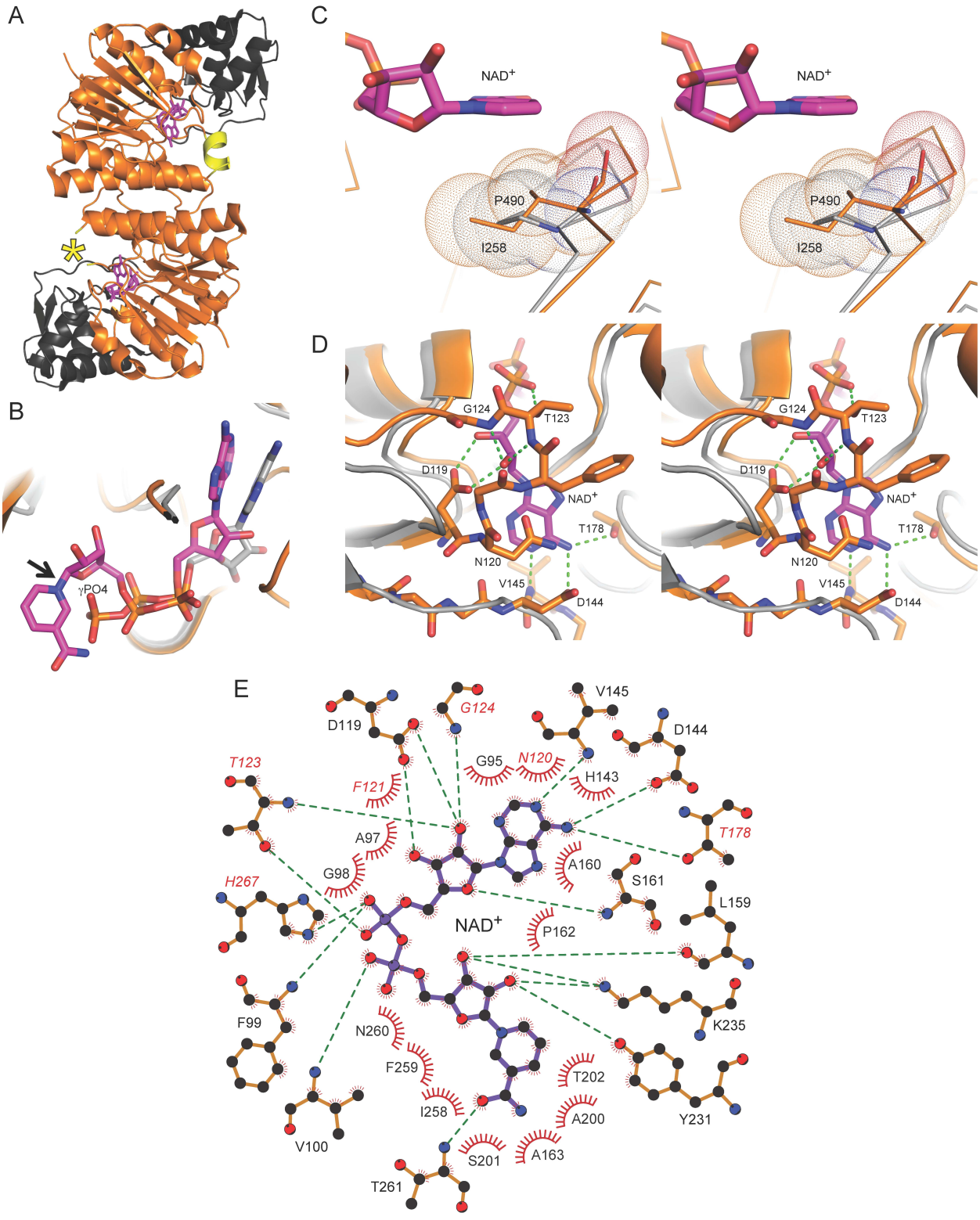
coincident with the plateaus of NADH and UX4O and the depletion of UGA (Figure 2.5C). In addition, the final V_{ss} of UDX production is reflected in the depletion of UX4O and NADH. This supports a mechanism in which UX4O and NADH rebind to ArnA to produce UDX.

NAD^+ is buried in the core of hUXS.

We solved the crystal structure of hUXS at 2.5 Å resolution, revealing a dimer in the asymmetric unit (Figure 2.6A). The extreme N-terminus of both chains is disordered (residues 85 to 87), as is the C-terminus of chain A (395 to 420) and chain B (399 to 420). In addition, residues 166-170 are disordered in chain B. The equivalent residues in chain A form a helix that is buried in a crystal contact. Crystal lattice formation selects for stable conformations and is too weak to deform protein structures,[164; 165] suggesting residues 166-170 are flexible in solution, and the helix we observe is a consequence of crystal packing (discussed below). The structure of hUXS is divided into a nucleotide sugar-binding domain (NSBD) and a catalytic domain (Figure 2.6A). Because the crystal structure of hUXS in complex with UDP as a substrate analog (hUXS:UDP) has recently been described,[44] we will focus our analysis on the conformational changes associated with cofactor and substrate binding.

Recombinant hUXS crystallizes with the co-purified NAD^+ cofactor bound (Figure 2.6). While no structures of *E. coli* ArnA contain NAD^+, one structure was solved with ATP bound as a cofactor analog (PDB entry 1Z7E).[25] hUXS and *E. coli* ArnA share 28% sequence identity and superimpose 305 corresponding Cα atoms with an rmsd of 1.9 Å. The superposition shows that hUXS and ArnA bind cofactor in a similar binding cleft (Figure 2.6B). In hUXS, the *re* face of the nicotinamide ring is supported by packing interactions with the side chain and carbonyl oxygen of Ile258 (Figure 2.6C). ArnA conserves a surprisingly similar nicotinamide packing

Figure 2.6. NAD^+ is buried in the hUXS catalytic domain. (A) The dimeric structure of hUXS with the nucleotide sugar binding (NSBD) and catalytic domains identified (black and orange, respectively). $Helix_{165\text{-}168}$ (yellow) is ordered in one subunit and disordered in the other (yellow asterisk). NAD^+ (purple sticks) is depicted in the catalytic domain. (B) Structural overlay of hUXS (orange) and the ArnA:UGA:ATP complex (PDB entry 1Z7E, grey) showing the adenine bases in NAD^+ (purple) and ATP (grey) are bound in altered conformations, and the smaller γ-phosphate is non-isosteric with the nicotinamide ring and ribose bound in hUXS. The NAD^+ χ_N torsion angle is identified (arrow). (C) Stereodiagram showing Ile258 from hUXS (orange) and Pro490 from ArnA (grey) form a similar van der Waals packing surface (dots) in the nicotinamide binding pocket. (D) Stereodiagram of the NAD^+ binding pockets of hUXS and ArnA colored as above to illustrate the hUXS cofactor binding $Loop_{120\text{-}124}$ and hydrogen bonding network (dashed lines). (E) Ligplot of hydrogen bonding interactions (dashed lines) and packing interactions (feathered lines) between hUXS and NAD^+. Red italics identify interactions in hUXS that are not present in ArnA.


A
B
C
D
E
NAD+
P490
I258
G124
T123
D119
N120
V145
D144
T178
γPO4
G95
H143
F121
A97
G98
H267
A160
S161
L159
P162
F99
N260
F259
I258
T202
A200
Y231
K235
V100
S201
A163
T261

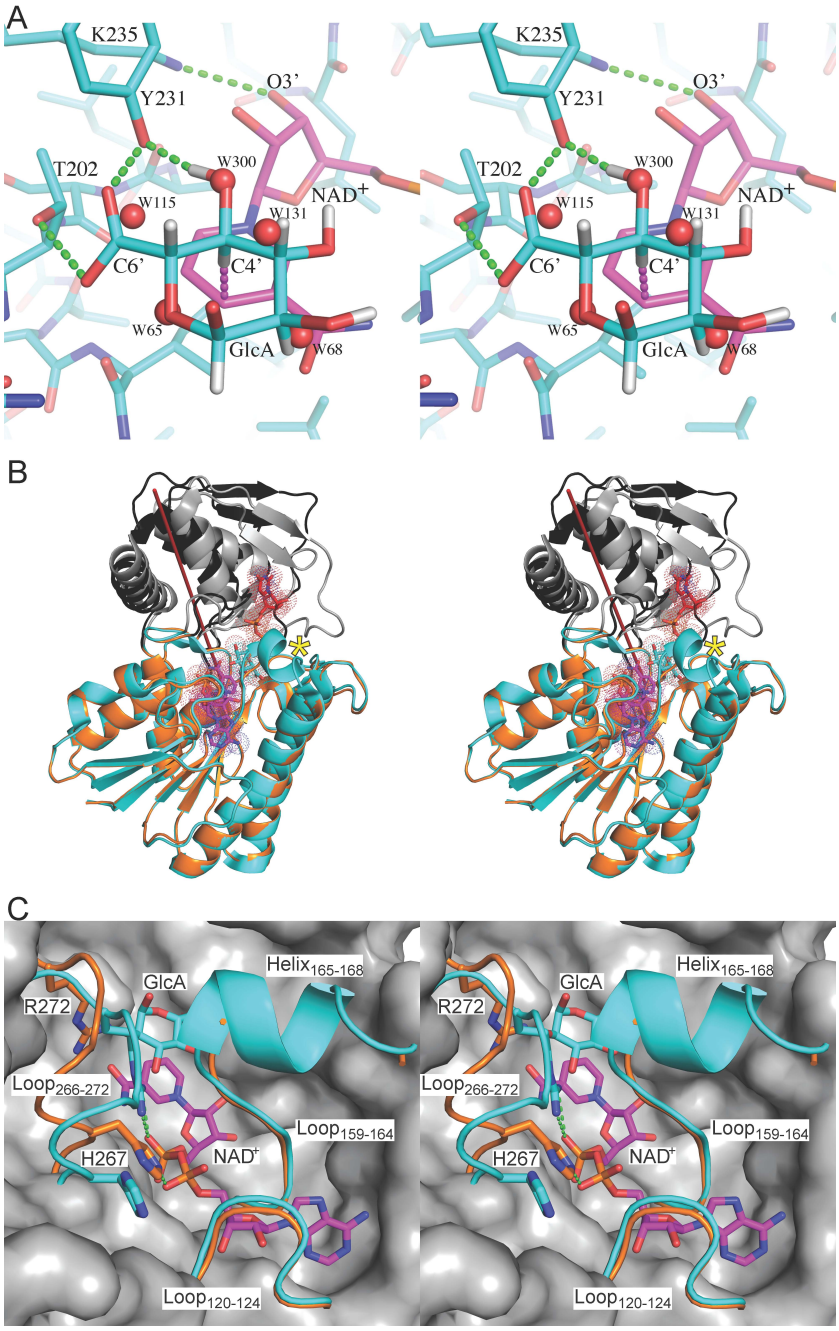
surface through an unorthodox substitution; Pro490 of ArnA is structurally equivalent to the non-isosteric Ile258 of hUXS (Figure 2.6C). To accomplish this, ArnA has undergone small structural rearrangements that shift the main chain Cα atom of Pro490 by ~1.2 Å relative to Ile258 in the hUXS structure. Recently, the Seattle Structural Genomics Center for Infectious Disease deposited the atomic coordinates of an annotated UGA decarboxylase from *Burkholderia pseudomallei* (PDB entry 3SLG) which conserves the same proline packing surface associated with the nicotinamide binding (not shown).

The adenosyl base of the cofactor is significantly shifted between the hUXS and ArnA structures, primarily due to differences in $Loop_{120\text{-}124}$ (hUXS numbering) (Figure 2.6D). In hUXS, the amide group of Asn120 packs on top of the adenylate base, and Thr123 and Gly124 contribute a total of three hydrogen bonds. In addition, Thr178 and His267 contribute hydrogen bonds to the N6 of the adenylate base and α-phosphate, respectively (Figure 2.6D). In hUXS, only ~4% of the cofactor is exposed to solvent. The $Loop_{120\text{-}124}$, Thr178 and His267 interactions are not conserved in ArnA (Figure 2.6D,E). As a consequence, ArnA binds the adenosine base and ribosyl in an altered conformation and exposes more of the adenylate to solvent (Figure 2.6B,D). The remaining hydrogen bonds and packing interactions with the cofactor are conserved in both hUXS and ArnA (Figure 2.6E).

Modeled GlcA binding in the hUXS active site.

To better understand the conformational changes associated with NADH and UX4O release, we modeled GlcA in the hUXS active site (Figure 2.7A). As a starting model, we used the crystal structure of a hUXS complex containing UDP and NAD^+ (PDB entry 2B69).[44] The chemistry of hydride transfer provides restraints for modeling a likely binding mode for GlcA in

Figure 2.7. UDP binding in hUXS induces conformational changes to the NSBD and active site. (A) Stereodiagram of the closed hUXS:UDP complex active site with GlcA modeled above the bound NAD^+ (purple). Carbon atoms C4′ and C6′ of the GlcA model are indicated, and hydrogens are included for clarity. Red spheres indicate the waters (labeled W) that occupy the active site in the absence of substrate. Hydrogen bonds are depicted as green dashes, and the path of hydride transfer from the sugar C4′ to the nicotinamide C4 is indicated by purple dashes. (B) Stereodiagram of the open hUXS monomer (orange with black NSBD) and NAD^+ (purple sticks) superimposed onto closed hUXS (teal with grey NSBD) and UDP (red sticks). Van der Waal surfaces (dots) are shown for bound ligands and modeled GlcA (teal sticks). The DYNDOM[76] calculated axis about which the NSBD rotates is depicted as a red line and the disordered $Helix_{165\text{-}168}$ is labeled with an asterisk. (C) Stereodiagram of the open (orange) and closed (teal) hUXS active sites showing the catalytic shunt elements $Loop_{120\text{-}124}$, $Loop_{266\text{-}272}$, $Loop_{159\text{-}164}$ and $Helix_{165\text{-}168}$ relative to native NAD^+ cofactor (purple sticks) and modeled GlcA substrate (teal sticks). UDP binding stabilizes $Helix_{165\text{-}168}$ and induces R272 and H267 to change conformation, thereby altering the hydrogen bond (green dashes) network to the NAD^+ diphosphates. Chain B of open hUXS was used in the superposition to illustrate the effect of $Helix_{165\text{-}168}$ disorder on active site accessibility.


A
K235
Y231
O3'
T202
W300
W115
W131
NAD+
C6'
C4'
W65
GlcA
W68
B
C
GlcA
Helix165-168
R272
Loop266-272
Loop159-164
H267
NAD+
Loop120-124

the active site, specifically: (i) hydride transfer to NAD^+ requires that the carbon and hydrogen atoms of the C4′ position of GlcA are almost in-line with the C4 atom of the nicotinamide (C4′-H′-C4 angle slightly less than 180°);[166] (ii) the angle between the GlcA C4′ and the nicotinamide C4-N1 atoms varies between 81° to 145°;[119; 167; 168] and (iii) the distance between the C4′-C4 atoms should be approximately 3.5 Å.[119; 168] Using only these constraints, we modeled the position and orientation of GlcA such that the C1′ atom could bond with the β-phosphate of UDP in the hUXS:UDP structure. Support for our modeled conformation can be seen in the ordered water structure of the active site, which closely approximates the volume, shape and hydrogen bonding requirements of GlcA (Figure 2.7A). Similarly ordered water structures are often seen compensating for missing sugar molecules in carbohydrate binding enzymes.[169; 170; 171] Even though no side chains were adjusted in this modeling exercise, the GlcA conformation is consistent with the active site architecture and hydrogen bonding requirements.

Substrate binding in hUXS reveals structural elements involved in the release of reaction intermediates.

Our crystal structure represents the NAD^+-bound, 'open' conformation of hUXS. A comparison with the 'closed' hUXS:UDP structure shows that UDP binding causes the NSBD to undergo a ~34° domain rotation to close the active site (Figure 2.7B). The superposition identifies three flexible regions at the NSBD/catalytic domain junction that contribute to substrate and cofactor retention: i) $Helix_{165\text{-}168}$, ii) $Loop_{159\text{-}164}$ and iii) $Loop_{266\text{-}272}$ (Figure 2.7C). $Helix_{165\text{-}168}$ undergoes a disorder-to-order transition to fold over the sugar substrate and close off the active site from solvent (Figure 2.7C). A similar substrate induced folding of the equivalent helix in dTDP-glucose 4,6-dehydratase (RmlB) has been reported.[119] $Loop_{159\text{-}164}$ folds over and

retains the cofactor. $Loop_{159-164}$ is more rigid when $Helix_{165-168}$ folds, based on a 70% reduction in Cα *B*-factors between chain A and B in our hUXS structure (average *B*-factors for entire chains 39.5 and 39.9 $Å^2$, respectively). $Loop_{266-272}$ is part of the NSBD, and shifts to shield the substrate and cofactor as the domain rotates to the closed conformation. This shift breaks a hydrogen bond between the imidazole of His263 and the α-phosphate, and forms a new hydrogen bond between the guanidinium of Arg272 and the β-phosphate of NAD^+ (Figure 2.7C).

The ArnA:ATP structure reveals an induced-fit conformational change associated with cofactor binding.

All of the conformational changes associated with substrate binding in hUXS are conserved in ArnA (see previous section). In addition, ArnA allows us to examine changes associated with the transition from the apo- to the holoenzyme. A comparison of the catalytic domains from apo-ArnA and the ArnA:ATP:UGA ternary complex (PDB entries 2BLL and 1Z7E, respectively) identifies three regions that flank the NAD^+-binding site and shift in response to cofactor binding (Figure 2.8). First, the N-terminus of $Helix_{326-337}$ shifts by 2.2 Å to facilitate the conserved interaction between the Rossmann GxxGxxG motif and the pyrophosphate of NAD^+. A second structural element, $Loop_{347-349}$, moves 2.6 Å toward the cofactor binding site to facilitate conserved hydrogen bonds between Asp347 (Asp119 in hUXS) and the C2′ and C3′ hydroxyls of the adenine ribose (Figures 6E and 8). $Loop_{347-349}$ also shifts towards the cofactor binding pocket to interact with the adenosine moiety. Finally, the N-terminal loop of $Helix_{373-381}$ shifts away from the dimer interface by 2.8 Å and maintains packing interactions with $Loop_{347-349}$. The side chain of Asp368 (Asp144 in hUXS) and the amide nitrogen of Ile369 form hydrogen bonds with the adenylate base N6 and N1, respectively. All

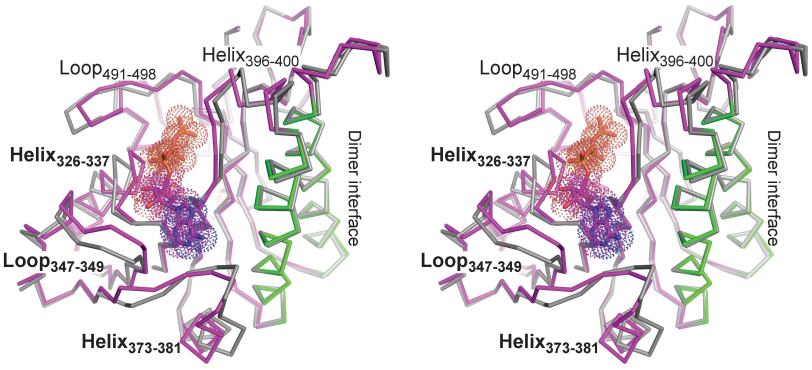


Figure 2.8. Cofactor binding in ArnA induces additional conformational changes in the active site. Stereodiagram of the apo-ArnA (grey; PDB entry 2BLL) and ArnA:ATP:UGA (purple; PDB entry 1Z7E) catalytic domains showing the organization of mobile elements around the cofactor analog ATP (van der Waals spheres). ArnA $Helix_{396-400}$ and $Loop_{491-498}$ are analogous to hUXS $Helix_{165-168}$ and $Loop_{266-272}$, respectively (Figure 7C). Bold letters indicate elements in the cofactor binding site that undergo conformational changes not present in the hUXS crystal structures. The dimerization helices (green) and interface are labeled in the ArnA:ATP:UGA complex to indicate the proximity of mobile elements to the adjacent subunit.

three cofactor interactions mediated by the ArnA elements ($Helix_{326-337}$, $Loop_{347-349}$ and $Helix_{373-381}$) are conserved in the hUXS structures.

Discussion

The bifurcated mechanism of hUXS explains the stimulatory effect of exogenous NAD^+.

Previous studies have reported that exogenous NAD^+ can increase hUXS activity by as much as 40%.[44] Given that recombinant hUXS purifies with a tightly bound NAD^+ cofactor, and the sequential oxidation and reduction steps of the catalytic cycle regenerates the cofactor (Figure2.1),[7] the mechanism by which exogenous NAD^+ could stimulate activity was not clear. It has been proposed that heterologously expressed hUXS produces a significant amount of apoenzyme.[44] Thus, the stimulatory effect of exogenous NAD^+ is due to the apoenzyme binding cofactor. We have shown that our recombinant hUXS purifies with stoichiometric amounts of cofactor (Figure 2.2A). We also show that hUXS can catalytically produce NADH and UX4O in the presence of saturating NAD^+ and substrate (Figure 2.2). The use of CZE was key to observing the slow release of the reaction intermediates. Previous studies[3; 4; 6; 37] relied on the lower resolution techniques HPLC or TLC to assay UXS, which makes the separation and identification of the UX4O intermediate difficult (Figure 2.3A,B). This suggests that HPLC should be used with caution in future studies of UXS and related SDR enzymes.[172] The release of the reaction intermediates implies a bifurcated mechanism in which the production of UX4O and NADH competes, albeit weakly, with the synthesis of UDX (Figure 2.4A). More importantly, the bifurcated scheme we propose offers a plausible explanation for the apparent stimulatory effect of exogenous NAD^+. In the absence of exogenous NAD^+, the hUXS:NADH:UX4O complex will be in dynamic equilibrium with apo-hUXS (Figure 2.4A). Thus, we expect a small fraction of

hUXS to be losing the reduced cofactor during turnover, decreasing the apparent activity of the enzyme. In fact, we observe that at low protein concentrations, hUXS is unstable and loses activity (unpublished observations). Therefore, the apparent increase in activity in the presence of exogenous NAD^+ could be due to the rescue of apo-hUXS from an otherwise abortive catalytic cycle.

ArnA conserves the bifurcated mechanism and substrate cooperativity of UXS.

In ArnA, UDX production displays a lag in activity that is well modeled by Frieden's equation for transient kinetics (Figure 2.5B).[163] However, we do not believe the production of UDX by ArnA involves enzyme hysteresis. Instead, we note that the half-life of the lag coincides with high levels of the products NADH and UX4O and depletion of UGA (Figure 2.5C). We also observe that the production of UDX is reflected in the depletion of NADH and UX4O. This suggests that the released products NADH and UX4O rebind to the ArnA active site to produce UDX. This is similar to a recent report of an ArnA-like protein that also rebinds NADH and UX4O to produce UDX.[28] The lag is likely due to the low concentrations of NADH and UX4O early in the progress curves, and the competition with UGA for the active site. The progress curves of NADH and UX4O production do not display a lag, effectively ruling out enzyme hysteresis as the source of the lag.

Our results show that the catalytic machinery for UDX synthesis is already present in ArnA. We believe that a fundamental difference between hUXS and ArnA is the equilibrium between E:NADH:UX4O and E* in the bifurcated mechanism, which acts as a catalytic shunt to favor one product over another (Figure 2.4A). In hUXS, the production of UX4O and NADH is ~2 orders of magnitude slower than UDX synthesis, indicating that the shunt favors the retention

of the intermediates. In contrast, ArnA quickly releases UX4O and NADH, followed by a slower rate of UDX formation. Our data shows that as the products UX4O and NADH accumulate in the ArnA reaction, E* is driven to the E:NADH:UX4O complex to produce UDX (Figures 4A, 5C). Thus, the equilibrium shunt favors the E:NADH:UX4O complex in hUXS, and the E* state in ArnA.

Gatzeva-Topalova *et al.* have proposed that the function of the hUXS $Loop_{120\text{-}124}$ is to retain NADH, and its absence in ArnA suggests a mechanism by which NADH could be released during the catalytic cycle.[24; 25] Consistent with this proposal, truncation of the equivalent loop in UDP-galactose 4-epimerase (GalE) results in an E* form of GalE that reversibly binds NAD^+ during catalysis.[138] Our detailed structural analysis suggests that $Loop_{120\text{-}124}$ is but one small part of the catalytic shunt (Figures 4D, 7C). We have identified several structural elements that also contribute to the induced fit binding of cofactor and substrate (Figures 7, 8). We believe that small changes in any of these elements are likely to influence the equilibrium between E:NADH:UX4O and E* in the bifurcated mechanisms of hUXS and ArnA.

Additional evidence that hUXS and ArnA share the same bifurcated mechanism comes from the cooperative binding of the cofactor in both enzymes (Figure 2.4). The observed cooperativity means that the conformational changes that occur upon binding of one NAD^+ molecule to the E* complex are reflected in the structure of the other subunit in the dimer. Our analysis of hUXS crystal structures indicates that significant local unfolding of the shunt is required to expose the active site and bind substrate (Figure 2.7C). Analysis of ArnA structures, which capture separate ArnA* and NAD^+-mimic bound conformations, reveal the structural elements associated with the induced fit conformational change that accompanies cofactor binding (Figure 2.8). Specifically, $Helix_{396\text{-}400}$ ($Helix_{165\text{-}168}$ in hXS) is positioned to interact with

the N-termini of both dimer interface helices and might propagate cooperative binding in ArnA*. A crystal structure of hUXS* is currently unavailable. However, conservation of the cooperative conformational change in hUXS and ArnA suggests a common mechanism for communication between subunits using equivalent structural elements.

Acknowledgements

We thank Dr. Marcelo Sousa for the ArnA pMS159 construct. We also thank Dr. Maor Bar-Peled for the gift of the UDP-4-keto-xylose standard. The authors would additionally like to thank Dr. Debra Mohnen and the CarboSource Services facility for help with HPLC purification. This work was supported in part by the Integrated Technology Resource for Biomedical Glycomics (P41RR018502, L.W. senior investigator). Funding from the American Cancer Society Research Scholar Grant (RSG0918401DMC) and the University of Georgia Research Alliance to Z.A.W. is gratefully acknowledged.

Data Deposition

The atomic coordinates and structure factors for the open form of hUXS have been deposited in the Protein Data Bank (PDB ID code 4GLL).

Abbreviations

hUXS, human UDP-xylose synthase; UGA, UDP-α-D-glucuronic acid; UX4O, UDP-α-D-4-keto-xylose; UDX, UDP-α-D-xylose; ArnA, *Escherichia coli* UDP-α-D-glucuronic acid decarboxylase; CZE, capillary zone electrophoresis; CID-MS^n, collision induced dissociation mass spectrometry; NSBD, nucleotide sugar binding domain; GalE, UDP-α-D-galactose 4-epimerase

CHAPTER 4

HUMAN UDP-α-D-XYLOSE SYNTHASE FORMS A HIGH ACTIVITY TETRAMER[2]

[2] Polizzi, S.J., Walsh, R.M., Jr., Le Magueres, P., Criswell, A.R. and Wood, Z.A. To be submitted to *Biochemistry*.

Abstract

Human UDP-α-D-xylose synthase (hUXS) decarboxylates UDP-α-D-glucuronic acid to form UDP-α-D-xylose (UDX). UDX is believed to regulate UDP-sugar pools by allosterically inhibiting UXS and UDP-α-D-glucose dehydrogenase (hUGDH). Although hUGDH inhibition involves changes in the hexamer quaternary structure, the biologically relevant hUXS oligomer has not been established. Here we show that perturbation of the hUXS quaternary structure regulates enzyme activity. In solution, hUXS undergoes a dimer-tetramer self-association (K_d = 4.7 μM) that is rapid on the time scale of sedimentation experiments. Reaction kinetics indicate that self-association impacts both paths of the bifurcated hUXS mechanism. Release of the reaction intermediate NADH is stimulated by crowding with increasing human UXS concentrations and can be adequately modeled by a binding curve using the orthogonally derived tetramer K_d. At an enzyme concentration ~1/10 of the tetramer K_d, UDP-xylose formation is stimulated >10-fold by the crowding agent trimethylamine *N*-oxide in a manner that resembles a binding isotherm. We identify a tetramer in the open hUXS lattice that is not present in the hUXS:UDP complex. Small angle x-ray scattering (SAXS) studies indicate the hUXS quaternary structure in solution is i) distinct from the tetramers reported to form in the related short-chain dehydrogenase/reductase (SDR) enzymes and ii) likely to undergo structural dynamics. We combine our structural and kinetic data to model hUXS regulation *in vivo*.

Introduction

Every cell in the body receives developmental signals and mechanical support from the extracellular matrix during its lifetime.[173] Proteoglycans are a major component of the extracellular matrix and facilitate growth factor presentation[141] and connective tissue integrity.[142] Proteoglycan functions are often mediated by the diversity of glycosaminoglycan chains covalently bound to the core protein.[9; 54] Xylose attachment to the Ser residue of an acceptor protein is the first committed step in glycosaminoglycan linkage to most proteoglycans.[9; 54] UDP-xylose (UDX) and UDP-glucuronic acid (UGA) are the activated donor substrates required for the first and last steps in proteoglycan linker synthesis, respectively.[174; 175] UDX biosynthesis proceeds via the conversion UDP-glucose to UGA by UDP-glucose dehydrogenase (UGDH)[14] and subsequent decarboxylation by UDP-xylose synthase (UXS).[7] UDX is believed to be the master regulator of UDP-sugar pools via allosteric inhibition of UGDH and UXS.[18] Recently, UXS activity was directly linked to proper proteoglycan maturation and tissue stability during the earliest stages of vertebrate development.[7]

UXS is a member of the short chain dehydrogenase/reductase (SDR) family of enzymes. SDR enzymes share a common three dimensional structure containing a classical Rossmann fold for NAD(P)(H) cofactor binding and the catalytic residues Ser (or Thr), Tyr-XXX-Lys.[150] UXS contains an additional ~100 amino acids organized into a substrate binding domain and can be further classified as an 'extended' SDR enzyme.[72] The extended SDR subfamily is dominated by nucleotide sugar modifying enzymes catalyzing epimerization, dehydration, and decarboxylation reactions.[73] Catalysis is believed to share several common steps, summarized as: 1) tyrosine mediated proton abstraction from the C-4 hydroxyl, 2) hydride transfer from the substrate C-4 to the NAD^+ cofactor, 3) reorientation of the C-4-keto-sugar within the active site, and 4) hydride

transfer from the NADH cofactor back to the substrate to regenerate NAD^+.[44; 78] While UDP-galactose epimerase and dTDP-glucose 4,6-dehydratase rapidly catalyze all four steps with transient intermediates,[105; 114] UGA-decarboxylases can either release NADH and UDP-4-keto-xylose (UX4O) or catalyze a second hydride transfer to form UDX.[23; 28; 176]

Extended SDR enzymes have been structurally well characterized. A human UXS construct lacking the N-terminal 84 residues (hUXS) has been crystallized in substrate free[176] and UDP bound conformations.[44] In each conformation, hUXS forms a canonical SDR dimer within the asymmetric unit[176] or with a symmetry related molecule, respectively.[44] Typical of SDR dimers, each monomer contributes two alpha helices to the oligomer interface. The hUXS dimer interface is restricted to the catalytic domains and related by a two-fold rotation axis perpendicular to the four-helix bundle.[73] The SDR dimers formed by GDP-mannose 4,6-dehydratase and CDP-D-glucose 4,6-dehydratase have also been crystallized in a tetrameric form with near 222 symmetry[111; 125; 130; 131] but solution studies have not indicated an altered role for the larger complexes.

The aim of this paper is to identify the biologically relevant oligomeric state of hUXS. Here we show that hUXS forms a high activity tetramer in solution. We derive the K_d for the hUXS dimer-tetramer association system and identify a novel, asymmetric tetramer in the unliganded hUXS crystal structure. We also present a small-angle x-ray scattering (SAXS) structure that is consistent with formation of a more symmetric, extended tetramer in solution. Supporting the biological relevance of the tetramer, we show crowding conditions stimulate UDX production and the orthogonally measured K_d describes NADH release. In conclusion, we present a model of hUXS regulation.

Materials and Methods

Sedimentation Velocity Analytical Ultracentrifugation (AUC)

hUXS was purified as previously reported[7] and concentrated to 790 μM (ε_{260} = 37485 M^{-1} cm^{-1}) in storage buffer containing 0.2 M NaCl and 25 mM Tris pH 8.0. For sedimentation velocity experiments, hUXS was diluted to 53 μM and dialyzed into PBS buffer pH 7.5 using a Microdialyzer System (Pierce), then diluted with PBS to the working concentration. 400 μL of diluted sample or reference PBS buffer was loaded into 12mm double-sector Epon centerpieces equipped with quartz windows and equilibrated at rest for approximately 1 hour at 20 º C in an An60 Ti four hole rotor. Experiments were performed using an Optima XLA analytical ultracentrifuge with a rotor speed of 50,000 rpms and a wavelength of 230 nm for 0.5, 1, and 2 μM samples or 280 nm for 3, 13, and 26 μM samples. Data was collected over 6 hour periods using a radial step size of 0.003 cm. The partial specific volume (V-bar) of 0.7353 was calculated from the amino acid sequence. The buffer viscosity of 1.019 cPoise and density of 1.00564 g/ml were calculated using the program SEDNTERP[177] (www.bbri.org/RASMB). All data analysis for sedimentation velocity experiments was performed using the programs SEDFIT and SEDPHAT (www.analyticalultracentrifugation.com). Using SEDFIT, continuous sedimentation coefficient distribution $c(s)$ analyses were restrained by maximum entropy regularization at P = 0.68 confidence interval.[178] The baseline, meniscus, frictional coefficient, systematic time-invariant and radial-invariant noise were all fit. The weight-averaged sedimentation coefficient (s_w) was determined by integrating each $c(s)$ distribution over the range of interacting species 3-8 S, then converted to standard AUC conditions (water at 20 °C) and plotted versus hUXS loading concentration. The resulting isotherm data was globally fit in SEDPHAT (with standard conditions) to an A+A→AA self-association model[179] in order to

derive the minimum and maximum sedimentation coefficient *s* values, and the association constant K_A.

X-ray Crystal and Small-angle X-ray Scattering (SAXS) Structural Analyses

Crystal symmetry mates and structural alignments of open hUXS (PDB entry 4GLL),[176] the hUXS:UDP complex (2B69)[44] and the GDP-mannose 4,6-dehydratase:NADPH:GDP-rhamnose complex (1N7G)[130] were generated in Coot.[158] Buried surface area and interface residues were calculated using PISA[180] (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html). Structural figures were generated with PyMOL (http://www.pymol.org). The hydrodynamic property *s* based on atomic coordinates was calculated with HYDROPRO[181] (www.bbri.org/RASMB) using standard AUC conditions for direct comparison to the SEDPHAT isotherm.

For SAXS experiments, hUXS was exchanged into Tris buffer (200 mM NaCl, 25 mM Tris pH 8.0) or phosphate buffer (200 mM NaCl, 50 mM phosphate pH 7.5) with or without 5% glycerol, and diluted to the working concentration. Sample and buffer data were collected at 4 °C in 2, 5 and/or 30 min exposures using a BioSAXS-1000 instrument (Rigaku) with a microfocus rotating anode FR-E+ x-ray source (1.54Å wavelength) and a Pilatus 100K hybrid pixel array detector (DECTRIS). Sample data was corrected for buffer scattering with PRIMUS[182] and the experimental radius of gyration (R_g) was computed (for $sR_g < 1.3$) using the AUTORG[183] tool. Corrected data were used to calculate the pair distribution function using GNOM.[184] A minimum of 10 *ab initio* molecular envelopes from each data set were generated with DAMMIN[185] and averaged using DAMAVER.[186] The quaternary structure was modeled by simulated annealing of two hUXS dimers to scattering data with SASREF,[187] followed by normal mode analysis using

the foXS server (http://modbase.compbio.ucsf.edu/foxs).[188] Hypothetical scattering profiles and R_g from atomic coordinates were fit with CRYSOL[189] and assessed by chi analysis. Structures were docked with *ab initio* envelopes using Chimera[190] (www.cgl.ucsf.edu/chimera).

Enzyme Assays

Standard reactions, unless otherwise stated, contained 50 mM Tris pH 8.0, 10 mM DTT (Research Products International), 1 mM EDTA, and 2 mM UDP-glucuronic acid (UGA; Sigma-Aldrich). All reactions were conducted at 25 °C and after a 5 min pre-incubation, were initiated by addition of a small volume of substrate. UDX formation was measured in standard reactions additionally containing 0.3 μM hUXS and 0.25-1.5 M trimethylamine *N*-oxide (TMAO; Sigma-Aldrich). Samples were taken at regular time intervals and snap frozen in liquid nitrogen, vortexed with 1:1 v/v chloroform to thaw, centrifuged to obtain the aqueous phase, boiled for 5 minutes and re-extracted with 1:1 v/v chloroform. The aqueous layer was loaded to a capillary electrophoresis system (Agilent Technologies) in an uncoated, fused-silica capillary (560 mm x 50 μm) with an extended light path (Agilent Technologies) at 18 °C, separated via capillary zone electrophoresis in 20 mM borate pH 9.3 run buffer for 20 min at 22 kV capillary voltage, and monitored at 260nm wavelength light. Absorbance chromatograms were analyzed with Agilent CE ChemStation and UDX production was confirmed by comigration with known standards. UDX peak area was manually integrated and converted to molar concentration using a standard curve of known UDP-glucose concentrations and the molar extinction coefficient of UDP (9820 M^{-1} cm^{-1}) at 260nm. NADH formation was measured in standard reactions additionally containing 2.5 mM NAD^+ (Sigma-Aldrich) and varying concentrations of hUXS. Samples were placed in a 1 cm path length quartz cuvette and absorbance at 340 nm was monitored at regular

time intervals for 10 minutes using a UV-Vis Spectrophotometer (Agilent Technologies). Absorbance was converted to concentration using the molar extinction coefficient of NADH (ε_{340} = 6220 M^{-1} cm^{-1}). Initial velocities for UDX and NADH production were calculated by linear regression and corrected for enzyme concentration to obtain specific activity (SA). Binding kinetics were modeled with Prism (GraphPad Software, Inc) the equation:

$$y = SA_{min} + ((SA_{max} - SA_{min})*x)/(K_d + x)$$

Results

Human UXS is an association-dissociation system in solution.

In order to determine the oligomeric state of human UXS, we expressed an enzyme construct lacking the N-terminal membrane-spanning domain (hUXS; residues 85-420) as previously reported.[7] Using sedimentation velocity analytical ultracentrifugation (AUC) we assayed hUXS for self-association (Figure 3.1A). At a concentration of 0.5 μM hUXS, the sedimentation coefficient distribution $c(s)$ shows a broad protein peak with a maximum at 4.6 svedburgs (S). Since the profile indicates tailing at sedimentation coefficient (s) values corresponding to larger complexes, we chose to examine higher protein concentrations. At 2 and 3 μM hUXS, sedimentation is modeled as either a mixture of two species with defined maxima at 4.6 S and 6 S, or a single maxima at 6 S with a broad profile tailing toward lower s values, respectively (Figure 3.1A). At the highest concentration tested, 26 μM, hUXS sediments as a more symmetrical peak at 7 S that is continuous with a slower sedimenting species at 5 S (Figure 3.1A). In order to calculate a dissociation constant (K_d) for the conversion between oligomeric species, we constructed a binding isotherm by plotting the weight averaged sedimentation coefficient (s_w) from each $c(s)$ distribution as a function of enzyme concentration (Figure 3.1B).

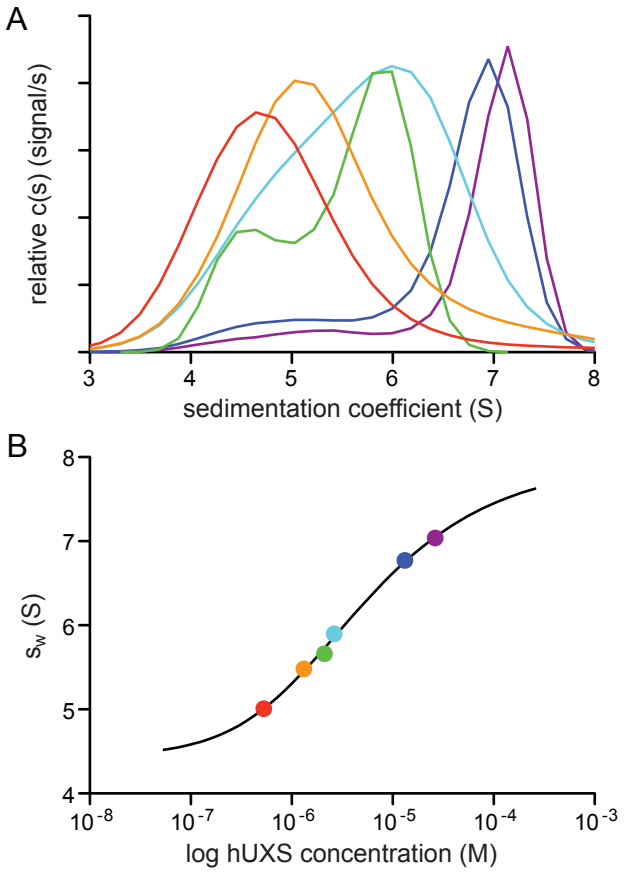


Figure 3.1. hUXS forms a tetramer in solution. (A) Overlay of *c*(*s*) distributions calculated from sedimentation velocity experiments at varying hUXS concentrations using SEDFIT. Absorbance is reported at 230 nm (0.5 μM hUXS, red; 1 μM, orange; 2 μM, green) or 280 nm wavelength light (3 μM, cyan; 13 μM, blue; 26 μM, purple). Distribution ranges (3-8 S) covering the hUXS dimer and tetramer are scaled for ease of viewing. (B) Plot of weight averaged *s* values calculated by integrating each *c*(*s*) distribution (above) and reported under standard AUC conditions. Data are modeled to an A+A→AA self-association binding isotherm using SEDPHAT.[191]

We modeled an A+A→AA self-association system to the binding isotherm data. The resulting model converges at best-fit values for the *s* minimum (s_{min} = 4.4 S), *s* maximum (s_{max} = 7.9 S) and K_d ($1/K_A$ = 4.7 μM) (Table 3.1).

hUXS assembles into a dimer and a tetramer.

The 'open' hUXS crystal structure (PDB entry 4GLL)[176] shows a dimer in the asymmetric unit with each monomer contributing ~1500Å^2 of solvent accessible surface area to the interface (Figure 3.2A). The hUXS dimer represents the canonical quaternary structure of extended SDR enzymes and has previously been characterized.[73; 176] Anticipating a monomer-dimer association, we generated a hydration shell from the unliganded hUXS atomic coordinates and calculated apparent *s* values (Table 3.1). The apparent *s* values of the monomer (3.3 S) and dimer (5.0 S) are less than the isotherm s_{min} and s_{max}, respectively. Instead, the apparent *s* of the dimer is more consistent with the modeled s_{min}, and suggests a tetramer is the dominant oligomeric species above 13 μM hUXS (Figure 3.1).

We examined packing interactions within the open hUXS unit cell for evidence of higher order oligomers. We located a tetramer formed by dimers rotated 120° and packed into a fiber along the crystallographic 3_2 axis (Figure 3.2A,B). The tetramer resembles a butterfly with two upper wings consisting of chains *A* and *B′* from different dimers, and two lower wings composed of the remaining chains *B* and *A′* (Figure 3.2C). The upper wings extend in opposite directions due to the overall twist of the oligomer (Figure 3.2C). Adjacent dimers create an interface with ~4200Å^2 of buried surface area, mainly between the lower wings. Unlike the canonical SDR dimer interface, the tetramer-building interface contains asymmetrical contacts due to a pseudo two-fold axis bisecting the complex (Figure 3.2C). Asymmetry is most prominent where the

Table 3.1. Hydrodynamic properties of the hUXS oligomers

Oligomer [a]	S_{calc} [b]	S_{exp} [c]	K_d [c]
monomer	3.32	---	---
dimer	5.01	4.44	---
asymmetric tetramer	8.05	7.94	4.68

[a] Oligomeric state present in the open hUXS crystal structure (PDB entry 4GLL).

[b] Sedimentation coefficient (*s*) values were calculated with HYDROPRO[181] using atomic coordinates and standard AUC conditions and are reported in svedbergs.

[c] *s* values from the experimental isotherm were modeled with SEDPHAT[191] and assigned to the dimer-tetramer association. *s* values were calculated under standard AUC conditions for direct comparison with S_{calc}. K_d is reported in μM.

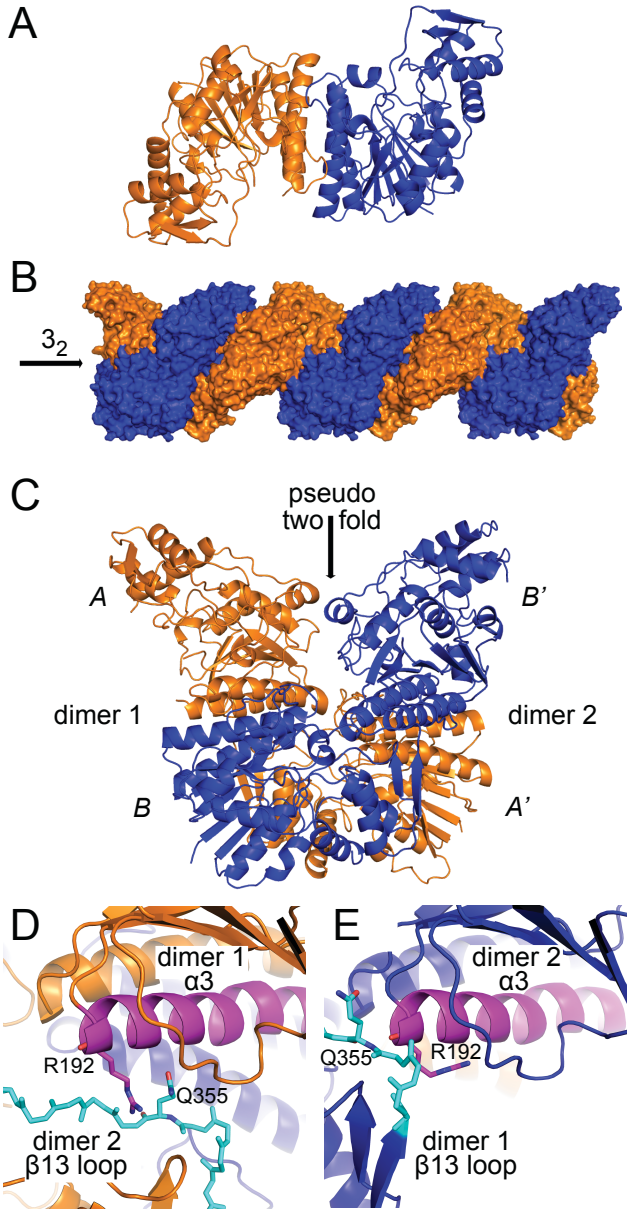


Figure 3.2. hUXS forms a crystallographic fiber of asymmetric tetramers. (A) Non-equivalent hUXS monomers (orange and blue; PDB entry 4GLL) associate along a four alpha helix bundle to form typical SDR dimer interfaces. (B) Dimers in the unit cell rotate by 120° and pack along the crystallographic 3_2 axis to form an ordered fiber. (C) Adjacent dimers form a tetramer with asymmetric contacts due to a pseudo two-fold axis. (D) Close up view of overlapping dimer and tetramer interfaces highlights asymmetric interactions at the α3 helix (residues 172-193, magenta). Gln355 of the β13 loop (residues 351-365, cyan sticks) packs against the side of the α3 helix at one contact and (E) moves 17Å to pack against the C-terminal Arg192 (magenta sticks) of the α3 helix at the analogous contact.

tetramer interface overlaps the four-helix bundle of the SDR dimer. In one overlapping region, the α3 helix (Pro172-Val193) from dimer 1 is packed against the β13 loop (Lue351-Ile365) from dimer 2 (Figure 3.2D). In the analogous overlapping region, the α3 helix from dimer 2 and the β13 loop from dimer 1 have rotated relative to one another (Figure 3.2E). As a result, Gln355 of the β13 loop breaks side chain interactions with α3, adopts an asymmetrical loop conformation roughly 17Å away, and forms a main chain interaction with Arg192 at the C-terminus of the α3 helix (Figure 3.2E).

Small-angle x-ray scattering analysis of the hUXS tetramer.

Since the 'asymmetric' tetramer we identified in the hUXS crystal structure assembled into a fiber and our AUC experiments did not indicate fiber formation in solution, we further probed the molecular structure of hUXS using small angle x-ray scattering (SAXS). In order to assay hUXS concentrations ranging between 26-260 μM (equivalent to 1-10mg/ml), we created a 790 μM stock enzyme solution. During sample preparation, we observed that placing concentrated hUXS on ice caused visible precipitation of the enzyme, which could be reversed by warming the protein aliquots in hand (data not shown). Therefore, we first addressed hUXS aggregation and possible fiber formation during SAXS experiments. We found that the hUXS radius of gyration (R_g) increases with X-ray exposure and enzyme concentration (Table 3.2), consistent with aggregation.[192] However, addition of 5% glycerol to hUXS samples consistently lowers the experimental Guinier R_g by 8-10Å. By reducing the enzyme concentration to 85 μM and adding 5% glycerol, we were able to generate solution scattering data with a linear Guinier plot (Figure 3.3), suggesting decreased aggregation. The resultant scattering profile and R_g of

Table 3.2. Concentration and X-ray dose dependence of hUXS R_g during SAXS

		R_g after X-ray exposure [b]		
Enzyme [a]	[E] μM	30 min	5 min	2 min
hUXS	340	57.9 (0.4)	55.4 (1.2)	51.3 (2.2)
	170	51.9 (0.7)	---	---
	85	51.6 (0.9)	---	---
hUXS + glycerol	340	49.4 (0.3)	46.5 (0.9)	46.4 (1.5)
	170	43.0 (1.0)	---	---
	85	41.7 (0.7)	---	---

[a] hUXS was buffered with 200mM NaCl and 50mM Phosphate pH 7.5 with or without 5% glycerol.

[b] Experimental R_g and error (in parenthesis) calculated from the Gunier plot for $sR_g < 1.3$ using AUTORG.[183] Samples were maintained at 4°C in a BioSAXS-1000 instrument and exposed to a microfocus rotating anode FR-E+ X-ray source (1.54Å wavelength).

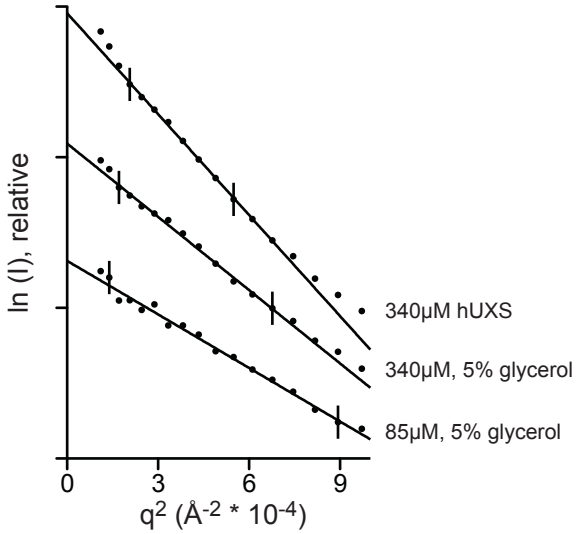


Figure 3.3. Gunier analysis of hUXS preparations for SAXS. Scattering data for samples containing 340 μM hUXS without glycerol (top), 340 μM with 5% glycerol (middle) and 85 μM with 5% glycerol (bottom) are scaled for ease of comparison. The Gunier region used for calculating the radius of gyration ($sR_g < 1.3$) is denoted by hash marks. The corresponding Gunier region slope is graphed for the extended data range to illustrate agreement with experimental data.

hUXS are more consistent with predictions from the tetramer atomic coordinates (χ = 4.1, R_g = 35.7Å) than the dimer (χ = 11.3, R_g = 30Å; Table 3.2, Figure 3.4). Based on the scattering profile, the molecular weight of hUXS in solution is estimated to be 169.6 kDa. A similar value is derived from the amino acid composition of the tetramer (4*37.9 = 151.6 kDa), but not extended fibers. However, the theoretical scatting profile of the asymmetric tetramer does not fully satisfy the hUXS conformation in solution and deviations occur at $0.08 < q < 0.13$ (Figure 3.4).

Since the hUXS crystal structure contains an asymmetric oligomer, we asked whether a more symmetric model could account for inconsistencies between the crystal and solution states. We modeled the hUXS tetramer to the scattering data using rigid body refinement of two dimers (Figure 3.4). In order to allow rearrangement of the crystallographic tetramer during simulated annealing, we applied distance restraints to the asymmetric interfaces: i) 10Å limit between Asn185 from the α3 helix of dimer 1 and Glu353 from the β13 loop of dimer 2, ii) analogous 10Å limit between Asn185 from the α3 helix of dimer 2 and Glu353 from the β13 loop of dimer 1, and iii) 20Å limit between the 'thumb' helices (Pro165-Asn171) of the lower wings B and A′. The refined model contains no main chain clashes and minimal side chain overlaps that could be resolved with alternate rotomers. The model is based upon the observed crystallographic tetramer and agrees more closely with the solution scattering data (χ = 2.2, Figure 3.4). Dimers 1 and 2 shift about the pseudo two-fold axis and form a more 'symmetrical' tetramer along the α3 helix:β13 loop packing surface. Surprisingly, the interface displays charge complimentary and is predicted to contain 7 more salt bridges than occur in the asymmetric tetramer (Table 3.3). Compared to the asymmetric tetramer, the upper wings of the symmetric model shift in opposite directions and cause a slight reduction in total buried surface area (~3750Å^2). Normal mode analysis of the symmetric tetramer reveals that further motion of the upper wings is favorable in

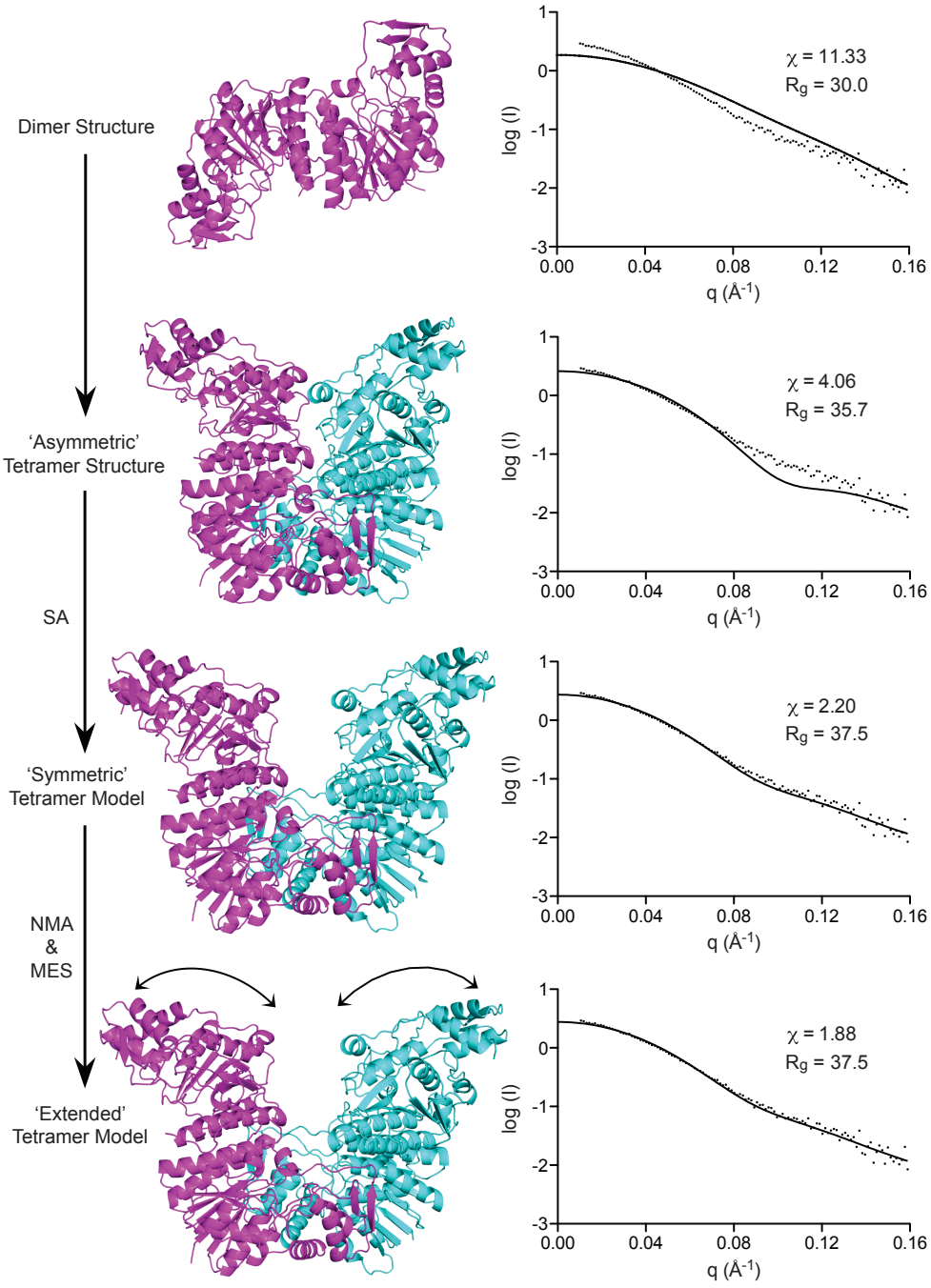


Figure 3.4. The hUXS solution state is consistent with an extended tetramer conformation. Refinement of the hUXS oligomeric state to SAXS data is depicted as a flow chart (straight arrows). Atomic coordinates present in the hUXS crystal structure (dimer, 'asymmetric' tetramer) or refined with simulated annealing (SA, 'symmetric' tetramer), normal mode analysis (NMA) and a minimal ensemble search (MES, 'extended' tetramer) are depicted in cartoon with dimers colored magenta and cyan. Curved arrows over the conformation selected by MES indicate dynamics observed in the lowest mode of NMA. A theoretical scattering profile for each structure (solid curves) and the corresponding radius of gyration (R_g) are modeled to experimental SAXS data using CRYSOL[189] and the quality of fit is derived by chi (χ) analysis.

Table 3.3. Interactions at the asymmetric and symmetric hUXS tetramer interfaces

Tetramer Form	Interface Interactions [a]	Dimer1 Chain	Residue [Atom]	Distance (Å)	Dimer2 Chain	Residue [Atom]
Asymmetric	Hydrogen Bonds					
	1	*A*	TYR 151 [OH]	2.63	*A′*	GLN 355 [NE2]
	2	*A*	LYS 191 [NZ]	3.23	*A′*	ASP 208 [O]
	3	*A*	ARG 192 [NH2]	2.26	*A′*	ASP 356 [O]
	4	*A*	LYS 247 [O]	3.07	*A′*	ASN 223 [ND2]
	5	*A*	ASN 316 [O]	3.85	*B′*	SER 215 [OG]
	6	*A*	ASN 316 [ND2]	3.31	*B′*	ASP 364 [OD2]
	7	*B*	GLU 147 [OE2]	2.48	*A′*	TYR 290 [OH]
	8	*B*	TYR 170 [O]	2.84	*A′*	ARG 228 [NE]
	9	*B*	ARG 228 [NH1]	3.32	*B′*	LYS 247 [O]
	10	*B*	GLN 355 [N]	2.85	*B′*	ARG 192 [O]
	Salt Bridges					
	1	*A*	LYS 191 [NZ]	3.23	*A′*	ASP 208 [OD2]
	2	*A*	GLU 249 [OE1]	2.95	*A′*	LYS 360 [NZ]
Symmetric	Hydrogen Bonds					
	1	*A*	GLU 249 [OE1]	3.82	*A′*	ASP 356 [N]
	2	*B*	ARG 125 [NH2]	3.66	*A′*	GLU 394 [O]
	3	*B*	ARG 125 [NH2]	3.54	*A′*	GLU 394 [OE2]
	4	*B*	LYS 126 [NZ]	2.95	*A′*	LEU 282 [O]
	5	*B*	ASN 128 [N]	3.34	*A′*	GLN 398 [O]
	6	*B*	GLU 130 [OE1]	3.78	*A′*	TYR 397 [N]
	7	*B*	ASN 142 [ND2]	2.98	*A′*	GLN 280 [O]
	8	*B*	HIS 143 [N]	2.00	*A′*	GLU 285 [OE1]
	9	*B*	ASN 171 [NZ]	3.73	*A′*	TYR 290 [OH]
	10	*B*	LYS 174 [NZ]	3.07	*A′*	TYR 290 [OH]
	11	*B*	GLN 283 [O]	3.80	*A′*	PHE 121 [N]
	12	*B*	GLN 283 [O]	3.60	*A′*	HIS 143 [N]
	13	*B*	GLU 285 [N]	2.37	*A′*	ASN 142 [OD1]
	14	*B*	LYS 393 [O]	3.55	*A′*	TYR 397 [OH]
	15	*B*	GLU 394 [OE1]	2.81	*A′*	ARG 125 [NH1]
	Salt Bridges					
	1	*B*	GLU 353 [OE2]	3.09	*B′*	ARG 192 [NE]
	2	*B*	GLU 353 [OE1]	3.92	*B′*	ARG 192 [NE]
	3	*B*	GLU 353 [OE2]	2.88	*B′*	ARG 192 [NH1]
	4	*B*	GLU 353 [OE1]	2.47	*B′*	ARG 192 [NH1]
	5	*B*	GLU 394 [OE1]	2.81	*A′*	ARG 125 [NH1]
	6	*B*	ARG 125 [NH2]	3.54	*A′*	GLU 394 [OE1]
	7	*A*	ARG 192 [NH2]	3.71	*A′*	GLU 353 [OE1]
	8	*A*	ARG 192 [NE]	2.37	*A′*	GLU 353 [OE2]
	9	*A*	ARG 192 [NH1]	3.50	*A′*	GLU 353 [OE2]

[a] Interface contacts and residue distances predicted using PISA[180] based on atomic coordinates for the asymmetric tetramer found in the hUXS crystal structure or the symmetric tetramer refined by simulated annealing to SAXS data with SASREF.[187]

solution (Movie S1 of the Supporting Information). In order to determine the most likely solution state of hUXS, we searched for a minimal ensemble of conformations that could explain the scattering data using the asymmetric tetramer, the symmetric tetramer, and the conformational trajectories from the two lowest normal modes. In each case, a minimal ensemble search chose a single conformation from the lowest normal mode corresponding to an 'extended' tetramer with separated upper wings. The extended tetramer coordinates show the highest agreement with experimental scattering data ($\chi = 1.9$, $R_g = 37.5$Å; Figure 3.4) and are not consistent with the tetrameric form of nucleotide sugar dehydratases **(**Figure 3.5).

We generated an average *ab initio* molecular envelope for hUXS without symmetry constraints and docked the dynamic tetramer model (Figure 3.6). Overall, the shape of the envelope accommodates the hUXS tetramer. Regions of the envelope extending beyond the docked crystal structure are observed for samples buffered with pH 7.5 phosphate (Figure 3.6) and pH 8.0 Tris (data not shown). Unmodeled regions of the envelope vary between samples containing 85μM (Figure 3.6A) or 170μM hUXS (Figure 3.6B) and may be the result of aggregates[193] remaining in the presence of 5% glycerol.

hUXS activity increases under crowded conditions.

In order to determine the effect of self-association on hUXS activity, we monitored the concentration dependence of UGA conversion to UDX via capillary zone electrophoresis. We were unable to accurately measure the linear rate of UDX production over the hUXS dilution range used in AUC experiments (0.5-26 μM) due to rapid catalysis at high protein concentrations (data not shown). Instead, we fixed the hUXS concentration at the lower end of the AUC isotherm (0.3μM, Figure 3.1B) and assayed UDX production with increasing amounts of the

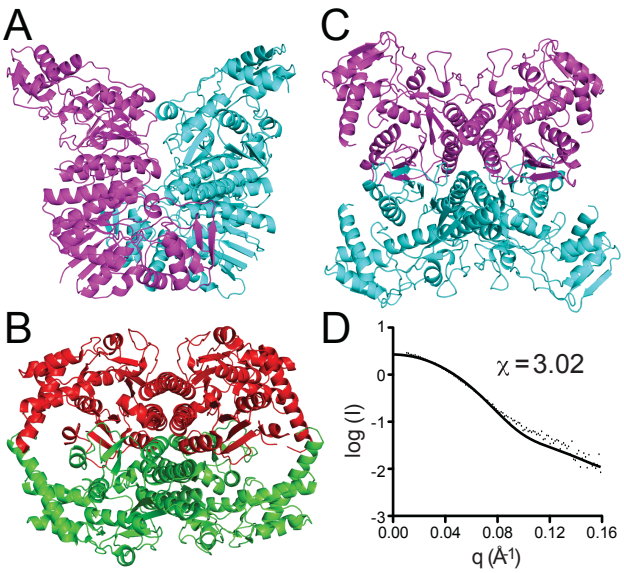


Figure 3.5. The hUXS tetramer is novel among nucleotide sugar modifying SDR enzymes. (A) Asymmetric hUXS tetramer reproduced from Figure 3.4 showing the four-helix bundles from each dimer are approximately aligned end to end. (B) Arabidopsis GDP-mannose 4,6-dehydratase crystal structure (PDB entry 1N7G) showing the four-helix bundles from each dimer (red, green) associate in parallel to form a distinct tetramer. (C) hUXS dimers modeled by superposition with the tetramer from (B). (D) The theoretical scattering profile (solid curve) for the hUXS model in (C) is fit to experimental SAXS data as in Figure 3.4.

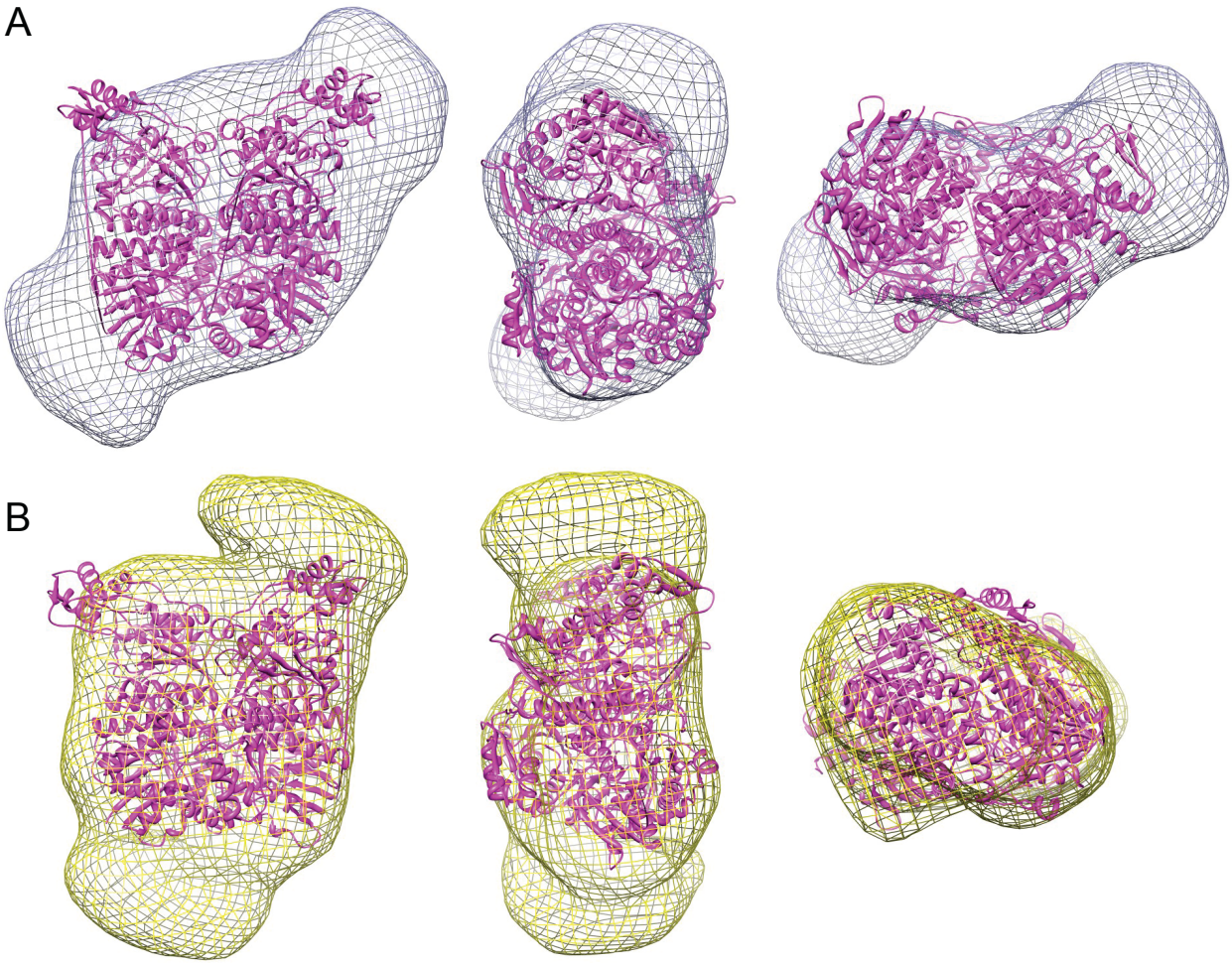


Figure 3.6. hUXS SAXS envelopes are consistent with a tetrameric solution state. (A) *Ab initio* molecular envelopes from samples containing 85μM (grey wire mesh) or (B) 170μM hUXS (yellow wire mesh) generated using DAMAVER.[186] Envelopes are docked with the crystallographic tetramer (magenta) and shown in several orientations to highlight agreement with the overall dimensions of a tetramer rather than a dimer. Extensions beyond the common core envelope vary in shape and orientation between enzyme concentrations and are possibly a consequence of aggregation behavior[193] remaining in the presence of 5% glycerol.

crowding agent trimethylamine *N*-oxide (TMAO). Crowding agents have been shown to increase effective protein concentrations, thereby stimulating native protein association[194] and association based catalytic activities.[195] We observe a rapid increase in UDX production over a 0.5-1 M TMAO concentration range, followed by a plateau in activity (Figure 3.7A). hUXS activity responds to TMAO in a manner resembling a binding isotherm and results in a greater than 10-fold increase in activity under crowded conditions.

We further examined the concentration dependence of hUXS activity by assaying NADH release with UV spectrometry. As previously reported, NADH is released as a minor product compared to UDX[176] and therefore can be measured at higher hUXS concentrations. As enzyme concentration increases, we observe an increase in hUXS specific activity (Figure 3.7B). A similar effect occurs in both Tris pH 8 and phosphate pH 7.5 buffered reactions. To determine whether the change in specific activity is explained by hUXS self-association, we modeled a binding curve using the orthogonally derived K_d from AUC (4.7μM, Table 3.1). The resulting model adequately describes the data and predicts a 10-20% increase in NADH production under Tris pH 8 or phosphate pH 7.5 reaction conditions, respectively.

Discussion

The active form of hUXS is a novel tetramer.

hUXS has recently been shown to release the major catalytic product UDX and, to a lesser extent, the intermediates NADH and UX4O.[176] We report that both hUXS specific activities are concentration dependent (Figure 3.7). UDX and NADH release increase with crowding conditions before plateauing at maximum activities. UDX formation increases by over 10-fold when crowded with the osmolyte TMAO (Figure 3.7A). NADH release, however, only

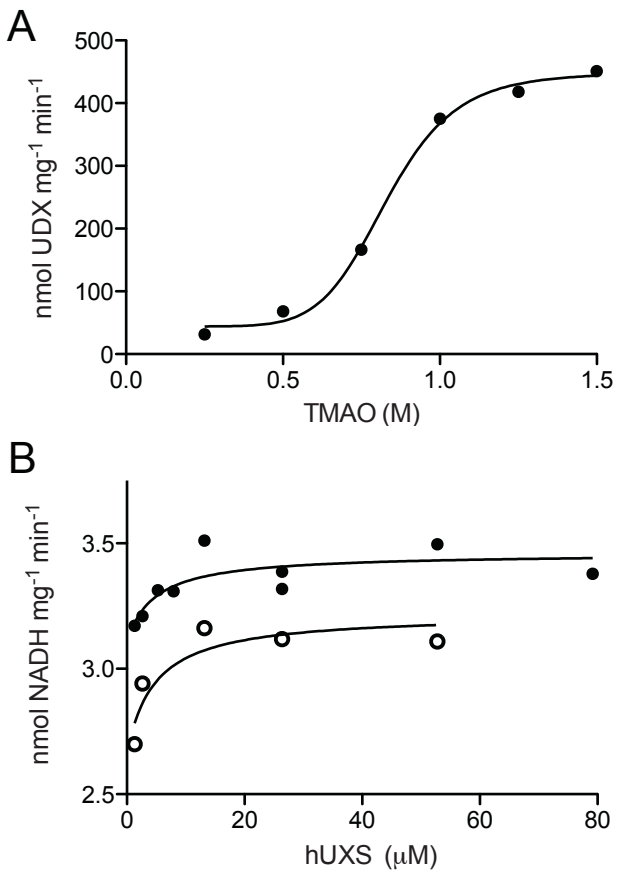


Figure 3.7. The hUXS tetramer is the active species. (A) UDP-xylose (UDX) production is shown as a function of crowding in reactions containing 0.26 μM hUXS and increasing trimethylamine *N*-oxide (TMAO) concentrations. UDX specific activity data are modeled by a binding isotherm for illustrative purposes only. (B) NADH release is shown as a function of crowding in reactions with increasing hUXS concentrations. NADH specific activity in Tris pH 8.0 (●) and phosphate pH 7.5 (o) are modeled by a binding curve using the orthogonally derived K_d for self-association (4.68 μM).

increases by 10-20% when crowded by increasing hUXS concentrations (Figure 3.7B). NADH release was proposed to occur via conformational changes in the intermediate complex.[176] It is possible that crowding conditions limit the conformational freedom required for NADH release, thus attenuating the large increase in specific activity observed for UDX. Alternatively, the hUXS oligomeric states may have similar affinities for substrate binding and oxidation, but varying propensities for reduction and product release. Despite differences in NADH and UDX activation, both kinetic trends resemble binding curves (Figure 3.7) and suggest hUXS exhibits self-association behavior.

In support of the inferred binding behavior, we show hUXS is a rapid association-dissociation system in solution (Figure 3.1). At each hUXS concentration assayed by AUC, the resulting $c(s)$ distribution displays asymmetric or connected peaks (Figure 3.1A), consistent with a mixture of self-associating species.[191] The dissociation of hUXS species is rapidly reversible on the time-scale of sedimentation, since the peaks migrate to higher s values when the hUXS concentration is increased (Figure 3.1A).[196; 197] Modeling the s_w value from each $c(s)$ distribution with a binding isotherm (Figure 3.1B) generates a K_d (4.7 μM) for the self-association reaction (Table 3.1). A binding curve based on the K_d for self-association also adequately models the change in NADH specific activity (Figure 3.7B). The observation that orthogonally measured processes share a single K_d provides strong evidence that oligomerization, rather indirect crowding effects, enhances hUXS activity. UDX specific activity is also well modeled by a binding isotherm (Figure 3.7A), and further supports a transition between low activity and high activity oligomers.

The K_d for self association also relates kinetic activation to enzyme complexes sedimenting at 4.4 S and 7.9 S (Figure 3.1B). A hUXS dimer has been crystallized in two

different space groups[44; 176] and is consistent with the 4.4 S, low activity species (Table 3.1). Here, we report the presence of a tetramer in the unliganded hUXS crystal structure (Figure 3.2) that is consistent with the largest oligomer predicted by AUC (Table 3.1, Figure 3.1). However, SAXS data indicate the 'asymmetric' tetramer conformation in the crystal structure is different from the activated tetramer state in solution. Solution scattering is best fit by an 'extended' tetramer model resulting from the conformational dynamics of a 'symmetric' tetramer. The refined tetramer interface shows increased symmetry along the α3 helices and striking charge complementarity (Table 3.3). The symmetric salt bridges are maintained during dynamics and may facilitate the reversible association of dimers. Peripheral electrostatic contacts have been shown to contribute to binding specificity,[198] suggesting hUXS dynamics may select for self association interactions in solution. We interpret differences between the solution and crystal states of hUXS to result from the inherent flexibility of the tetramer and the selective nature of the crystal lattice. *Ab initio* molecular envelopes derived from the SAXS data account for an hUXS homotetramer but are too large for a single dimer structure (Figure 3.6). Collectively, our structural and kinetic analyses of hUXS support the formation of a canonical SDR dimer with low activity and a novel tetramer with high activity. The hUXS tetramer does not resemble the previously crystallized tetramer structures of nucleotide sugar dehydratases (Figure 3.5). To our knowledge, hUXS is the first extended SDR enzyme to have a distinct function ascribed to the tetrameric state.

A possible model of UXS regulation based on oligomeric states.

Based on the observation that tetramer formation activates UDX biosynthesis (Figure 3.7), we propose that modulating the hUXS oligomeric state may serve as a regulation mechanism

(Figure 3.8). In support of this model, we have demonstrated that a soluble hUXS construct forms a rapidly reversible tetramer with moderate binding affinity (K_d ~5μM). Full-length hUXS is not predicted to be soluble, however, but is believed to anchor within the lumen of the endoplasmic reticulum/golgi[16; 64] via an N-terminal transmembrane (TM) domain.[17] Modeling suggests the TM domain limits translational freedom while facilitating dimer and tetramer quaternary changes. Limiting spatial degrees of freedom has been shown to increase reaction rates,[199] and therefore is anticipated to lower the tetramer K_d in full-length hUXS. Further, the lipid entropy model of membrane complex formation predicts that exclusion of the TM domain within the lipid bilayer may enhance self-assembly[200] and decrease the K_d. An obvious consequence of our model is that protein concentration determines whether full-length hUXS adopts a low or high activity state. Assuming membrane association causes differences in the K_d of other UXS tetramers, our model may also explain the observation of both soluble and anchored UXS isoforms in plants.[3; 38; 40; 59] Specifically, membrane bound UXS may form high activity tetramers at concentrations that promote soluble UXS to form low activity dimers. Additional experiments are required to probe the differences between naturally occurring soluble and membrane bound UXS isoforms.

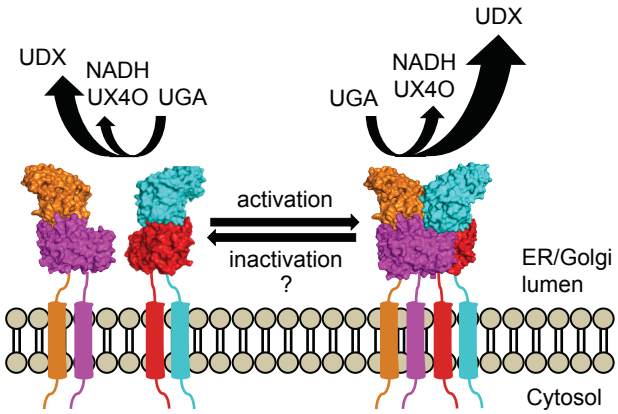


Figure 3.8. A possible model of hUXS regulation. hUXS monomers (orange, magenta, cyan, red) are anchored to the luminal membrane of the endoplasmic reticulum (ER)/golgi by a single transmembrane helix (cylinder). Dimers are shown as the basic oligomeric unit and catalyze the conversion of UGA to the intermediates, NADH and UX4O, or the favored product, UDX. Lateral diffusion within the membrane facilitates association of the asymmetric tetramer, thereby increasing UDX production. Dissociation returns hUXS to the less active state and may be a consequence of enzyme crowding or small molecule binding to an allosteric site in the tetramer interface.

Acknowledgements

We thank Rigaku for generously donating use of their SAXS facility (The Woodlands, TX). We also kindly thank Huaying Zhao (National Institute of Biomedical Imaging and Bioengineering, NIH) and Savvas Savvides (Ghent University) for helpful AUC and SAXS discussions, respectively. Funding from the University of Georgia Research Alliance and the American Cancer Society Grant RSG0918401DMC to Z.A.W. is gratefully acknowledged.

Abbreviations

UGDH, UDP-glucose-dehydrogenase; UXS, UDP-xylose synthase; UDX, UDP-xylose; UGA, UDP-glucuronic acid; SDR, short-chain dehydrogenase/reductase; UX4O, UDP-4-keto-xylose; hUXS, human UXS construct lacking N-terminal 84 amino acids; SAXS, small-angle x-ray scattering; AUC, sedimentation velocity analytical ultracentrifugation; *s*, sedimentation coefficient; $c(s)$, continuous *s* distribution; s_w, weight averaged *s*; TMAO, trimethylamine *N*-oxide; TM, transmembrane

CHAPTER 5

CONCLUSION

I. The hUXS:NAD^+ crystal structure reveals a novel cofactor conformation in UGA decarboxylases.

When we began studying UXS, it was established that members of the extended SDR enzyme family conserve a Gly-rich motif (GXXGXXG) involved in retaining NAD(P)(H) during sequential hydride transfers.[73] It was also generally accepted that three conserved catalytic residues (Ser (or Thr), Tyr-XXX-Lys) facilitate the oxidation and subsequent reduction of nucleotide sugars.[78] However, it remained unclear why the SDR catalytic machinery was conserved in the UGA decarboxylase subfamily, since some members use NAD(H) as cosubstrate to catalyze a single hydride transfer (e.g. ArnA and U4kpxs) while others recycle NAD^+ during a complete oxidation-reduction cycle (e.g. UXS and UAXS) (Table 1.2 and Figure 1.9).

We began to address this ambiguity by recombinantly expressing a soluble, active construct of hUXS (Appendix A)[7] and solving the crystal structure.[176] Consistent with other members of the extended SDR family, hUXS co-purifies and crystallizes with one well-ordered NAD^+ cofactor per enzyme active site (Figures 2.2, 2.6). NAD^+ is coordinated in part by residues Gly95-Gly101, comprising the SDR motif GXXGXXG. NAD^+ makes additional contacts with residues Tyr231 and Lys235, constituting the SDR motif Tyr-XXX-Lys (Figure 2.6). Our structure also indicates that NAD^+ adopts the *syn* conformation expected for class B dehydrogenases.[91] Unexpectedly, however, the χ_N torsion angle between the nicotinamide and

ribose of NAD^+ in hUXS varies from other class B dehydrogenases by 30° or more (unpublished observation). Nicotinamide rotation about the χ_N torsion angle is caused by a packing interaction with the extended side chain of Ile258 in hUXS (Figure 2.6), relative to the shorter Cys188 rotamer found in the equivalent position of *Streptococcus suis* RmlB (unpublished observation). Although a crystal structure of ArnA with NAD(H) bound is currently unavailable, Pro490 provides a packing surface similar to the Ile258 interaction in hUXS (Figure 2.6) and is also predicted to rotate the nicotinamide ring by ~30°.

We have modeled GlcA binding based on the NAD^+ conformation in hUXS (Figure 2.7), rather than the non-isosteric ATP analog in ArnA (Figure 2.6)[25] or the rotated nicotinamide in other class B dehydrogenases (Figure 1.19).[26] Our model is consistent with the geometry of hydride transfer[119; 166; 167; 168] and supported by the ordered water structure of the active site (Figure 2.7).[169; 170; 171] The productive conformation of GlcA in our model facilitates interactions with the catalytic Thr and Tyr residues as anticipated (Figure 2.7). Our modeled GlcA carboxylate does not interact directly with either the conserved Glu204 or Arg361 residues, as previously proposed using a RmlB-based model of ArnA,[26] or the ArnA:UGA:ATP complex,[25] respectively. Instead, Glu204 forms a salt-bridge with Arg361 in both hUXS structures that is maintained during molecular dynamic simulations of the catalytic cycle.[44] The residues forming the Glu-Arg salt bridge in hUXS are conserved in UGA decarboxylases, suggesting the salt bridge may be a defining feature of the decarboxylase active site.[44]

II. Substrate and cofactor induce distinct conformational changes in UGA decarboxylases.

Prior to our work, UGA decarboxylases had only been crystallized in apo-enzyme[24; 26] and ternary complexes.[25; 44] Therefore, our binary hUXS:NAD^+ complex allows us to distinguish

between the conformational changes associated with UGA and NAD^+ binding, respectively, in decarboxylases for the first time. Our comparison of the apo-ArnA and ArnA:ATP:UGA structures reveals three conserved structural elements that change conformations due to NAD^+ (or mimic) binding (Figure 2.8). One element, ArnA $Loop_{347\text{-}349}$, changes conformation to orient Asp347 for interactions with the adenine ribose. The N-terminal region of a second element, ArnA $Helix_{373\text{-}381}$, also changes conformation and facilitates hydrogen bonding between Asp360 and the adenylate. The third element, ArnA $Helix_{326\text{-}337}$, moves the GXXGXXG motif toward the NAD^+ binding site to pack against substrate. Reversible NAD(P)(H) binding has previously been described in related dehydratases, which contain charged residues within the variable (X) regions of the GXXGXXG motif (Figure 1.7).[116; 118] ArnA reversibly binds NAD^+ without a charged amino acid in the GXXGXXG motif,[24] indicating exchange may occur via another mechanism. The short ArnA $Loop_{347\text{-}349}$ increases the solvent accessibility of the catalytic domain, relative to other members of the extended SDR family, and may provide a mechanism to exchange NAD^+ (Figure 1.21).[24] hUXS contains neither charged residues in the GXXGXXG motif (Figure 1.7), nor a shortened NAD^+-binding loop (Figure 2.6), and purifies with a tightly bound NAD^+ in each active site (Figure 2.2). Consistent with the importance of the NAD^+-binding loop, truncation of the extended loop found in *P. calidifontis* GALE to an ArnA-like structure results in reversible NAD^+ binding[138] without the need for charged GXXGXXG residues.

Both hUXS structures contain NAD^+ bound in the catalytic domain and show little deviation in the conserved cofactor interactions, as might be expected. Instead, deviations in the structures can be attributed to sugar substrate binding. Our comparison of the open and closed hUXS structures indicates that binding the UDP moiety of UGA induces a rotation of the NSBD (Figures 1.8). In contrast to the modest rotations induced by substrate in human GALE and

RmlB structures (<14°) (Figures 1.12, 1.13, 1.16), the NSBD of hUXS rotates ~34° to close the active site (Figure 2.7). The large rotation is also accompanied by conformational changes in three dynamic elements located at the junction of the NSBD and catalytic domains of our hUXS:GlcA model (Figure 2.7). The first element, hUXS $Loop_{266-272}$, breaks interactions with the NAD^+ backbone and reorders over the catalytic center formed by the nicotinamide and GlcA moieties. $Loop_{266-272}$ buttresses the second element, $Helix_{165-168}$, which undergoes a disorder to order transition upon substrate binding. A similar substrate induced transition has been described in the equivalent element of RmlB (Figure 1.16).[119] The final hUXS element, $Loop_{159-164}$, is N-terminal to $Helix_{165-168}$ and forms part of the NAD^+ binding site (Figure 2.7). B-factor analysis indicates that ordering $Helix_{165-168}$ stabilizes $Loop_{159-164}$ and the active site architecture.[176] Just as hUXS conserves the interactions induced by NAD^+ binding, ArnA conserves similar structural elements that respond to substrate sugar binding.

III. UGA decarboxylases are bifunctional UDP-4-keto-xylose/UDP-xylose synthases.

Our observation that hUXS and ArnA generally conserve the same catalytic machinery and conformational responses to substrate led us to re-examine the catalytic cycle of both types of UGA decarboxylases. To this end, we recombinantly expressed the decarboxylase domain of *E. coli* ArnA for comparison with hUXS. Consistent with the published division in UGA decarboxylases (Figure 1.9), we confirm that hUXS uses a bound NAD^+ cofactor to form UDX (Figure 2.2) and ArnA uses exogenous NAD^+ to form NADH and UX4O (Figure 2.4). Challenging the reported dichotomy, however, we show that in the presence of NAD^+ and UGA substrates, hUXS exhibits an ArnA-like release of NADH and UX4O after a single hydride

transfer (Figure 2.2). Further blurring the distinction, we show that ArnA will rebind concentrated levels of UX4O and NADH and catalyze UDX formation (Figure 2.5).

The previously uncharacterized activities of UGA decarboxylases cast a new light on the description of U4kpxs as a bifunctional UDP-4-keto-xylose/UDP-xylose synthase.[28] U4kpxs was originally identified as an ArnA enzyme located within the canonical cluster of Lipid A metabolic genes.[27] After CZE analysis of the U4kpxs reaction in our lab revealed that UDX had been synthesized, Gu et al.[28] recharacterized the decarboxylase as a bifunctional UDP-4-keto-xylose/UDP-xylose synthase (personal correspondence). Our observation that the archetypical ArnA and UXS families form both products suggests that all UGA decarboxylases may be inherently bifunctional UDP-4-keto-xylose/UDP-xylose synthases (Figure 2.4A). The reported accumulation of UX4O in NAD^+-dependent reactions catalyzed by UDP-apiose/UDP-xylose synthase[52] further supports the presence of a bifurcated pathway in the three known classes of UGA decarboxylases.

Our finding that the UGA decarboxylases conserve a bifurcated pathway may also provide a reason for the metabolic separation of decarboxylase and formyltranserase steps in full length ArnA. Williams et al. observed that the decarboxylase and formyltransferase activities of full length ArnA are unusual, in that the enzyme ArnB catalyzes an intervening step.[26] If ArnA were to retain UX4O and NADH in the active site long enough to facilitate transfer to a second functional domain, it is possible that the conserved catalytic machinery would catalyze the reduction of UX4O to UDX. Instead, UX4O and NADH are initially released from the active without detectible UDX accumulation (Figure 2.5). Gel filtration studies indicate that the separately expressed decarboxylase and formyltransferase domains of ArnA do not interact with

each other or ArnB,[26] further supporting a model of rapid UX4O release rather than retention and transfer to a downstream enzyme in Lipid A metabolism.

IV. Identification of a catalytic and structural shunt in UGA decarboxylases.

Despite sharing a bifurcated mechanism (Figure 2.4 A), hUXS and ArnA qualitatively favor two very different reactions. hUXS produces predominantly UDX (Figure 2.2) while ArnA converts almost all available UGA to UX4O (Figure 2.5). We derived a bifurcated rate equation to quantify the catalytic shunt of UGA toward each product.[176] Modeling substrate saturation data with our rate equation identifies a fundamental difference between hUXS and ArnA. At comparable enzyme concentrations, ArnA and hUXS display apparent V_{max} values of 5100 and 56.4 nmol L^{-1} sec^{-1}, respectively, for NADH production (Table 2.2). This indicates that once reaching the E:NADH:UX4O intermediate state, the ArnA shunt favors release of the intermediates and a return to the apoenzyme (E*) state (Figure 2.4). In hUXS, UDX production is ~2 orders of magnitude faster than NADH and UX4O release,[176] indicating the hUXS shunt favors E:NADH:UX4O over the E* state. Additional evidence that hUXS and ArnA share a catalytic shunt is provided by the conservation of a cooperative conformational change. hUXS and ArnA both show positive cooperativity toward NAD^+ with Hill coefficients of 1.8 and 1.7, respectively (Figure 2.4, Table 2.2). Our analysis of the substrate induced conformational changes in ArnA suggests that a large portion of the catalytic domain folds in response to ligand. Structural equivalents of ArnA $Helix_{326-337}$, $Loop_{347-349}$ and $Helix_{373-381}$ are present in hUXS and likely to be the underlying basis for the E:NADH:UX4O to E* transition in UGA decarboxylases.

Cooperativity has previously been described in extended SDR enzymes that reversibly bind NAD(P)(H), similar to ArnA. Fluorescence studies of the CDP-glucose 4,6-dehydratase

(E_{od}) from *Yersinia pseudotuberculosis* (*Yp*E_{od}), indicate that NAD(H) binding is negatively cooperative (Figure 1.14).[116] Alternatively, fluorescence experiments of GDP-mannose 4,6-dehydratase (GMD) enzymes from the chlorella viruses PBCV-1 and ATCV-1 indicate that NADPH binding is positively cooperative (Figure 1.15).[118] NAD^+ binding data are currently unavailable for UGA decarboxylases and our studies do not preclude the contribution of cooperative binding to the observed kinetic activation. However, cooperative binding in the dehydratases has been examined in relationship to the N-terminal GXXGXXG motif,[116] highlighting variations in the amino acids found at the first and last X positions (Figure 1.7). Chlorella virus GMD enzymes possess a negatively charged (Asp) amino acid within the motif (Figure 1.7), while *Yp*E_{od} contains positively charged (His and Lys) residues.[116] The neutral ArnA motif has been simulated in *Yp*E_{od} by mutations that remove each charged residue from the GXXGXXG sequence (i.e. HΔG or KΔI). Either point mutation is reported to shift *Yp*E_{od} closer to a noncooperative NAD^+ binding mode,[116] suggesting that the uncharged motif in UGA decarboxylases may not contribute significantly to cooperative binding. Additionally, wild type *Yp*E_{od} displays classical Michaelis-Menten reaction kinetics,[116] suggesting an alternative mechanism may be at work in the cooperative catalytic cycles of hUXS and ArnA.

Our analysis of conformational changes in hUXS has identified alternative structural elements that form the catalytic shunt and influence the underlying equilibrium between E:NADH:UX4O and E*. hUXS $Loop_{120\text{-}124}$ excludes NAD^+ from solvent (Figure 2.6) and is predicted to favor the E:NADH:UX4O complex over substrate release. The equivalent loop in ArnA is truncated and has been proposed to allow NADH exchange from the E:NADH:UX4O complex.[24] Consistent with this proposal, mutating the hUXS-like loop in *Pyrobaculum calidifontis* GALE to resemble an ArnA-like structure results in reversible NAD^+ binding.[138]

hUXS Loop$_{266\text{-}272}$ folds over the NAD^+ binding site in both hUXS structures and is also likely to favor the E:NADH:UX4O complex over substrate release. The equivalent loop in ArnA contains a six residue insertion and occupies the NAD^+ binding site in the apo-ArnA structure. The ArnA conformation led Gatzeva et al. to propose that the loop may compete with ordered NAD^+ binding,[24] which might favor the E* state relative to hUXS. Eixelsberger et al. have previously proposed that the Glu204-Arg361 salt bridge serves as a gate to close the hUXS active site during catalysis. It is conceivable that such a gate would influence the E:NADH:UX4O to E* equilibrium.[44] However, as the authors note, the salt bridge residues are conserved in ArnA and other UGA decarboxylases. Our structure of open hUXS contains the salt bridge interaction in both chains, suggesting this interaction is unlikely to be part of the shunt. Additional elements that vary in sequence and flexibility, such as hUXS Helix$_{165\text{-}168}$, are more likely to impact the equilibrium.

V. The role of quaternary structure in UX4O and UDX formation.

Prior to our work, extended SDR enzymes were accepted to function as dimers.[73; 81] However, within the UGA decarboxylase class, only *E. coli* ArnA had been examined. Previous crystallographic studies show full-length ArnA forms a hexamer centered around three typical SDR dimers (Figure 1.22 A).[25] The full-length ArnA hexamer is built upon the α* helices from adjacent decarboxylase domains. The decarboxylase hexamer orients the active site away from the oligomer interfaces such that the active site is accessible via solvent channel (Figure 4.1). Gel filtration indicates the oligomeric state of full-length ArnA does not change in the presence of substrate,[25] confirming that UX4O production occurs in the context of the hexamer. Separating ArnA from full length ArnA is reported to abolish the hexameric solution state without

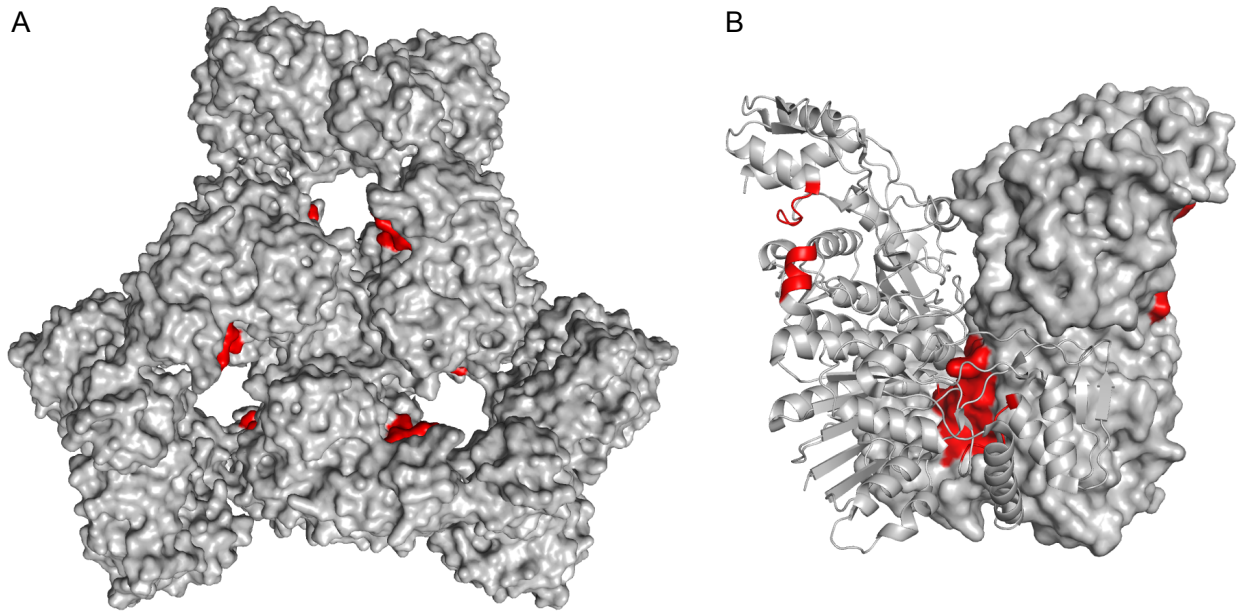


Figure 4.1. Active site orientation in the UGA decarboxylase oligomeric structures. (A) Surface rendering of the full length *E. coli* ArnA hexamer (grey) (PDB 1Z7E). ArnA $Helix_{396\text{-}400}$ and $Loop_{498\text{-}510}$ (red) form the part of the active site and are oriented toward the interior of solvent channels. (B) The open hUXS tetramer with the respective elements, $Helix_{165\text{-}168}$ and $Loop_{26\text{-}272}$, colored as above. Only the hUXS dimer contributing an ordered $Helix_{165\text{-}168}$ to the tetramer interface is rendered as a surface in order to illustrate the burial of two of the four active sites.

impacting the apparent K_m or V_{max} for UGA and NAD^+.[24] The results of Gatzeva-Topalova *et al.*[24] are consistent with the body of extended SDR literature, in that the dimer is active and distinct functions are not associated with higher oligomeric states.

We provide evidence that changes in the hUXS oligomeric state are associated with a large increase in UDX production and have little impact on release of the NADH and UX4O intermediates. Sedimentation velocity experiments indicate that hUXS transiently undergoes a dimer to tetramer self-association (Figure 3.1, Table 3.1). At concentrations below the K_d (4.7μM), dimeric hUXS converts UGA and NAD^+ substrates to detectable levels of NADH (Figure 3.7 B). Increasing the hUXS concentration causes a 10-20% increase in NADH release that can be modeled by a binding isotherm using the orthogonally determined K_d for self-association. Due to rapid UDX formation at hUXS concentrations above the K_d, we were unable to duplicate the AUC and NADH isotherms for UDX formation. Since crowding agents have been shown to stimulate protein association[194] and association based catalytic activities,[195] we chose a dimeric hUXS concentration and assayed UDX production in the presence of the small molecule crowder TMAO. We show that UDX formation across a range of TMAO concentrations resembles a binding isotherm, consistent with the dimer-tetramer association (Figure 3.7 A). Strikingly, the higher order oligomer increases UDX production by greater than 10-fold.

Analysis of our unliganded hUXS crystal structure reveals a possible oligomeric explanation for the increase in UDX production relative to NADH. Unlike ArnA and other extended SDR enzymes, hUXS has evolved a short loop (loop*) between beta strands β3 and β4 in place of the α* helix (Figure 1.8). Loop* is expected to provide a less extensive packing surface than the α* helix and may not facilitate formation of an ArnA-like hexamer in hUXS

structures (Figure 1.22). Instead, unliganded hUXS crystallizes with a dimer in the asymmetric unit that packs into a fiber along the crystallographic 3_2 axis (Figure 3.2 A,B). Adjacent dimers bury ~4200Å^2 of surface area in the tetramer interface (Figure 3.2 C), including the $Loop_{266-272}$ region from two of the four protomers (Figure 4.1). Packing also buries $Helix_{165-168}$ from the protomers in the tetramer interface. Steric constraints prevent both helices from ordering in the interface, however, and one $Helix_{165-168}$ remains disordered. $Loop_{266-272}$ and $Helix_{165-168}$ conformations suggest that the hUXS tetramer interface occludes two of the four active sites from bulk solvent, in contrast to the ArnA hexamer quaternary structure. In the context of hUXS catalysis, shielding the active site may increase UDX production 10-fold by stabilizing interactions between $Loop_{266-272}$ and NAD(H) and/or limiting the conformational changes in $Helix_{165-168}$. In broader terms, the hUXS tetramer may serve as another structural element of the catalytic shunt and influence the equilibrium between E:NADH:UX4O and E*.

VI. UDP binding at a putative allosteric site prevents the tetramer crystal form.

Since the hUXS tetramer appears to be an active species, we asked whether the higher order oligomer was also present in the hUXS:UDP crystal form. Reexamination of the hUXS:UDP unit cell shows the typical extended SDR dimer[73] but not the extensive contacts needed to build the butterfly-shaped tetramer or crystalline fibers (Figure 4.2A). One feature unique to the hUXS:UDP crystal form is the binding of a UDP molecule at an allosteric site along the dimer interface (Figure 4.2A).[44] The allosteric pocket is well defined in both hUXS crystal forms and allows us assess the impact of UDP binding on tetramer formation. We aligned the hUXS:UDP and tetramer structures and found an overlap in the allosteric site and the oligomer building interface that may account for different quaternary states (Figure 4.2D). In the

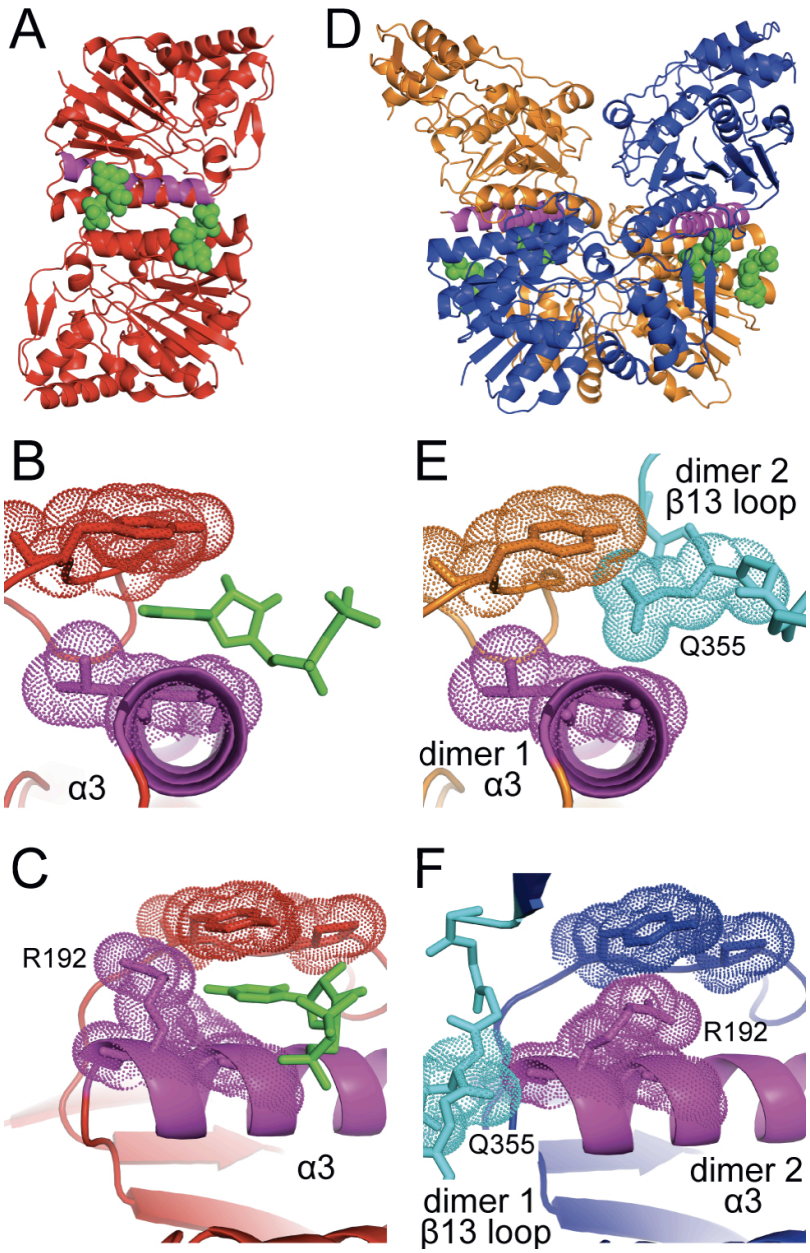


Figure 4.2. UDP binding at a putative allosteric site disrupts the tetramer interface. (A) Crystal structure of hUXS (red, PBD entry 2b69) showing UDP (green spheres) bound at the dimer interface α3 helix (magenta, colored in top monomer only). (B) Close up view of the putative allosteric site showing UDP (green sticks) binds in a cleft formed by Pro148 and Tyr150 (red sticks, dots) and (C) the α3 helix residues Gly188, Lue189, and Arg192 (magenta sticks, dots). (D) Conceptual model of the asymmetric tetramer with UDP bound in the allosteric sites to show the tetramer interface and UDP binding sites overlap at the α3 helix. Chain A and Chain B of each dimer are colored orange and blue, respectively. (E) Close up view of the tetramer interface oriented as (B) to show Gln355 from the β13 loop (cyan sticks, dots) occupies the allosteric site of an adjacent dimer. (F) Close up view of the analogous tetramer interface oriented as in (C) to show the asymmetric conformation of Arg192, rather than the β13 loop, occludes the second allosteric site.

dimeric state, a UDP molecule is bound at the allosteric site against the α3 helix from one monomer with the pyrophosphate oriented across the dimer interface (Figure 4.2B,C). As we discussed previously, the tetrameric structure displays asymmetric conformations at the α3 helix (Figure 3.2). In one conformation, Gln355 from dimer 2 reaches across the tetramer interface toward the α3 helix of dimer 1. The side chain and main chain of Gln355 occlude the binding sites for the ribose and β-phosphate of UDP, respectively, and appear to stabilize the tetramer (Figure 4.2E). In a second conformation, the β13 loop is removed from the allosteric site and packs against the C-terminus of the α3 helix. Consequently, Arg192 of the α3 helix swings into the pocket above Gly188, and packs against the allosteric site (Figure 4.2F). We also examined the allosteric site in the 'extended' tetramer model and found similar interactions with the β13 loop that are predicted to occlude the allosteric site.

Given our model of hUXS regulation via quaternary changes (Figure 3.7), it is tempting to speculate that the allosteric site servers a regulatory role. Since each allosteric site is occluded by formation of the asymmetric and extended tetramer structures, UDP binding might prevent formation of the high activity tetramer, effectively replacing the "?" in our current model of inactivation. It is currently unclear whether UDP acts as an allosteric effector or merely serves as a mimic for another UDP-sugar. Sugar moieties extending from the β-phosphate of UDP are likely to cause additional steric clashes that would destabilize the tetramer and disrupt the high activity complex. UDX, along with UGA, UDP-glucose, UDP-galactose, UDP-N-acetylglucosamine and UDP-N-acetylgalactosamine,[15; 201; 202; 203] are believed to accumulate in the endoplasmic reticulum/golgi with hUXS. Of the possible effectors, UDX has been shown to allosterically inhibit UXS from chicken chondrocytes and wheat germ[1; 2] and may do so by

binding at the putative allosteric site. Investigations into the role of the allosteric site and potential effectors are currently underway.

VII. A structure based classification system for UGA decarboxylases.

Our work provides a structural basis for the equilibrium between E:NADH:UX4O and E*, as well as a possible means to regulate hUXS activity. We have selected 5 criteria that are likely to influence the equilibrium shunt: (i) bifunctional decarboxylase/formyltransferase; (ii) separate decarboxylase domain; (iii) NADH binding loop; (iv) loop* and (v) allosteric site. Using our criteria to establish subclasses of UGA decarboxylases, we can generate hypotheses about the propensity of each group to form UX4O and UDX (Table 4.1).

In our scheme, full length *E. coli* ArnA is a Type I UGA decarboxylase and represents the extreme case of a UX4O synthase. *E. coli* ArnA is a bifunctional decarboxylase/formyl-transferase fusion protein that forms a stable hexamer in solution.[25] The hexamer active site is exposed to bulk solvent for substrate and product exchange (Figure 4.1). $Loop_{347\text{-}349}$ serves as a poor NADH binding loop (Figure 2.6) and facilitates NADH and UX4O exchange. The equilibrium in *E. coli* ArnA favors E* and leads to UX4O release before UDX is formed. Type I decarboxylases containing both decarboxylase and formyltransferase domains are also annotated in *Salmonella* species, *Pseudomonas aeruginosa*, *Burkholderia cepacia*,[204] *Burkholderia pseudomallei*,[205] and *Wigglesworthia glossinidia*.[206]

Our classification system identifies non-fused ArnA decarboxylase domains as Type II UGA decarboxylases. The ArnA gene in *Ralstonia solanacearum*, which naturally contains an internal stop codon and 10 spacer nucleotides,[27] encodes distinct transformylase and decarboxylase domains. The resulting decarboxylase, U4kpxs, lacks the additional

Table 4.1. Classification scheme for UGA decarboxylases.

		Classification Criteria					
Class	Example	Fusion	Separate	NADH Loop	Loop*	Allosteric Site	Product
Type I	ArnA	X	−	−	−	−	UX4O
Type II	U4kpxs	−	X	−	−	−	UX4O
Type III	CalS9	−	X	X	−	−	UX4O/UDX
Type IV	SmUXS	−	X	X	X	−	UDX
Type V	hUXS	−	X	X	X	X	UDX

transformylase packing surfaces that stabilize the hexameric state (Figure 1.22). Gel filtration studies indicate that deletion of the formyltransferase domain in *E. coli* ArnA can disrupt the hexamer.[24] We have also identified dimer, tetramer and hexamer species in preliminary AUC experiments of the *E. coli* ArnA decarboxylase domain, suggesting U4kpxs may naturally form lower ordered oligomers. In the lower order oligomers, the active sites of Type II decarboxylases are not likely to be sterically protected by the large hexamer structure. Under the crowded conditions of a cell, we predict U4kpxs is likely to encounter packing interactions along the active site. In the event that packing occurs during catalysis, equilibrium may be slightly shifted toward E:NADH:UX4O. Sequence analysis indicates that U4kpxs contains a truncated NADH binding loop, however, and is predicted to produce mainly UX4O. A Type II UGA decarboxylase is also annotated in *Ralstonia metallidurans*.[207]

Type III UGA decarboxylases in our scheme are single decarboxylase domain enzymes with an extended NADH binding loop. The decarboxylase identified in *M. echinospora* ssp. *calichensis*, CalS9, contains three consecutive proline residues in the sequence region corresponding to hUXS $Loop_{120\text{-}124}$. The prolines may form part of a flexible loop near the adenine of NAD^+ and increase substrate retention. CalS9 was only assayed in the presence of exogenous NAD^+,[35] however, leaving the possibility that NAD^+ may not be bound tightly as a cofactor. Mass spectrum data indicate that the major product of CalS9 reactions is UDX, although a species with the expected m/z ratio of UX4O is also clearly present.[35] CalS9 reactions were incubated for 90 minutes, and as we have shown with ArnA, decarboxylase production can change from UX4O to UDX depending on the availability of substrate (Figure 2.5). CalS9 may also contain a modified α* helix that could impact oligomer formation. The CalS9 residues corresponding to the α* helix are truncated, similar to the hUXS loop*, but contain four

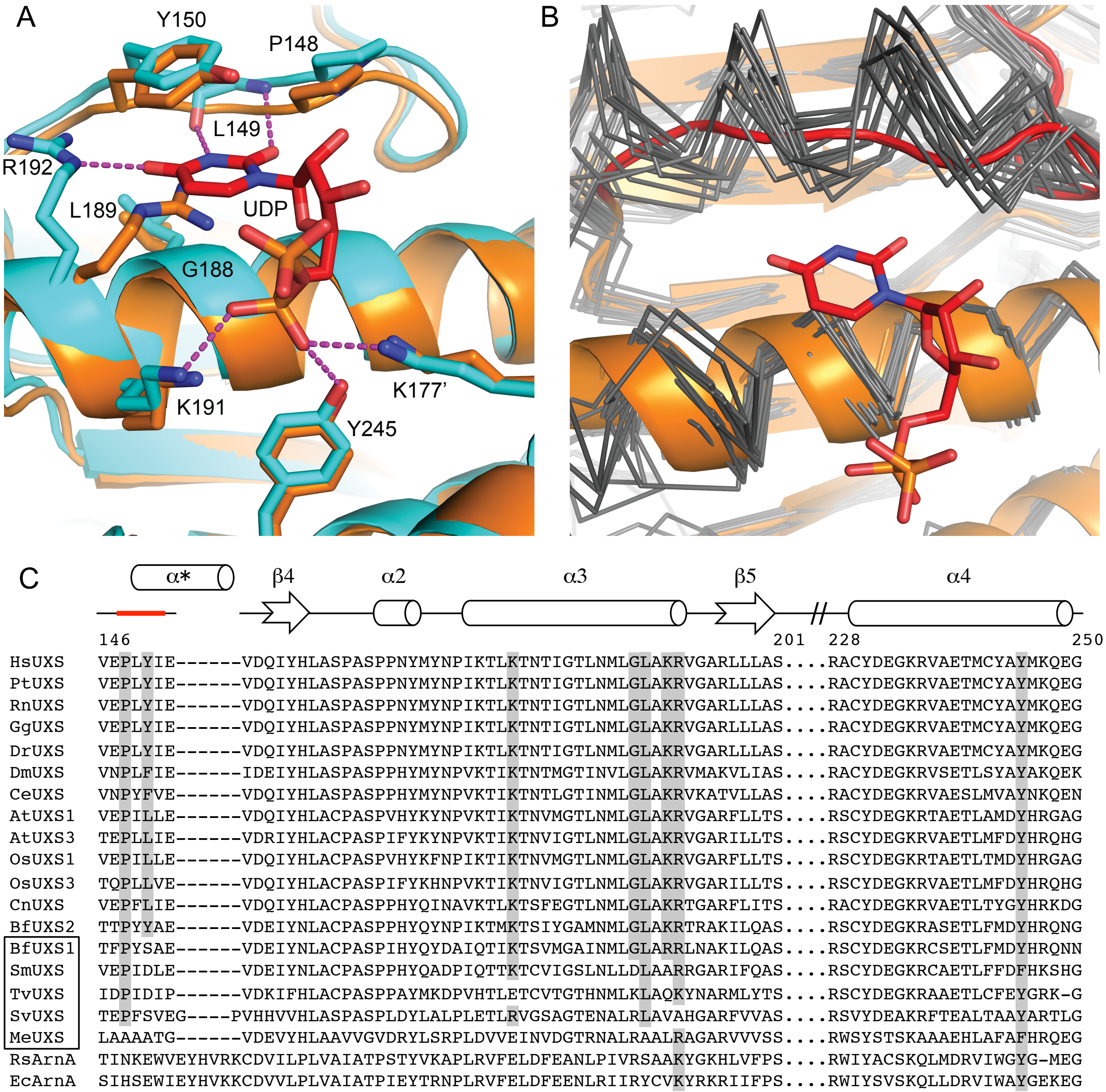
consecutive Ala residues. Alanine is the most helix stabilizing amino acid,[208; 209] suggesting the modified α* helix may still facilitate some hexamer formation. Together, the flexible NADH loop and weakened hexamer interface are predicted to alter the E:NADH:UX4O to E* equilibrium toward UDX formation. Our prediction is that Type III UGA decarboxylases do not show an initial lag in UDX formation and release UX4O and UDX with similar initial velocities.

Type IV UGA decarboxylases contain an NADH binding loop and loop* structure similar to hUXS. As a consequence, these decarboxylases are more likely to retain NAD^+ as a cofactor. The reduced packing surface of loop* is predicted to hinder formation of the hexamer and solvent accessible active sites. Instead, lower order self-assemblies (i.e. the hUXS tetramer) or hetero-associations (i.e. molecular crowding) may close the active site and promote UDX formation. The decarboxylase identified in *S. meliloti* is characteristic of a Type IV UGA decarboxylase. The extended NADH binding loop residues in dimeric *Sm*UXS (NFSTG) are similar in sequence to hUXS (NFFTG) and may explain the production of UDX in the absence of exogenous NAD^+ (Table 1.2).[36] Class IV UGA decarboxylases from *T. vaginalis*, *S. viridochromogenes* (AviE2), and *B. fragilis* (*Bf*UXS1) are predicted to favor the E:NADH:UX4O state and form primarily UDX.

Finally, the Type V UGA decarboxylases represent the extreme case of UDX synthases. Type V decarboxylases are activated by tetramer self-association and favor the E:NADH:UX4O state. In addition, this class of decarboxylase may be regulated by perturbation of the quaternary structure. We believe that a UDP binding cleft remote from the active site serves as an authentic allosteric site. Our modeling suggests that UDP bound in the allosteric site is likely to disrupt the tetramer interface, and may explain the absence of the asymmetric tetramer in the crystal structure of the hUXS:UDP complex. The allosteric site is formed by GLAKR and PxB motifs

(where B is the bulky side chain of Tyr, Phe or Lue) (Figure 4.3). The allosteric site also coordinates the UDP α-phosphate via interactions with one Tyr and two Lys side chains. The motifs and phosphate coordinating residues are conserved in all animal, plant and fungi UXS enzymes examined. Interestingly, *B. fragilis* contains two UXS isoforms. *Bf*UXS1 and *Bf*UXS2 differ in the allosteric site motifs and are classified as Type IV and Type V UGA decarboxylases, respectively.

Figure 4.3. Conservation of the allosteric site interactions. (A) Overlay of the allosteric site of unliganded hUXS (orange) and the hUXS:UDP complex (cyan, PDB entry 2B69) showing hydrogen bonding interactions (magenta dashes) between UDP (red) and the main chain atoms of L149 (shown only in 2B69 for clarity) or side chains of Y245, R192, K191 and K177' (from the adjacent protomer). L189, G188, Y150 and P148 provide packing surfaces for UDP (red). (B) Structural divergence in the allosteric loop (red) of hUXS (orange ribbons). The C_α traces (grey) of representative SDR enzymes from all kingdoms are superimposed to illustrate the widespread conservation of an α-helix in this region, and the structural divergence that occurs in UXS. SDR structures used: human GDP-mannose 4,6-dehydratase (PDB entry 1T2A), *E. coli* dTDP-glucose 4,6-dehydratase (1BXK), *Salmonella enterica* dTDP-glucose 4,6- dehydratase (1KEU), *Helicobacter pylori* UDP-GlcNAC 4,6-dehydratase (2GN9), *Arabidopsis thaliana* GDP-mannose 4,6-dehydratase (1N7G), *Yersinia pseudotuberculosis* CDP-glucose 4,6-dehydratase (1RKX), *Streptococcus suis* dTDP-glucose 4,6-dehydratase (1OC2), *Bordetella bronchiseptica* putative SDR (2Q1T), *Trypanosoma brucei* UDP-galactose 4-epimerase (2CNB), human UDP-galactose 4-epimerase (1EK5), *Bacillus anthracis* UDP-glucose 4-epimerase (2C20) and *A. thaliana* sulfolipid biosynthesis protein (1QRR). (C) Sequence alignment of the allosteric site residues from known UGA decarboxylases. UXS and ArnA enzymes were assigned based on published evidence of UDX and UDP-4-keto-xylose synthesis, respectively, or high sequence identity (>70%). Alpha helices (cylinders), beta strands (arrows) and residue numbers (146-201 and 228-250) correspond to hUXS. A short loop* (red line) before β-strand 4 is conserved in all examined members of the UXS family, but diverges from the α-helix (α*) found in ArnA and other members of the extended SDR family. Annotated UXS enzymes without the full allosteric site fingerprint are boxed. Aligned enzymes are human HsUXS (organism *Homo sapiens*, Accession NP079352, 100% sequence identity to hUXS), chimpanzee PtUXS (*Pan troglodytes*, BAK63103, 99%), rat RnUXS (*Rattus norvegicus*, NP647552, 97%), chicken GgUXS (*Gallus gallus*, XP416926, 94%), zebrafish DrUXS (*Danio rerio*, AAH74058.1, 89%), fruit fly DmUXS (*Drosophila melanogaster*, NP648182, 72%), flatworm CeUXS (a.k.a. SQV-1, *Caenorhabditis elegans*, NP501418, 76%), *A. thaliana* AtUXS membrane-bound isoform 1 (NP190920, 63%) and soluble isoform 3 (NP200737, 64%), *Oryza sativa* OsUXS membrane-bound isoform 1 (BAD12491, 63%) and soluble isoform 3 (BAB84334, 63%), *Cryptococcus neoformans* var. neoformans CnUXS (AAK59981, 62%), *Bacteroides fragilis* BfUXS isoform 1 (CAH07883, 55%) and isoform 2 (CAH07260, 59%), *Sinorhizobium meliloti* SmUXS (ACY30251, 58%), *Trichomonas vaginalis* TvUXS (CCC58322, 57%), *Streptomyces viridochromogenes* SvUXS (a.k.a. AviE2, AAK83183, 50%), *Micromonospora echinospora* MeUXS (a.k.a. CalS9, AAM70333, 32%), *Ralstonia solanacearum* RaArnA (a.k.a. U4kpxs, GQ369438, 27%), and *Escherichia coli* EcArnA (AAL23678, 26% with hUXS, 53% with RaArnA).


A
Y150
P148
L149
R192
L189
UDP
G188
K191
K177'
Y245
B
C
α*
β4
α2
α3
β5
α4
146
201
228
250
HsUXS VEPLYIE------VDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNMLGLAKRVGARLLLAS....RACYDEGKRVAETMCYAYMKQEG
PtUXS VEPLYIE------VDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNMLGLAKRVGARLLLAS....RACYDEGKRVAETMCYAYMKQEG
RnUXS VEPLYIE------VDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNMLGLAKRVGARLLLAS....RACYDEGKRVAETMCYAYMKQEG
GgUXS VEPLYIE------VDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNMLGLAKRVGARLLLAS....RACYDEGKRVAETMCYAYMKQEG
DrUXS VEPLYIE------VDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNMLGLAKRVGARLLLAS....RACYDEGKRVAETMCYAYMKQEG
DmUXS VNPLFIE------IDEIYHLASPASPPHYMYNPVKTIKTNTMGTINVLGLAKRVMAKVLIAS....RACYDEGKRVSETLSYAYAKQEK
CeUXS VNPYFVE------VDQIYHLASPASPPHYMYNPVKTIKTNTLGTINMLGLAKRVKATVLLAS....RACYDEGKRVAESLMVAYNKQEN
AtUXS1 VEPILLE------VDQIYHLACPASPVHYKYNPVKTIKTNVMGTLNMLGLAKRVGARFLLTS....RSCYDEGKRTAETLAMDYHRGAG
AtUXS3 TEPLLIE------VDRIYHLACPASPIFYKYNPVKTIKTNVIGTLNMLGLAKRVGARILLTS....RSCYDEGKRVAETLMFDYHRQHG
OsUXS1 VEPILLE------VDQIYHLACPASPVHYKFNPIKTIKTNVMGTLNMLGLAKRVGARFLLTS....RSCYDEGKRTAETLTMDYHRGAG
OsUXS3 TQPLLVE------VDQIYHLACPASPIFYKHNPVKTIKTNVIGTLNMLGLAKRVGARILLTS....RSCYDEGKRVAETLMFDYHRQHG
CnUXS VEPFLIE------VDQIYHLACPASPPHYQINAVKTLKTSFEGTLNMLGLAKRTGARFLITS....RACYDEGKRVAETLTYGYHRKDG
BfUXS2 TTPYYAE------VDEIYNLACPASPPHYQYNPIKTMKTSIYGAMNMLGLAKRTRAKILQAS....RSCYDEGKRASETLFMDYHRQNG
BfUXS1 TFPYSAE------VDEIYNLACPASPIHYQYDAIQTIKTSVMGAINMLGLARRLNAKILQAS....RSCYDEGKRCSETLFMDYHRQNN
SmUXS VEPIDLE------VDEIYNLACPASPPHYQADPIQTTKTCVIGSLNLLDLAARRGARIFQAS....RSCYDEGKRCAETLFFDFHKSHG
TvUXS IDPIDIP------VDKIFHLACPASPPAYMKDPVHTLETCVTGTHNMLKLAQKYNARMLYTS....RSCYDEGKRAAETLCFEYGRK-G
SvUXS TEPFSVEG----PVHHVVHLASPASPLDYLALPLETLRVGSAGTENALRLAVAHGARFVVAS....RSVYDEAKRFTEALTAAYARTLG
MeUXS LAAAATG------VDEVYHLAAVVGVDRYLSRPLDVVEINVDGTRNALRAALRAGARVVVSS....RWSYSTSKAAAEHLAFAFHRQEG
RsArnA TINKEWVEYHVRKCDVILPLVAIATPSTYVKAPLRVFELDFEANLPIVRSAAKYGKHLVFPS....RWIYACSKQLMDRVIWGYG-MEG
EcArnA SIHSEWIEYHVKKCDVVLPLVAIATPIEYTRNPLRVFELDFEENLRIIRYCVKYRKRIIFPS....RWIYSVSKQLLDRVIWAYGEKEG

REFERENCES

1. John, K. V., Schutzbach, J. S. & Ankel, H. (1977). Separation and allosteric properties of two forms of UDP-glucuronate carboxy-lyase. *J Biol Chem* **252**, 8013-7.
2. John, K. V., Schwartz, N. B. & Ankel, H. (1977). UDP-glucuronate carboxy-lyase in cultured chondrocytes. *J Biol Chem* **252**, 6707-10.
3. Harper, A. D. & Bar-Peled, M. (2002). Biosynthesis of UDP-xylose. Cloning and characterization of a novel Arabidopsis gene family, UXS, encoding soluble and putative membrane-bound UDP-glucuronic acid decarboxylase isoforms. *Plant Physiol* **130**, 2188-98.
4. Kobayashi, M., Nakagawa, H., Suda, I., Miyagawa, I. & Matoh, T. (2002). Purification and cDNA cloning of UDP-D-glucuronate carboxy-lyase (UDP-D-xylose synthase) from pea seedlings. *Plant Cell Physiol* **43**, 1259-65.
5. Suzuki, K., Watanabe, K., Masumura, T. & Kitamura, S. (2004). Characterization of soluble and putative membrane-bound UDP-glucuronic acid decarboxylase (OsUXS) isoforms in rice. *Arch Biochem Biophys* **431**, 169-77.
6. Pattathil, S., Harper, A. D. & Bar-Peled, M. (2005). Biosynthesis of UDP-xylose: characterization of membrane-bound AtUxs2. *Planta* **221**, 538-48.
7. Eames, B. F., Singer, A., Smith, G. A., Wood, Z. A., Yan, Y. L., He, X., Polizzi, S. J., Catchen, J. M., Rodriguez-Mari, A., Linbo, T., Raible, D. W. & Postlethwait, J. H. (2010). UDP xylose synthase 1 is required for morphogenesis and histogenesis of the craniofacial skeleton. *Dev Biol* **341**, 400-15.
8. Kadirvelraj, R., Sennett, N. C., Polizzi, S. J., Weitzel, S. & Wood, Z. A. (2011). Role of Packing Defects in the Evolution of Allostery and Induced Fit in Human UDP-Glucose Dehydrogenase. *Biochemistry* **50**, 5780-9.
9. Silbert, J. E. & Sugumaran, G. (2002). Biosynthesis of chondroitin/dermatan sulfate. *IUBMB Life* **54**, 177-86.
10. Wilson, J. E. (2003). Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. *The Journal of experimental biology* **206**, 2049-57.
11. Ninfali, P., Accorsi, A., Palma, F., Fazi, A., Piatti, E., Chiarantini, L. & Fornaini, G. (1984). Human erythrocyte phosphoglucomutase: comparison of the kinetic properties of PGM1 and PGM2 isoenzymes. *Biochimie* **66**, 617-23.
12. Duggleby, R. G., Chao, Y. C., Huang, J. G., Peng, H. L. & Chang, H. Y. (1996). Sequence differences between human muscle and liver cDNAs for UDPglucose pyrophosphorylase and kinetic properties of the recombinant enzymes expressed in Escherichia coli. *European journal of biochemistry / FEBS* **235**, 173-9.
13. Sommer, B. J., Barycki, J. J. & Simpson, M. A. (2004). Characterization of human UDP-glucose dehydrogenase. CYS-276 is required for the second of two successive oxidations. *J Biol Chem* **279**, 23590-6.
14. Sennett, N. C., Kadirvelraj, R. & Wood, Z. A. (2011). Conformational flexibility in the allosteric regulation of human UDP-alpha-D-glucose 6-dehydrogenase. *Biochemistry*.
15. Muraoka, M., Kawakita, M. & Ishida, N. (2001). Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity. *FEBS Lett* **495**, 87-93.
16. Kearns, A. E., Vertel, B. M. & Schwartz, N. B. (1993). Topography of glycosylation and UDP-xylose production. *The Journal of biological chemistry* **268**, 11097-104.

17. Moriarity, J. L., Hurt, K. J., Resnick, A. C., Storm, P. B., Laroy, W., Schnaar, R. L. & Snyder, S. H. (2002). UDP-glucuronate decarboxylase, a key enzyme in proteoglycan synthesis: cloning, characterization, and localization. *J Biol Chem* **277**, 16968-75.
18. Bakker, H., Oka, T., Ashikov, A., Yadav, A., Berger, M., Rana, N. A., Bai, X., Jigami, Y., Haltiwanger, R. S., Esko, J. D. & Gerardy-Schahn, R. (2009). Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase. *J Biol Chem* **284**, 2576-83.
19. Ankel, H. & Feingold, D. S. (1965). Biosynthesis of Uridine Diphosphate D-Xylose. I. Uridine Diphosphate Glucuronate Carboxy-lyase of Wheat Germ*. *Biochemistry* **4**, 2468-2475.
20. Schutzbach, J. S. & Feingold, D. S. (1970). Biosynthesis of uridine diphosphate D-xylose. IV. Mechanism of action of uridine diphosphoglucuronate carboxy-lyase. *J Biol Chem* **245**, 2476-82.
21. Ankel, H. & Feingold, D. S. (1966). Biosynthesis of uridine diphosphate D-xylose. II. Uridine diphosphate D-glucuronate carboxy-lyase of Cryptococcus laurentii. *Biochemistry* **5**, 182-9.
22. Persson, B., Kallberg, Y., Bray, J. E., Bruford, E., Dellaporta, S. L., Favia, A. D., Duarte, R. G., Jornvall, H., Kavanagh, K. L., Kedishvili, N., Kisiela, M., Maser, E., Mindnich, R., Orchard, S., Penning, T. M., Thornton, J. M., Adamski, J. & Oppermann, U. (2009). The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. *Chemico-biological interactions* **178**, 94-8.
23. Breazeale, S. D., Ribeiro, A. A. & Raetz, C. R. (2002). Oxidative decarboxylation of UDP-glucuronic acid in extracts of polymyxin-resistant Escherichia coli. Origin of lipid a species modified with 4-amino-4-deoxy-L-arabinose. *J Biol Chem* **277**, 2886-96.
24. Gatzeva-Topalova, P. Z., May, A. P. & Sousa, M. C. (2004). Crystal structure of Escherichia coli ArnA (PmrI) decarboxylase domain. A key enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry* **43**, 13370-9.
25. Gatzeva-Topalova, P. Z., May, A. P. & Sousa, M. C. (2005). Structure and mechanism of ArnA: conformational change implies ordered dehydrogenase mechanism in key enzyme for polymyxin resistance. *Structure* **13**, 929-42.
26. Williams, G. J., Breazeale, S. D., Raetz, C. R. & Naismith, J. H. (2005). Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *J Biol Chem* **280**, 23000-8.
27. Breazeale, S. D., Ribeiro, A. A., McClerren, A. L. & Raetz, C. R. (2005). A formyltransferase required for polymyxin resistance in Escherichia coli and the modification of lipid A with 4-Amino-4-deoxy-L-arabinose. Identification and function oF UDP-4-deoxy-4-formamido-L-arabinose. *J Biol Chem* **280**, 14154-67.
28. Gu, X., Glushka, J., Yin, Y., Xu, Y., Denny, T., Smith, J., Jiang, Y. & Bar-Peled, M. (2010). Identification of a bifunctional UDP-4-keto-pentose/UDP-xylose synthase in the plant pathogenic bacterium Ralstonia solanacearum strain GMI1000, a distinct member of the 4,6-dehydratase and decarboxylase family. *J Biol Chem* **285**, 9030-40.
29. Coyne, M. J., Fletcher, C. M., Reinap, B. & Comstock, L. E. (2011). UDP-glucuronic acid decarboxylases of Bacteroides fragilis and their prevalence in bacteria. *Journal of bacteriology*.

30. Rosenberger, A. F., Hangelmann, L., Hofinger, A. & Wilson, I. B. (2011). UDP-xylose and UDP-galactose synthesis in Trichomonas vaginalis. *Molecular and biochemical parasitology*.
31. Singh, B. N., Hayes, G. R., Lucas, J. J., Sommer, U., Viseux, N., Mirgorodskaya, E., Trifonova, R. T., Sassi, R. R., Costello, C. E. & Fichorova, R. N. (2009). Structural details and composition of Trichomonas vaginalis lipophosphoglycan in relevance to the epithelial immune function. *Glycoconjugate journal* **26**, 3-17.
32. Ma, L., Lu, H., Sprinkle, A., Parsek, M. R. & Wozniak, D. J. (2007). Pseudomonas aeruginosa Psl is a galactose- and mannose-rich exopolysaccharide. *Journal of bacteriology* **189**, 8353-6.
33. Forsberg, L. S. & Carlson, R. W. (2008). Structural characterization of the primary O-antigenic polysaccharide of the Rhizobium leguminosarum 3841 lipopolysaccharide and identification of a new 3-acetimidoylamino-3-deoxyhexuronic acid glycosyl component: a unique O-methylated glycan of uniform size, containing 6-deoxy-3-O-methyl-D-talose, n-acetylquinovosamine, and rhizoaminuronic acid (3-acetimidoylamino-3-deoxy-D-gluco-hexuronic acid). *The Journal of biological chemistry* **283**, 16037-50.
34. Hofmann, C., Boll, R., Heitmann, B., Hauser, G., Durr, C., Frerich, A., Weitnauer, G., Glaser, S. J. & Bechthold, A. (2005). Genes encoding enzymes responsible for biosynthesis of L-lyxose and attachment of eurekanate during avilamycin biosynthesis. *Chem Biol* **12**, 1137-43.
35. Simkhada, D., Oh, T. J., Pageni, B. B., Lee, H. C., Liou, K. & Sohng, J. K. (2009). Characterization of CalS9 in the biosynthesis of UDP-xylose and the production of xylosyl-attached hybrid compound. *Appl Microbiol Biotechnol* **83**, 885-95.
36. Gu, X., Lee, S. G. & Bar-Peled, M. (2011). Biosynthesis of UDP-xylose and UDP-arabinose in Sinorhizobium meliloti 1021: first characterization of a bacterial UDP-xylose synthase, and UDP-xylose 4-epimerase. *Microbiology* **157**, 260-9.
37. Bar-Peled, M., Griffith, C. L. & Doering, T. L. (2001). Functional cloning and characterization of a UDP- glucuronic acid decarboxylase: the pathogenic fungus Cryptococcus neoformans elucidates UDP-xylose synthesis. *Proc Natl Acad Sci U S A* **98**, 12003-8.
38. Pan, Y., Wang, X., Liu, H., Zhang, G. & Ma, Z. (2010). Molecular Cloning of Three UDP-Glucuronate Decarboxylase Genes That Are Preferentially Expressed In Gossypium Fibers From Elongation to Secondary Cell Wall Synthesis. *Journal of Plant Biology* **53**, 367-373.
39. Zhang, Q., Shirley, N., Lahnstein, J. & Fincher, G. B. (2005). Characterization and expression patterns of UDP-D-glucuronate decarboxylase genes in barley. *Plant physiology* **138**, 131-41.
40. Bindschedler, L. V., Wheatley, E., Gay, E., Cole, J., Cottage, A. & Bolwell, G. P. (2005). Characterisation and expression of the pathway from UDP-glucose to UDP-xylose in differentiating tobacco tissue. *Plant Mol Biol* **57**, 285-301.
41. Hwang, H. Y. & Horvitz, H. R. (2002). The SQV-1 UDP-glucuronic acid decarboxylase and the SQV-7 nucleotide-sugar transporter may act in the Golgi apparatus to affect Caenorhabditis elegans vulval morphogenesis and embryonic development. *Proc Natl Acad Sci U S A* **99**, 14218-23.

42. Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R. E., Farnet, C. M. & Thorson, J. S. (2002). The calicheamicin gene cluster and its iterative type I enediyne PKS. *Science* **297**, 1173-6.
43. Wheatley, E. R., Davies, D. R. & Bolwell, G. P. (2002). Characterisation and immunolocation of an 87 kDa polypeptide associated with UDP-glucuronic acid decarboxylase activity from differentiating tobacco cells (Nicotiana tabacum L.). *Phytochemistry* **61**, 771-80.
44. Eixelsberger, T., Sykora, S., Egger, S., Brunsteiner, M., Kavanagh, K. L., Oppermann, U., Brecker, L. & Nidetzky, B. (2012). Structure and mechanism of human UDP-xylose synthase: evidence for a promoting role of sugar ring distortion in a three-step catalytic conversion of UDP-glucuronic acid. *The Journal of biological chemistry* **287**, 31349-58.
45. Weitnauer, G., Muhlenweg, A., Trefzer, A., Hoffmeister, D., Sussmuth, R. D., Jung, G., Welzel, K., Vente, A., Girreser, U. & Bechthold, A. (2001). Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the avi biosynthetic gene cluster of Streptomyces viridochromogenes Tu57 and production of new antibiotics. *Chemistry & biology* **8**, 569-81.
46. Vaishnav, V. V., Bacon, B. E., O'Neill, M. & Cherniak, R. (1998). Structural characterization of the galactoxylomannan of Cryptococcus neoformans Cap67. *Carbohydrate research* **306**, 315-30.
47. Moyrand, F., Klaproth, B., Himmelreich, U., Dromer, F. & Janbon, G. (2002). Isolation and characterization of capsule structure mutant strains of Cryptococcus neoformans. *Molecular microbiology* **45**, 837-49.
48. Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature reviews. Molecular cell biology* **6**, 850-61.
49. Matern, U. & Grisebach, H. (1977). UDP-apiose/UDP-xylose synthase. Subunit composition and binding studies. *European journal of biochemistry / FEBS* **74**, 303-12.
50. Baron, D., Wellmann, E. & Grisebach, H. (1972). Purification and properties of an enzyme from cell suspension cultures of parsley catalyzing the synthesis of UDP-apiose and UDP-D-xylose from UDP-D-glucuronic acid. *Biochimica et biophysica acta* **258**, 310-8.
51. Gebb, C., Baron, D. & Grisebach, H. (1975). Spectroscopic evidence for the formation of a 4-keto intermediate in the UDP-apiose/UDP-xylose synthase reaction. *European journal of biochemistry / FEBS* **54**, 493-8.
52. Guyett, P., Glushka, J., Gu, X. & Bar-Peled, M. (2009). Real-time NMR monitoring of intermediates and labile products of the bifunctional enzyme UDP-apiose/UDP-xylose synthase. *Carbohydr Res* **344**, 1072-8.
53. O'Neill, M. A., Ishii, T., Albersheim, P. & Darvill, A. G. (2004). Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual review of plant biology* **55**, 109-39.
54. Sugahara, K. & Kitagawa, H. (2002). Heparin and heparan sulfate biosynthesis. *IUBMB Life* **54**, 163-75.
55. Moloney, D. J., Shair, L. H., Lu, F. M., Xia, J., Locke, R., Matta, K. L. & Haltiwanger, R. S. (2000). Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *The Journal of biological chemistry* **275**, 9604-11.

56. Hase, S., Kawabata, S., Nishimura, H., Takeya, H., Sueyoshi, T., Miyata, T., Iwanaga, S., Takao, T., Shimonishi, Y. & Ikenaka, T. (1988). A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. *Journal of biochemistry* **104**, 867-8.
57. Nishimura, H., Kawabata, S., Kisiel, W., Hase, S., Ikenaka, T., Takao, T., Shimonishi, Y. & Iwanaga, S. (1989). Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl2-Glc) O-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. *The Journal of biological chemistry* **264**, 20320-5.
58. Hill, A. V. (1910). The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *The Journal of physiology* **40**, iv-vii.
59. Suzuki, K., Suzuki, Y. & Kitamura, S. (2003). Cloning and expression of a UDP-glucuronic acid decarboxylase gene in rice. *J Exp Bot* **54**, 1997-9.
60. Eimert, K., Villand, P., Kilian, A. & Kleczkowski, L. A. (1996). Cloning and characterization of several cDNAs for UDP-glucose pyrophosphorylase from barley (Hordeum vulgare) tissues. *Gene* **170**, 227-32.
61. Reiter, W. D. & Vanzin, G. F. (2001). Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant molecular biology* **47**, 95-113.
62. Jayasinghe, S., Hristova, K. & White, S. H. (2001). Energetics, stability, and prediction of transmembrane helices. *Journal of molecular biology* **312**, 927-34.
63. Parks, G. D. & Lamb, R. A. (1991). Topology of eukaryotic type II membrane proteins: importance of N-terminal positively charged residues flanking the hydrophobic domain. *Cell* **64**, 777-87.
64. Vertel, B. M., Walters, L. M., Flay, N., Kearns, A. E. & Schwartz, N. B. (1993). Xylosylation is an endoplasmic reticulum to Golgi event. *The Journal of biological chemistry* **268**, 11105-12.
65. Hwang, H. Y. & Horvitz, H. R. (2002). The Caenorhabditis elegans vulval morphogenesis gene sqv-4 encodes a UDP-glucose dehydrogenase that is temporally and spatially regulated. *Proc Natl Acad Sci U S A* **99**, 14224-9.
66. Griffith, C. L., Klutts, J. S., Zhang, L., Levery, S. B. & Doering, T. L. (2004). UDP-glucose dehydrogenase plays multiple roles in the biology of the pathogenic fungus Cryptococcus neoformans. *The Journal of biological chemistry* **279**, 51669-76.
67. Oka, T. & Jigami, Y. (2006). Reconstruction of de novo pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in Saccharomyces cerevisiae. *Febs J* **273**, 2645-57.
68. Herman, T., Hartwieg, E. & Horvitz, H. R. (1999). sqv mutants of Caenorhabditis elegans are defective in vulval epithelial invagination. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 968-73.
69. Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J. & Ghosh, D. (1995). Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34**, 6003-13.
70. Jornvall, H., Persson, M. & Jeffery, J. (1981). Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 4226-30.
71. Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Current opinion in structural biology* **5**, 775-83.

72. Kallberg, Y., Oppermann, U., Jornvall, H. & Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). *European journal of biochemistry / FEBS* **269**, 4409-17.
73. Kavanagh, K. L., Jornvall, H., Persson, B. & Oppermann, U. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* **65**, 3895-906.
74. Bruns, C. M., Hubatsch, I., Ridderstrom, M., Mannervik, B. & Tainer, J. A. (1999). Human glutathione transferase A4-4 crystal structures and mutagenesis reveal the basis of high catalytic efficiency with toxic lipid peroxidation products. *J Mol Biol* **288**, 427-39.
75. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-637.
76. Hayward, S. & Berendsen, H. J. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144-54.
77. Ghosh, D., Weeks, C. M., Grochulski, P., Duax, W. L., Erman, M., Rimsay, R. L. & Orr, J. C. (1991). Three-dimensional structure of holo 3 alpha,20 beta-hydroxysteroid dehydrogenase: a member of a short-chain dehydrogenase family. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 10064-8.
78. Field, R. A. & Naismith, J. H. (2003). Structural and mechanistic basis of bacterial sugar nucleotide-modifying enzymes. *Biochemistry* **42**, 7637-47.
79. Jornvall, H., Hoog, J. O. & Persson, B. (1999). SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS letters* **445**, 261-4.
80. Kallberg, Y., Oppermann, U., Jornvall, H. & Persson, B. (2002). Short-chain dehydrogenase/reductase (SDR) relationships: a large family with eight clusters common to human, animal, and plant genomes. *Protein science : a publication of the Protein Society* **11**, 636-41.
81. Frey, P. A. (1996). The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **10**, 461-70.
82. Allard, S. T., Giraud, M. F. & Naismith, J. H. (2001). Epimerases: structure, function and mechanism. *Cellular and molecular life sciences : CMLS* **58**, 1650-65.
83. Holden, H. M., Rayment, I. & Thoden, J. B. (2003). Structure and function of enzymes of the Leloir pathway for galactose metabolism. *The Journal of biological chemistry* **278**, 43885-8.
84. Timson, D. J. (2006). The structural and molecular biology of type III galactosemia. *IUBMB Life* **58**, 83-9.
85. Bauer, A. J., Rayment, I., Frey, P. A. & Holden, H. M. (1992). The molecular structure of UDP-galactose 4-epimerase from Escherichia coli determined at 2.5 A resolution. *Proteins* **12**, 372-81.
86. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L. & Holden, H. M. (2000). Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase. *Biochemistry* **39**, 5691-701.

87. Thoden, J. B. & Holden, H. M. (2005). The molecular architecture of galactose mutarotase/UDP-galactose 4-epimerase from Saccharomyces cerevisiae. *The Journal of biological chemistry* **280**, 21900-7.
88. Brahma, A. & Bhattacharyya, D. (2004). UDP-galactose 4-epimerase from Kluyveromyces fragilis. Evidence for independent mutarotation site. *European journal of biochemistry / FEBS* **271**, 58-68.
89. Bellamacina, C. R. (1996). The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *Faseb J* **10**, 1257-69.
90. Thoden, J. B., Frey, P. A. & Holden, H. M. (1996). Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose 4-epimerase from Escherichia coli: implications for the catalytic mechanism. *Biochemistry* **35**, 5137-44.
91. Wang, S. F. & Gabriel, O. (1970). Biological mechanisms involved in the formation of deoxysugars. VI. Role and function of enzyme-bound nicotinamide adenine dinucleotide in thymidine diphosphate D-glucose oxidoreductase. *The Journal of biological chemistry* **245**, 8-14.
92. Thoden, J. B., Frey, P. A. & Holden, H. M. (1996). Crystal structures of the oxidized and reduced forms of UDP-galactose 4-epimerase isolated from Escherichia coli. *Biochemistry* **35**, 2557-66.
93. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-34.
94. Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L. & Thornton, J. M. (1997). PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends in biochemical sciences* **22**, 488-90.
95. Wong, S. S. & Frey, P. A. (1977). Fluorescence and nucleotide binding properties of Escherichia coli uridine diphosphate galactose 4-epimerase: support for a model for nonstereospedific action. *Biochemistry* **16**, 298-305.
96. Thoden, J. B., Frey, P. A. & Holden, H. M. (1996). High-resolution X-ray structure of UDP-galactose 4-epimerase complexed with UDP-phenol. *Protein science : a publication of the Protein Society* **5**, 2149-61.
97. Thoden, J. B., Hegeman, A. D., Wesenberg, G., Chapeau, M. C., Frey, P. A. & Holden, H. M. (1997). Structural analysis of UDP-sugar binding to UDP-galactose 4-epimerase from Escherichia coli. *Biochemistry* **36**, 6294-304.
98. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L. & Holden, H. M. (2001). Human UDP-galactose 4-epimerase. Accommodation of UDP-N-acetylglucosamine within the active site. *The Journal of biological chemistry* **276**, 15131-6.
99. Konopka, J. M., Halkides, C. J., Vanhooke, J. L., Gorenstein, D. G. & Frey, P. A. (1989). UDP-galactose 4-epimerase. Phosphorus-31 nuclear magnetic resonance analysis of NAD+ and NADH bound at the active site. *Biochemistry* **28**, 2645-54.
100. Burke, J. R. & Frey, P. A. (1993). The importance of binding energy in catalysis of hydride transfer by UDP-galactose 4-epimerase: a 13C and 15N NMR and kinetic study. *Biochemistry* **32**, 13220-30.
101. Swanson, B. A. & Frey, P. A. (1993). Identification of lysine 153 as a functionally important residue in UDP-galactose 4-epimerase from Escherichia coli. *Biochemistry* **32**, 13231-6.

102. Liu, Y., Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Ruzicka, F. J., Holden, H. M. & Frey, P. A. (1997). Mechanistic roles of tyrosine 149 and serine 124 in UDP-galactose 4-epimerase from Escherichia coli. *Biochemistry* **36**, 10675-84.
103. Wei, Y., Lin, J. & Frey, P. A. (2001). 13C NMR analysis of electrostatic interactions between NAD+ and active site residues of UDP-galactose 4-epimerase: implications for the activation induced by uridine nucleotides. *Biochemistry* **40**, 11279-87.
104. Arabshahi, A., Flentke, G. R. & Frey, P. A. (1988). Uridine diphosphate galactose 4-epimerase. pH dependence of the reduction of NAD+ by a substrate analog. *The Journal of biological chemistry* **263**, 2638-43.
105. Berger, E., Arabshahi, A., Wei, Y., Schilling, J. F. & Frey, P. A. (2001). Acid-base catalysis by UDP-galactose 4-epimerase: correlations of kinetically measured acid dissociation constants with thermodynamic values for tyrosine 149. *Biochemistry* **40**, 6699-705.
106. Thoden, J. B. & Holden, H. M. (1998). Dramatic differences in the binding of UDP-galactose and UDP-glucose to UDP-galactose 4-epimerase from Escherichia coli. *Biochemistry* **37**, 11469-77.
107. Nayar, S. & Bhattacharyya, D. (1997). UDP-galactose 4-epimerase from Escherichia coli: existence of a catalytic monomer. *FEBS letters* **409**, 449-51.
108. Bhattacharyya, D. (1993). Reversible folding of UDP-galactose-4-epimerase from yeast Kluyveromyces fragilis. *Biochemistry* **32**, 9726-34.
109. Darrow, R. A. & Rodstrom, R. (1970). Uridine diphosphate galactose 4-epimerase from yeast. Studies on the relationship between quaternary structure and catalytic activity. *The Journal of biological chemistry* **245**, 2036-42.
110. Allard, S. T., Giraud, M. F., Whitfield, C., Graninger, M., Messner, P. & Naismith, J. H. (2001). The crystal structure of dTDP-D-Glucose 4,6-dehydratase (RmlB) from Salmonella enterica serovar Typhimurium, the second enzyme in the dTDP-l-rhamnose pathway. *Journal of molecular biology* **307**, 283-95.
111. Vogan, E. M., Bellamacina, C., He, X., Liu, H. W., Ringe, D. & Petsko, G. A. (2004). Crystal structure at 1.8 A resolution of CDP-D-glucose 4,6-dehydratase from Yersinia pseudotuberculosis. *Biochemistry* **43**, 3057-67.
112. Somoza, J. R., Menon, S., Schmidt, H., Joseph-McCarthy, D., Dessen, A., Stahl, M. L., Somers, W. S. & Sullivan, F. X. (2000). Structural and kinetic analysis of Escherichia coli GDP-mannose 4,6 dehydratase provides insights into the enzyme's catalytic mechanism and regulation by GDP-fucose. *Structure* **8**, 123-35.
113. Daude, N., Gallaher, T. K., Zeschnigk, M., Starzinski-Powitz, A., Petry, K. G., Haworth, I. S. & Reichardt, J. K. (1995). Molecular cloning, characterization, and mapping of a full-length cDNA encoding human UDP-galactose 4'-epimerase. *Biochemical and molecular medicine* **56**, 1-7.
114. Gross, J. W., Hegeman, A. D., Vestling, M. M. & Frey, P. A. (2000). Characterization of enzymatic processes by rapid mix-quench mass spectrometry: the case of dTDP-glucose 4,6-dehydratase. *Biochemistry* **39**, 13633-40.
115. Tonetti, M., Sturla, L., Bisso, A., Zanardi, D., Benatti, U. & De Flora, A. (1998). The metabolism of 6-deoxyhexoses in bacterial and animal cells. *Biochimie* **80**, 923-31.
116. He, X., Thorson, J. S. & Liu, H. W. (1996). Probing the coenzyme and substrate binding events of CDP-D-glucose 4,6-dehydratase: mechanistic implications. *Biochemistry* **35**, 4721-31.

117. Tonetti, M., Zanardi, D., Gurnon, J. R., Fruscione, F., Armirotti, A., Damonte, G., Sturla, L., De Flora, A. & Van Etten, J. L. (2003). Paramecium bursaria Chlorella virus 1 encodes two enzymes involved in the biosynthesis of GDP-L-fucose and GDP-D-rhamnose. *The Journal of biological chemistry* **278**, 21559-65.
118. Fruscione, F., Sturla, L., Duncan, G., Van Etten, J. L., Valbuzzi, P., De Flora, A., Di Zanni, E. & Tonetti, M. (2008). Differential role of NADP+ and NADPH in the activity and structure of GDP-D-mannose 4,6-dehydratase from two chlorella viruses. *The Journal of biological chemistry* **283**, 184-93.
119. Allard, S. T., Beis, K., Giraud, M. F., Hegeman, A. D., Gross, J. W., Wilmouth, R. C., Whitfield, C., Graninger, M., Messner, P., Allen, A. G., Maskell, D. J. & Naismith, J. H. (2002). Toward a structural understanding of the dehydratase mechanism. *Structure* **10**, 81-92.
120. Adair, G. (1925). The hemoglobin system. VI. The oxygen dissociation curve of hemoglobin. *J Biol Chem* **63**, 529-545.
121. Hegeman, A. D., Gross, J. W. & Frey, P. A. (2001). Probing catalysis by Escherichia coli dTDP-glucose-4,6-dehydratase: identification and preliminary characterization of functional amino acid residues at the active site. *Biochemistry* **40**, 6598-610.
122. Snipes, C. E., Brillinger, G. U., Sellers, L., Mascaro, L. & Floss, H. G. (1977). Stereochemistry of the dTDP-glucose oxidoreductase reaction. *The Journal of biological chemistry* **252**, 8113-7.
123. Gross, J. W., Hegeman, A. D., Gerratana, B. & Frey, P. A. (2001). Dehydration is catalyzed by glutamate-136 and aspartic acid-135 active site residues in Escherichia coli dTDP-glucose 4,6-dehydratase. *Biochemistry* **40**, 12497-504.
124. Hegeman, A. D., Gross, J. W. & Frey, P. A. (2002). Concerted and stepwise dehydration mechanisms observed in wild-type and mutated Escherichia coli dTDP-glucose 4,6-dehydratase. *Biochemistry* **41**, 2797-804.
125. Koropatkin, N. M. & Holden, H. M. (2005). Structure of CDP-D-glucose 4,6-dehydratase from Salmonella typhi complexed with CDP-D-xylose. *Acta crystallographica. Section D, Biological crystallography* **61**, 365-73.
126. Allard, S. T., Giraud, M. F., Whitfield, C., Messner, P. & Naismith, J. H. (2000). The purification, crystallization and structural elucidation of dTDP-D-glucose 4,6-dehydratase (RmlB), the second enzyme of the dTDP-L-rhamnose synthesis pathway from Salmonella enterica serovar typhimurium. *Acta crystallographica. Section D, Biological crystallography* **56**, 222-5.
127. Allard, S. T., Cleland, W. W. & Holden, H. M. (2004). High resolution X-ray structure of dTDP-glucose 4,6-dehydratase from Streptomyces venezuelae. *The Journal of biological chemistry* **279**, 2211-20.
128. Yu, Y., Russell, R. N., Thorson, J. S., Liu, L. D. & Liu, H. W. (1992). Mechanistic studies of the biosynthesis of 3,6-dideoxyhexoses in Yersinia pseudotuberculosis. Purification and stereochemical analysis of CDP-D-glucose oxidoreductase. *The Journal of biological chemistry* **267**, 5868-75.
129. Rosano, C., Zuccotti, S., Sturla, L., Fruscione, F., Tonetti, M. & Bolognesi, M. (2006). Quaternary assembly and crystal structure of GDP-D-mannose 4,6 dehydratase from Paramecium bursaria Chlorella virus. *Biochemical and biophysical research communications* **339**, 191-5.

130. Mulichak, A. M., Bonin, C. P., Reiter, W. D. & Garavito, R. M. (2002). Structure of the MUR1 GDP-mannose 4,6-dehydratase from Arabidopsis thaliana: implications for ligand binding and specificity. *Biochemistry* **41**, 15578-89.
131. Webb, N. A., Mulichak, A. M., Lam, J. S., Rocchetta, H. L. & Garavito, R. M. (2004). Crystal structure of a tetrameric GDP-D-mannose 4,6-dehydratase from a bacterial GDP-D-rhamnose biosynthetic pathway. *Protein science : a publication of the Protein Society* **13**, 529-39.
132. Koropatkin, N. M., Liu, H. W. & Holden, H. M. (2003). High resolution x-ray structure of tyvelose epimerase from Salmonella typhi. *The Journal of biological chemistry* **278**, 20874-81.
133. Bisso, A., Sturla, L., Zanardi, D., De Flora, A. & Tonetti, M. (1999). Structural and enzymatic characterization of human recombinant GDP-D-mannose-4,6-dehydratase. *FEBS letters* **456**, 370-4.
134. Sturla, L., Bisso, A., Zanardi, D., Benatti, U., De Flora, A. & Tonetti, M. (1997). Expression, purification and characterization of GDP-D-mannose 4,6-dehydratase from Escherichia coli. *FEBS letters* **412**, 126-30.
135. Nakayama, K., Maeda, Y. & Jigami, Y. (2003). Interaction of GDP-4-keto-6-deoxymannose-3,5-epimerase-4-reductase with GDP-mannose-4,6-dehydratase stabilizes the enzyme activity for formation of GDP-fucose from GDP-mannose. *Glycobiology* **13**, 673-80.
136. Major, L. L., Wolucka, B. A. & Naismith, J. H. (2005). Structure and function of GDP-mannose-3',5'-epimerase: an enzyme which performs three chemical reactions at the same active site. *J Am Chem Soc* **127**, 18309-20.
137. Pan, Y., Kee, C. W., Jiang, Z., Ma, T., Zhao, Y., Yang, Y., Xue, H. & Tan, C. H. (2011). Expanding the utility of Bronsted base catalysis: biomimetic enantioselective decarboxylative reactions. *Chemistry* **17**, 8363-70.
138. Sakuraba, H., Kawai, T., Yoneda, K. & Ohshima, T. (2011). Crystal structure of UDP-galactose 4-epimerase from the hyperthermophilic archaeon Pyrobaculum calidifontis. *Archives of biochemistry and biophysics* **512**, 126-34.
139. Gatzeva-Topalova, P. Z., May, A. P. & Sousa, M. C. (2005). Crystal structure and mechanism of the Escherichia coli ArnA (PmrI) transformylase domain. An enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry* **44**, 5328-38.
140. Esko, J. D. & Selleck, S. B. (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu Rev Biochem* **71**, 435-71.
141. Fuster, M. M. & Esko, J. D. (2005). The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer* **5**, 526-42.
142. Lamoureux, F., Baud'huin, M., Duplomb, L., Heymann, D. & Redini, F. (2007). Proteoglycans: key partners in bone cell biology. *Bioessays* **29**, 758-71.
143. Esko, J. D., Rostand, K. S. & Weinke, J. L. (1988). Tumor formation dependent on proteoglycan biosynthesis. *Science* **241**, 1092-6.
144. Belting, M., Borsig, L., Fuster, M. M., Brown, J. R., Persson, L., Fransson, L. A. & Esko, J. D. (2002). Tumor attenuation by combined heparan sulfate and polyamine depletion. *Proc Natl Acad Sci U S A* **99**, 371-6.

145. Bai, X., Zhou, D., Brown, J. R., Crawford, B. E., Hennet, T. & Esko, J. D. (2001). Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *J Biol Chem* **276**, 48189-95.
146. Rampal, R., Luther, K. B. & Haltiwanger, R. S. (2007). Notch signaling in normal and disease States: possible therapies related to glycosylation. *Current molecular medicine* **7**, 427-45.
147. Takeuchi, H., Fernandez-Valdivia, R. C., Caswell, D. S., Nita-Lazar, A., Rana, N. A., Garner, T. P., Weldeghiorghis, T. K., Macnaughtan, M. A., Jafar-Nejad, H. & Haltiwanger, R. S. (2011). Rumi functions as both a protein O-glucosyltransferase and a protein O-xylosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16600-5.
148. de Bernabe, D. B., Inamori, K., Yoshida-Moriguchi, T., Weydert, C. J., Harper, H. A., Willer, T., Henry, M. D. & Campbell, K. P. (2009). Loss of alpha-dystroglycan laminin binding in epithelium-derived cancers is caused by silencing of LARGE. *The Journal of biological chemistry* **284**, 11279-84.
149. Inamori, K., Yoshida-Moriguchi, T., Hara, Y., Anderson, M. E., Yu, L. & Campbell, K. P. (2012). Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. *Science* **335**, 93-6.
150. Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., Ladenstein, R., Jornvall, H. & Oppermann, U. (2002). Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *The Journal of biological chemistry* **277**, 25677-84.
151. Hald, E., Lehmann, P. & Ziegenhorn, J. (1975). Molar Absorptivities of β-NADH and β-NAD at 260 nm. *Clin Chem* **21**, 884-887.
152. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* (Walker, J. M., ed.), Vol. 18, pp. 571-607. Humana Press, Totowa, NJ.
153. Otwinowski, Z., and Minor, W. (1996). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymology* **276**, 307-326.
154. Kabsch, W. (2010). Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32.
155. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-21.
156. Brunger, A. T. (1997). Free R Value: Cross-validation in Crystallography. *Methods in Enzymology* **277 B**, 366-396.
157. Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-3.
158. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.
159. Diederichs, K. & Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography [published erratum appears in Nat Struct Biol 1997 Jul;4(7):592]. *Nat Struct Biol* **4**, 269-75.

160. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology* **215**, 403-10.
161. Holm, L. & Park, J. (2000). DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-7.
162. Richards, F. M. & Kundrot, C. E. (1988). Identification of structural motifs from protein coordinate data: secondary structure and first-level supersecondary structure. *Proteins* **3**, 71-84.
163. Frieden, C. (1970). Kinetic aspects of regulation of metabolic processes. The hysteretic enzyme concept. *The Journal of biological chemistry* **245**, 5788-99.
164. Zhang, X. J., Wozniak, J. A. & Matthews, B. W. (1995). Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. *Journal of molecular biology* **250**, 527-52.
165. Dickerson, R. E., Goodsell, D. S. & Neidle, S. (1994). "...the tyranny of the lattice...". *Proceedings of the National Academy of Sciences of the United States of America* **91**, 3579-83.
166. Wu, Y., Lai, D. K. W. & Houk, K. N. (1995). Transition structures of hydride transfer reactions of protonated pyridinium ion with 1,4-dihydropyridine and protonated nicotinamide with 1,4-dihydronicotinamide. *Journal of the American Chemical Society* **117**, 4100-08.
167. Mesecar, A. D., Stoddard, B. L. & Koshland, D. E., Jr. (1997). Orbital steering in the catalytic power of enzymes: small structural changes with large catalytic consequences. *Science* **277**, 202-6.
168. Fraaije, M. W. & Mattevi, A. (2000). Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem Sci* **25**, 126-32.
169. Naismith, J. H., Emmerich, C., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Raftery, J., Kalb, A. J. & Yariv, J. (1994). Refined structure of concanavalin A complexed with methyl alpha-D-mannopyranoside at 2.0 A resolution and comparison with the saccharide-free structure. *Acta Crystallogr D Biol Crystallogr* **50**, 847-58.
170. Bouckaert, J., Hamelryck, T. W., Wyns, L. & Loris, R. (1999). The crystal structures of Man(alpha1-3)Man(alpha1-O)Me and Man(alpha1-6)Man(alpha1-O)Me in complex with concanavalin A. *J Biol Chem* **274**, 29188-95.
171. Kadirvelraj, R., Foley, B. L., Dyekjaer, J. D. & Woods, R. J. (2008). Involvement of water in carbohydrate-protein binding: concanavalin A revisited. *Journal of the American Chemical Society* **130**, 16933-42.
172. Xu, G., Lehmann, R., Schleicher, E., Haring, H. U. & Liebich, H. (1998). Advantages in the analysis of UDP-sugars by capillary electrophoresis-comparison of the conventional HPLC method with two new capillary electrophoretic micro-procedures. *Biomed Chromatogr* **12**, 113-5.
173. Hynes, R. O. (2009). The extracellular matrix: not just pretty fibrils. *Science* **326**, 1216-9.
174. Gotting, C., Kuhn, J., Zahn, R., Brinkmann, T. & Kleesiek, K. (2000). Molecular cloning and expression of human UDP-d-Xylose:proteoglycan core protein beta-d-xylosyltransferase and its first isoform XT-II. *Journal of molecular biology* **304**, 517-28.
175. Kitagawa, H., Tone, Y., Tamura, J., Neumann, K. W., Ogawa, T., Oka, S., Kawasaki, T. & Sugahara, K. (1998). Molecular cloning and expression of glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *The Journal of biological chemistry* **273**, 6615-8.

176. Polizzi, S. J., Walsh, R. M., Jr., Peeples, W. B., Lim, J. M., Wells, L. & Wood, Z. A. (2012). Human UDP-alpha-D-xylose synthase and E. coli ArnA conserve a conformational shunt that controls whether xylose or 4-keto-xylose is produced. *Biochemistry*.
177. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical ultracentrifugation in biochemistry and polymer science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds.), pp. 90-125. The Royal Society of Chemistry, Cambridge.
178. Brown, P. H. & Schuck, P. (2006). Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* **90**, 4651-61.
179. Dam, J., Velikovsky, C. A., Mariuzza, R. A., Urbanke, C. & Schuck, P. (2005). Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions c(s). *Biophys J* **89**, 619-34.
180. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97.
181. Garcia De La Torre, J., Huertas, M. L. & Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J* **78**, 719-30.
182. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* **36**, 1277-1282.
183. Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. (2007). ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. *Journal of Applied Crystallography* **40**, s223-s228.
184. Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography* **25**, 495-503.
185. Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal* **76**, 2879-86.
186. Volkov, V. V. & Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **36**, 860-864.
187. Petoukhov, M. V. & Svergun, D. I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophysical journal* **89**, 1237-50.
188. Schneidman-Duhovny, D., Hammel, M. & Sali, A. (2010). FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic acids research* **38**, W540-4.
189. Svergun, D., Barberato, C. & Koch, M. H. J. (1995). CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* **28**, 768-773.
190. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605-12.
191. Schuck, P. (2003). On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**, 104-24.

192. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly reviews of biophysics* **40**, 191-285.
193. Grant, T. D., Luft, J. R., Wolfley, J. R., Tsuruta, H., Martel, A., Montelione, G. T. & Snell, E. H. (2011). Small angle X-ray scattering as a complementary tool for high-throughput structural studies. *Biopolymers* **95**, 517-30.
194. Phillip, Y., Sherman, E., Haran, G. & Schreiber, G. (2009). Common crowding agents have only a small effect on protein-protein interactions. *Biophysical journal* **97**, 875-85.
195. Chebotareva, N. A., Andreeva, I. E., Makeeva, V. F., Livanova, N. B. & Kurganov, B. I. (2004). Effect of molecular crowding on self-association of phosphorylase kinase and its interaction with phosphorylase b and glycogen. *Journal of molecular recognition : JMR* **17**, 426-32.
196. Gilbert, G. A. & Jenkins, R. C. (1956). Boundary problems in the sedimentation and electrophoresis of complex systems in rapid reversible equilibrium. *Nature* **177**, 853-4.
197. Brautigam, C. A. (2011). Using Lamm-Equation modeling of sedimentation velocity data to determine the kinetic and thermodynamic properties of macromolecular interactions. *Methods* **54**, 4-15.
198. Clackson, T. & Wells, J. A. (1995). A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383-6.
199. Bruice, T. C. & Pandit, U. K. (1960). Intramolecular Models Depicting the Kinetic Importance of "Fit" in Enzymatic Catalysis. *Proceedings of the National Academy of Sciences of the United States of America* **46**, 402-4.
200. Helms, V. (2002). Attraction within the membrane. Forces behind transmembrane protein folding and supramolecular complex assembly. *EMBO reports* **3**, 1133-8.
201. Zuber, C., Fan, J. Y., Guhl, B., Parodi, A., Fessler, J. H., Parker, C. & Roth, J. (2001). Immunolocalization of UDP-glucose:glycoprotein glucosyltransferase indicates involvement of pre-Golgi intermediates in protein quality control. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10710-5.
202. Kabuss, R., Ashikov, A., Oelmann, S., Gerardy-Schahn, R. & Bakker, H. (2005). Endoplasmic reticulum retention of the large splice variant of the UDP-galactose transporter is caused by a dilysine motif. *Glycobiology* **15**, 905-11.
203. Muraoka, M., Miki, T., Ishida, N., Hara, T. & Kawakita, M. (2007). Variety of nucleotide sugar transporters with respect to the interaction with nucleoside mono- and diphosphates. *The Journal of biological chemistry* **282**, 24615-22.
204. Zähringer, U., Lindner, B. & Rietschel, E. T. (1999). Chemical structure of Lipid A: Recent advances in structural analysis of biologically active molecules. In *Endotoxin in Health and Disease* (Brade, H., Opal, S. M., Vogel, S. N. & Morrison, D. C., eds.), pp. 93-114. Marcel Dekker, Inc., New York.
205. Edwards, T. E., Abendroth, J. & Sankaran, B. Seattle Structural Genomics Center for Infectious Disease (SSGCID). Crystal structure of PbgP3 protein from Burkholderia pseudomallei (unpublished).
206. Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M. & Aksoy, S. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia. *Nature genetics* **32**, 402-7.

207. Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottier, P., Camus, J. C., Cattolico, L., Chandler, M., Choisne, N., Claudel-Renard, C., Cunnac, S., Demange, N., Gaspin, C., Lavie, M., Moisan, A., Robert, C., Saurin, W., Schiex, T., Siguier, P., Thebault, P., Whalen, M., Wincker, P., Levy, M., Weissenbach, J. & Boucher, C. A. (2002). Genome sequence of the plant pathogen Ralstonia solanacearum. *Nature* **415**, 497-502.
208. Chakrabartty, A., Kortemme, T. & Baldwin, R. L. (1994). Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci* **3**, 843-52.
209. Pace, C. N. & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* **75**, 422-7.

APPENDIX A

UDP XYLOSE SYNTHASE 1 IS REQUIRED FOR MORPHOGENESIS AND HISTOGENESIS OF THE CRANIOFACIAL SKELETON[3]

[3] Reprinted from *Developmental Biology*, Vol 341, Eames, B. F., Singer, A., Smith, G. A., Wood, Z. A., Yan, Y. L., He, X., Polizzi, S. J., Catchen, J. M., Rodriguez-Mari, A., Linbo, T., Raible, D. W., and Postlethwait, J. H. Pages 400-415, Copyright (2010), with permission from the Society for Developmental Biology.

S.J. Polizzi contributed to the catalytic analysis of human UDP xylose synthase 1.

Abstract

UDP-xylose synthase (Uxs1) is strongly conserved from bacteria to humans, but because no mutation has been studied in any animal, we do not understand its roles in development. Furthermore, no crystal structure has been published. Uxs1 synthesizes UDP-xylose, which initiates glycosaminoglycan attachment to a protein core during proteoglycan formation. Crystal structure and biochemical analyses revealed that an R233H substitution mutation in zebrafish *uxs1* alters an arginine buried in the dimer interface, thereby destabilizing and, as enzyme assays show, inactivating the enzyme. Homozygous *uxs1* mutants lack Alcian blue-positive, proteoglycan-rich extracellular matrix in cartilages of the neurocranium, pharyngeal arches, and pectoral girdle. Transcripts for *uxs1* localize to skeletal domains at hatching. GFP-labeled neural crest cells revealed defective organization and morphogenesis of chondrocytes, perichondrium, and bone in *uxs1* mutants. Proteoglycans were dramatically reduced and defectively localized in *uxs1* mutants. Although *col2a1a* transcripts over-accumulated in *uxs1* mutants, diminished quantities of Col2a1 protein suggested a role for proteoglycans in collagen secretion or localization. Expression of *col10a1, indian hedgehog,* and *patched* were disrupted in mutants, reflecting improper chondrocyte/perichondrium signaling. Up-regulation of *sox9a, sox9b,* and *runx2b* in mutants suggested a molecular mechanism consistent with a role for proteoglycans in regulating skeletal cell fate. Together, our data reveal time-dependent changes to gene expression in *uxs1* mutants that support a signaling role for proteoglycans during at least two distinct phases of skeletal development. These investigations are the first to examine the effect of mutation on the structure and function of Uxs1 protein in any vertebrate embryos, and reveal that Uxs1 activity is essential for the production and organization of skeletal extracellular matrix, with consequent effects on cartilage, perichondral, and bone morphogenesis.

Introduction

The vertebrate skeleton provides structural support for muscle attachments and a protective casing for vulnerable internal organs. These functions rely on the coordinated secretion of dense extracellular matrix (ECM) by skeletal precursor cells during embryonic development. Proteinaceous components of skeletal ECM include collagens, elastin, and proteoglycans. Collagens anchor and reinforce the ECM; elastin provides flexibility (Velleman, 2000); and proteoglycans impact cell division, cell adhesion, and migration (Holt and Dickson, 2005; Kirn-Safran et al., 2004; Lander and Selleck, 2000 Knudson, 2001). Proteoglycans contain repeating disaccharides (glycosaminoglycans, or GAGs) linked to a protein core (Prydz and Dalen, 2000) and include four major classes: dermatan, keratan, chondroitin, and heparan sulfate. In addition to imparting hydrostatic properties to skeletal tissues through GAG sulfation and hydration, proteoglycans can play a role in cell signaling. For example, heparan sulfate proteoglycans (HSPGs) help cell receptors bind growth factors (Izvolsky et al., 2003; Lin et al., 1999) and although the mechanism remains unclear, *Cspg1* (*Aggrecan*)-deficient mutant chickens have dwarfed bones (Velleman and Clark, 1992).

Not all bones are created equal. Dermal bones differentiate osteoblasts directly via intramembranous ossification, but chondral bones form by endochondral ossification, during which developing chondrocytes and an overlying osteogenic epithelium, the perichondrium, interact (Eames et al., 2003). Understanding how these cell types signal each other and how proteoglycans play structural roles is important because impaired signaling between chondrocytes and osteoprogenitors can lead to osteoarthritis, a disease in which bone spurs replace cartilage in many people over age 65 (Ala-Kokko et al., 1990; Kizawa et al., 2005; Knowlton et al., 1990; Rothschild and Panza, 2007).

Proteoglycan biosynthesis initiates with the addition of a common tetrasaccharide linker to a core protein. Vertebrates use UDP-xylose, the first sugar in the linker, almost exclusively for proteoglycan synthesis (xylose is also added to EGF-repeat domains of some proteins (Bakker et al., 2009; Ishimizu et al., 2007)). UDP-xylose biosynthesis begins with the conversion of UDP-glucose into UDP-glucuronic acid by UDP-glucose dehydrogenase (Ugdh). Zebrafish with diminished Ugdh activity have defective craniofacial and coronary development (Neuhauss et al., 1996; Walsh and Stainier, 2001). Next, UDP-xylose synthase (Uxs1, also called UDP-glucuronic acid decarboxylase or UGD, EC 4.1.1.35) converts UDP-glucuronic acid into UDP-xylose (Kearns et al., 1993; Vertel et al., 1993). The GAG tetrasaccharide linker of proteoglycans is initiated by a xylosyltransferase, which adds UDP-xylose to a serine residue of the core protein. Galactose and glucuronic acid transferases then add two galactoses and one glucuronic acid, completing the tetrasaccharide linker. GAG synthesis continues as exostosins (Ext1a, Ext1b, Ext1c, Ext2, Extl2, and Extl3 in zebrafish) add disaccharide constituents (Kjellen and Lindahl, 1991; Knudson and Knudson, 2001; Lin, 2004). Zebrafish mutations in *ext2* (*dackel* (*dak*)), *extl3* (*boxer* (*box*)), and *solute carrier family 35, member b2* (*pinscher* (*pic*); previously termed *3'-phosphoadenosine 5'-phosphosulfate transporter 1*, or *papst1*) show that GAG synthesis and sulfation are important for axon sorting and cartilage morphogenesis (Clement et al., 2008; Lee et al., 2004; Schilling et al., 1996). The developmental roles of Uxs1, however, are poorly understood because vertebrate models that lack Uxs1 activity have not yet been investigated.

In a mutation screen for neural crest defects, we identified *man o'war* (*mow*), which, like the *sox9a* mutation *jellyfish* (Yan et al., 2002), fails to form craniofacial cartilages. Our molecular genetic analyses showed that the mow^{w60} mutation causes an amino acid replacement in the zebrafish *uxs1* gene and is allelic to the viral insert *hi3357* (Amsterdam et al., 2004;

Golling et al., 2002; Nissen et al., 2006). Sequence alignments reveal Uxs1 to be one of the most highly conserved non-mitochondrial proteins, preserving 57% amino acid identity between thebacterium *Rhodospirillum rubrum* and human. Using the crystal structure of human UXS1, we modeled the *mow*^{w60} substitution and found it to disrupt interactions at the enzyme's dimer interface, which should reduce or eliminate enzymatic activity. Indeed, our biochemical analyses revealed that an amino acid replacement homologous to the *mow*^{w60} allele destroys human UXS1 activity. Transcripts of *uxs1* are deposited maternally, and then *uxs1* is expressed zygotically in regions of the developing craniofacial skeleton. Histochemical and immunohistochemical investigations showed that wild-type *uxs1* is essential for the production and organization of many components of the ECM, including both proteoglycans and collagens. Additionally, confocal microscopy of GFP-labeled cranial neural crest cells revealed a critical role for *uxs1* in directing the morphology of chondrocytes, perichondrium, and bone during craniofacial development. Our molecular analyses demonstrated that chondrocyte maturation and Hedgehog signaling is dependent upon *uxs1*. Finally, our observation that the early chondrogenic markers *sox9a, sox9b,* and *runx2b* were up-regulated in *uxs1* mutants provide a mechanistic explanation for many of the defects in endochondral ossification, and furthermore suggest a novel feedback role for proteoglycans as skeletal progenitor cells undergo differentiation. Specifically, our data show proteoglycans to modify signaling pathways in early chondrogenic condensations and in later interactions between chondrocytes and perichondrium.

Materials and Methods

Mapping and cloning of mow^{w60}

Adult male AB zebrafish (*Danio rerio*) were treated with ethylnitrosourea to induce point

mutations and out-crossed to wild-type females. F2 families were produced and F3 larvae were screened for mutant phenotypes. The *mow*^{w60} allele showed reduced pharyngeal cartilages. For mapping, heterozygous *mow*^{w60} fish on an AB background were mated to WIK wild-type fish. F2 individuals were genotyped for 311 well-distributed simple sequence repeats (Knapik et al., 1998; Shimoda et al., 1999), identifying linked zmarker z3124. The zebrafish genome sequence nearby was screened for candidate genes involved in skeletal differentiation and mapping primers were designed in candidate genes. For *uxs1,* primers were designed to amplify a simple sequence repeat (SSR) in intron-7 (scaff346.117+GCAGCGTGAAAAAGCAAAGAC and scaff346.524-ACCGCCGCCTGTGACGA). cDNA for *uxs1* was amplified and isolated for sequencing using overlapping fragments amplified by primer sets designed from NM_173242 (Uxs1.114+TGACCGTTGGACAAGGGAGGATTTA, Uxs1.421-CTATTTGAAGAGCGGCTG CACGACTAT; Uxs1.309+AGCCGAAAATAAACTGCCCAGACTACTT, Uxs1.594- CATCC GCATCATCCTCCAGCACAC; Uxs1.394+CATAGTCGTGCAGCCGCTCTTCAAAT, Uxs1.757-GTCCCACTGCCTCATCTATCCTCTGCTC; Uxs1.854+TCACCGGTGGGGCAGG ATTC, Uxs1.1397-ACCACTCGCCCGTCGTTCAT; Uxs1.945+CGGCCGCAAGCGCAATGT AGA, Uxs1.1340- ACTCGCACCTCCACTCCTTCCTGTTTC; Uxs1.1251+TGGTCCCCGGG CCTGTTATGATG, Uxs1.1685-AGTTTGGCCCTGCGGATGTCG). To genotype *mow*^{w60} fish, we used primers MOW.928+CACCCCCAAAATGAGGACTACTG and MOW.1277-AGAGCT CGCAACGGCATAAGAT, which amplify a 349 bp fragment that yields 274 and 75 bp fragments from the *mow*^{w60} mutant amplicon but leaves the wild-type amplicon intact after digestion by *NspI*. To detect the *hi3357* pro-viral insertion, we used primers that flank the insertion site (Uxs1.e1.398+GTCGTGCAGCCGCTCTTCAAAT and Uxs1.e1.597-GCTCATC CGCATCATCCTCCAG) and yield a 199 bp fragment from wild type but no fragment from

homozygous *hi3357* mutants; a *wnt5a* amplicon verified DNA quality (Golling et al., 2002). All work with animals was approved by the appropriate Institutional Animal Care and Use Committee.

Sequence alignments

Uxs1 sequences: human, *Homo sapiens* NP_079352 (179/311, 57% identity to the bacterial protein); mouse, *Mus musculus* NP_080706 (179/311, 57%); chicken *Gallus gallus* XP_416926 (180/311, 57%); frog *Xenopus tropicalis* NP_001006849 (178/311, 57%); pufferfish *Tetraodon nigroviridis* CAG05807 (145/252, 57%); zebrafish *Danio rerio* NP_775349 (178/311, 57%); beetle *Tribolium castaneum* XP_969232 (180/305, 59%); fruitfly *Drosophila melanogaster* NP_648182 (186/311, 59%); fungus *Cryptococcus neoformans* XP_572003 (176/316, 55%); rice *Oryza sativa* EAY89464 (192/304, 63%); mustard *Arabidopsis thaliana* NP_180443 (192/305, 62%); bacterium *Rhodospirillum rubrum* YP_428334 (311/311, 100%). Tgds sequences: human *Homo sapiens* NP_055120; mouse *Mus musculus* EDL00567; chicken *Gallus gallus* XP_416988; frog *Xenopus laevis* NP_001088301; zebrafish *Danio rerio* NP_956111; mustard *Arabidopsis thaliana* NP_564633; rice *Oryza sativa* NP_001049724; bacterium *Rhodospirillum rubrum* YP_425086. Clustal-X alignment available on request; sequences were trimmed to include only unambiguously aligned sequences.

Comparison of human and R. rubrum *proteomes*

We downloaded the human proteome from Ensembl (Birney et al., 2004; Kasprzyk et al., 2004) using NCBI v36 of the human genome obtained from Ensembl version 41. From the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/), we obtained the

genome sequence of the bacterium *Rhodospirillum rubrum*, strain ATCC 11170. To rank the conservation of Uxs1 relative to other proteins, we performed a BLASTp search using each human protein as a query against a database of all *R. rubrum* proteins using WU-BLAST (http://blast.wustl.edu/) with the BLOSUM62 substitution matrix (Henikoff and Henikoff, 1992) and recorded only BLAST hits with an E-value less than 1 × 10−5. We used a gap opening penalty of 11 and a gap extension penalty of 1.

Purification of wild-type human UDP-xylose synthase

Human *UXS1* cDNA was purchased from ATCC® (catalog number 10658721). For protein expression, a construct lacking the 84 residue N-terminal transmembrane domain was cloned into a Pet-15b expression vector (Invitrogen) that was modified to include a tobacco etch virus protease cleavage site in place of the thrombin site following the N-terminal 6-His tag. The 84-residue truncation was necessary to express soluble protein, and was based on the deposited crystal structure of human UXS1 in complex with UDP (PDB entry 2B69). UXS1 expression was induced from 4 liters of logarithmically growing cells (OD600 ~1.0) after reducing temperature to 20° and adding IPTG (Sigma-Aldrich) to a final concentration of 0.1 mM. Cells were harvested after 6 hours and lysed by sonication. The His-tagged UXS1 was purified with Talon(R) affinity resin (Clontech) and eluted with 250 mM imidazole (Sigma-Aldrich). The His tag was removed with a codon- optimized tobacco etch virus protease (van den Berg et al., 2006). The cleaved protein was dialyzed into 20mM Tris pH 8.0, 250 mM NaCl and concentrated to 11 mg/ml.

Generation of human UDP-xylose synthase R236H

A human wild-type *UXS1* clone (above) was mutagenized with QuikChange site-directed mutagenesis kit (Agilent Technologies-Stratagene). DNA sequencing confirmed the human *UXS1* clone mutation, resulting in an enzyme incorporating histidine instead of arginine at amino acid position 236, equivalent to zebrafish *moww36* mutation, Uxs1 R233H. Purification proceeded as for wild-type UXS1 (above).

Determining UDP-xylose synthase activity

Human wild-type UXS1 or human UXS1 R236H (10μg) was incubated (18 hours, 37°C) with substrate UDP-glucuronic acid (1mM) in reaction buffer (100 mM sodium phosphate pH 6, 10 mM dithiothreitol, 1mM EDTA). Protein was precipitated with acetonitrile and soluble compounds were resolved by capillary zone electrophoresis for 25 minutes at 22 volts (3D-CE ChemStation, Agilent) in 50 mM sodium borate pH 9.0. Peaks were identified with known standards.

Reverse Transcriptase-PCR

Total RNA was extracted using Tri Reagent kit (Molecular Research Center Inc., Catalog # TR-118) according to manufacturer's instructions. Superscript III RNase H-reverse transcriptase (Invitrogen, #18080-044) and oligo(dT) primers synthesized first strand cDNA. Reverse transcriptase was heat inactivated and RNA was degraded with RNaseH (Biolabs, M0297S). Primers for *uxs1* were Uxs1.1251+ and Uxs1.1685- and for *actin* were as described (Krovel and Olsen, 2004).

Fluorescent cell labeling

Crosses produced homozygous *mow*^{w60} and *uxs1*^{hi3357} mutants carrying *Tg(fli1:EGFP)y1* (Lawson and Weinstein, 2002), which we imaged live by confocal microscopy. Larvae were grown in 15 mg/L PTU to minimize pigment. One day prior to observation, 200 µl of 0.3% Alizarin red was added to the water to visualize bone. Z-sections at 1.5 µm intervals were taken ventral to dorsal at 20x and 0.25µ at 60x on a Zeiss confocal microscope.

Whole-mount in situ hybridization

Expression analysis was performed as in (Jowett and Yan, 1996)). A 625 bp antisense probe in the 5' untranslated region (UTR) of *uxs1* was amplified from a 2 dpf (days post fertilization) cDNA library using primers Uxs1.1252 and Uxs1.1877. Probes for *sox9b*, *sox9a*, *ptc1*, *ptc2, runx2b*, *col2a1a*, *col10a1, runx2a, ihha, ihhb, gli3,* and *erm* were as described (Avaron et al., 2006; Yan et al., 1995; Yan et al., 2002; Yan et al., 2005).

Detection of ECM components

Alcian blue (cartilage) and Alizarin red (bone) double staining was performed as described (Walker and Kimmel, 2007). Wheat germ agglutinin (WGA) staining was performed using biotinylated succinylated WGA (Vector Laboratories, Catalog #B-1025S) as described (Lang et al., 2006). Whole-mount antibody staining for heparan sulfate (α-HS) (USBiological, Catalog# H1890), chondroitin sulfate (α-CS) (Sigma, product #C8035), and type II collagen (α-Col2a1) (Polysciences, Inc., Catalog# 23707) were conducted by fixation in 4% paraformaldehyde, dehydration with 100% methanol, rehydration with PBST followed by digestion with 0.1% trypsin at 37°C for 2 hr and 30 min for 5 dpf and 3 dpf respectively. Following two PBST rinses,

α-HS and α-Col2a1 samples were incubated with 0.5% hyaluronidase (USBiological, Catalog # D-H7981-01) in 1x PBST for 30 min at 37° C, followed by PBDT washes. Blocking solution was added to tubes containing 2% goat serum in PBDT for 2 hr. Antibodies α-HS, α-CS, and α-Col2a1 were added at 1:200, 1:100, and 1:100 dilution respectively and incubated overnight at 4°C. Secondary antibodies goat-α-rabbit-HRP for α-Col2a1, goat α-mouse-Alexa488 for α-CS and goat-α-mouse-AP for α-HS were used at 1:1000. Histochemical detection of antibody was carried out with VECTASTAIN ABC Reagent kit (Vector Laboratories, Cat. No. PK-4000) and NBT/BCIP, respectively.

Results

Chondrogenesis and osteogenesis require mow *function*

A screen of mutagenized chromosomes identified *man-o'-war* (*mow*^{w60}), a mutant phenotype that, like *jellyfish* (*sox9a*) (Piotrowski et al., 1996; Yan et al., 2002), lacks Alcian blue- positive, neural crest-derived cartilages. Compared to wild-type siblings at five days post fertilization (5 dpf), homozygous *mow*^{w60} mutants lacked tissue anterior to the eyes, had reduced lower jaw, a shortened body axis, and small pectoral fins (Fig. 5.1 A,B,E,F; and data not shown). Histological staining of skeletal tissues showed that 7 dpf homozygous *mow*^{w60} embryos lacked robust Alcian blue-positive pharyngeal and neurocranial cartilages (Fig. 5.1 C,D,G,H), including Meckel's cartilage, ceratohyal, palatoquadrate, and ceratobranchials. Alizarin red, which marks ossification centers, identified reduced bone formation in *mow*^{w60} embryos. For example, the parasphenoid, 5th ceratobranchial, and opercle showed Alizarin red staining in *mow*^{w60} embryos, but these skeletal elements were much smaller than those in wild types. Compared to wild types, homozygous *mow*^{w60} mutants also had fewer Alizarin red- positive skeletal elements, such as the

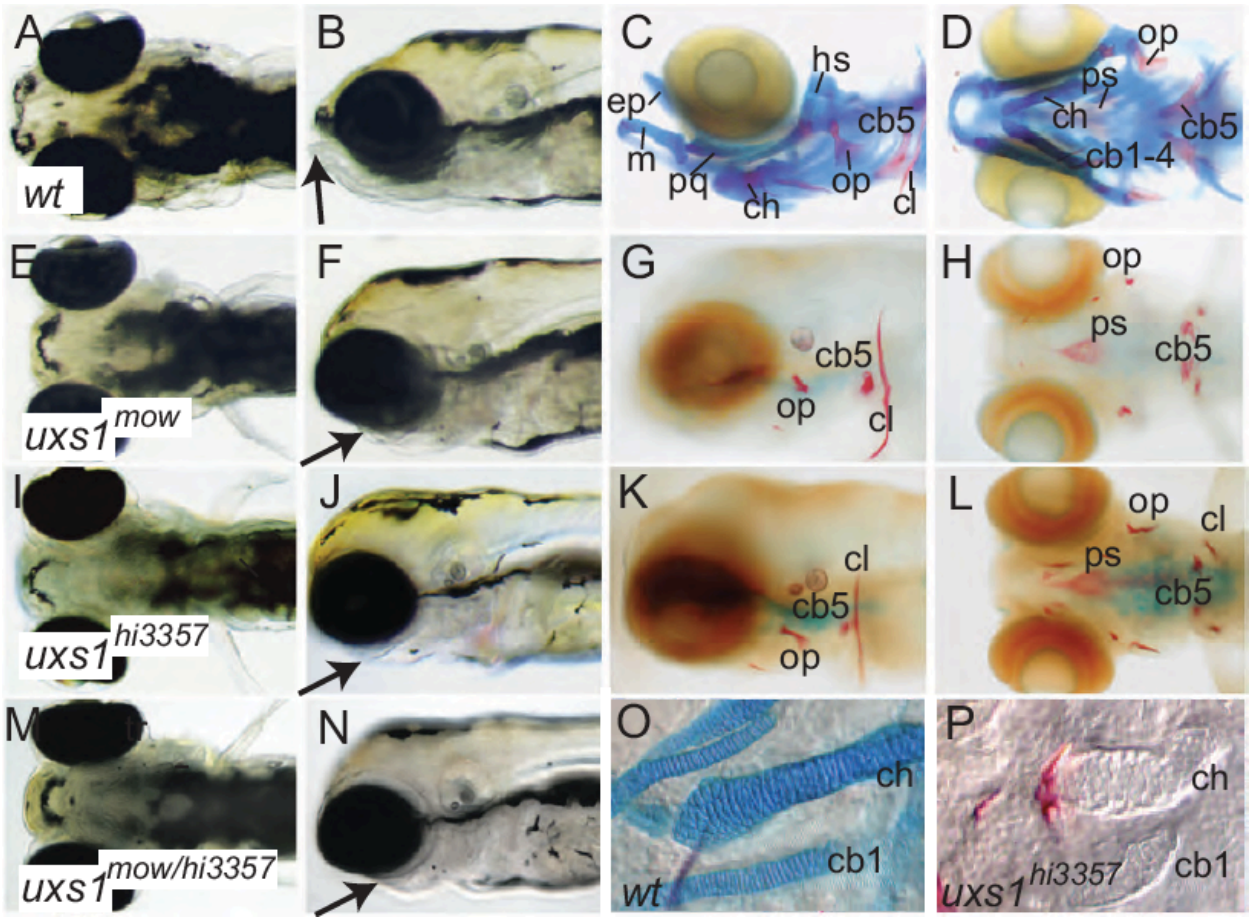


Figure 5.1. Craniofacial and skeletal phenotypes of zebrafish larvae. Ventral and lateral views of live (A,B,E,F,I,J,M,N) and Alcian blue-, Alizarin red-stained (C,D,G,H,K,L,O,P) animals. Compared to wild types (A,B), mutant animals (E,F *mow*^{w60} allele; I,J *hi3357* allele) had reduced lower jaws (arrows) at 5 dpf. Reduced lower jaw (arrow) in *mow*^{w60}*/hi3357* double heterozygotes (M,N) showed failure of complementation. Alcian blue and Alizarin red staining for cartilage (blue) and bone (red) revealed the lack of cartilage and reduced bones in mutants (G,H,K,L) compared to wild type (C,D) at 7 dpf. Nomarski optics on dissected pharyngeal skeletons suggested that mutant cartilages (P) condensed in the same areas as wild types (O), but did not secrete Alcian blue-positive matrix. Abbreviations: cb1-5, ceratobranchials 1 to 5; ch, ceratohyal; cl, cleithrum; ep, ethmoid plate; hs, hyosymplectic; m, Meckel's cartilage; op, opercle; pq, palatoquadrate; ps, parasphenoid.

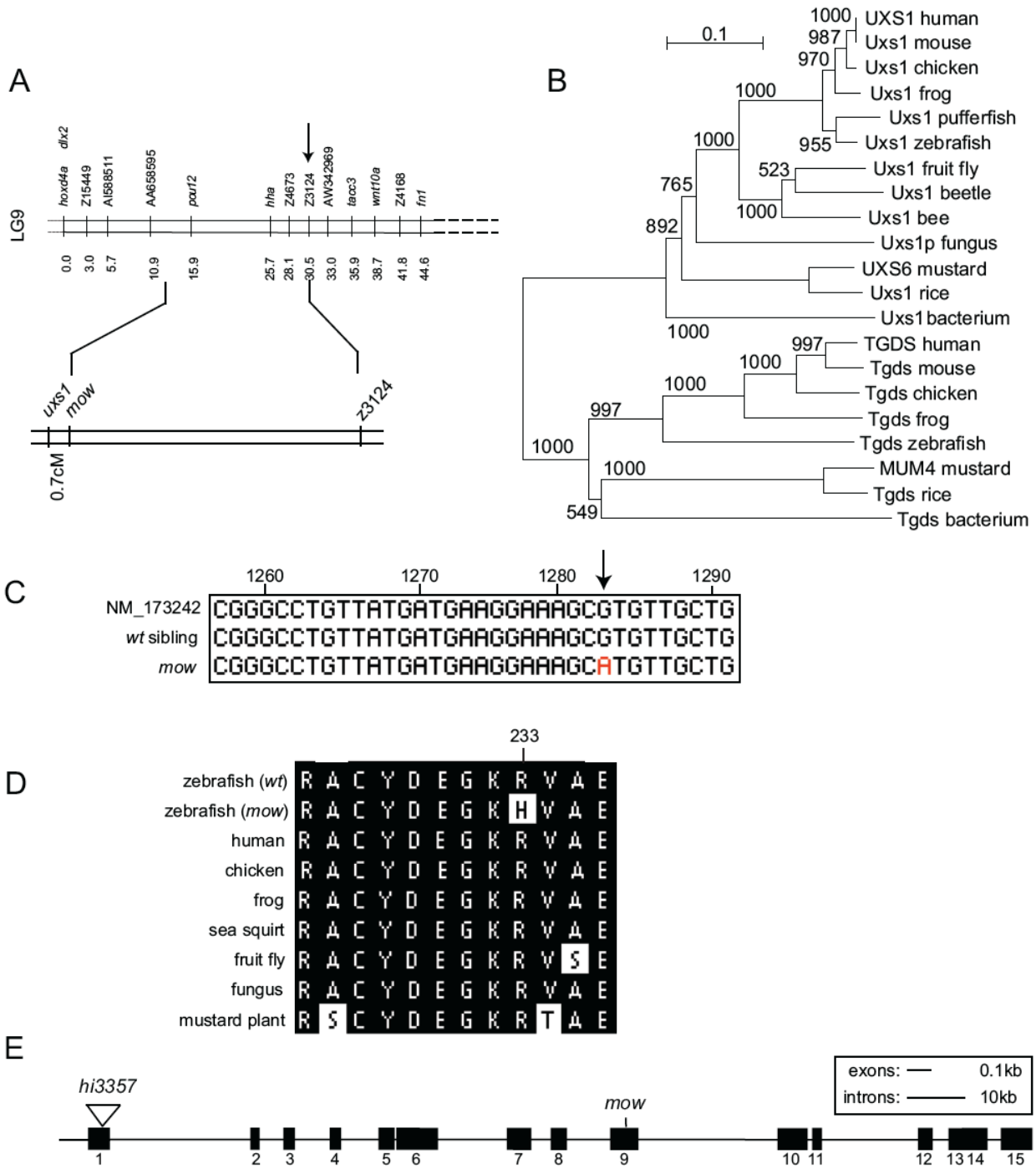
ceratohyal and hyosymplectic. Because the parasphenoid, opercle, and cleithrum develop by intramembranous ossification, and the hyosymplectic, ceratohyal, and 5th ceratobranchial develop by endochondral ossification (Cubbage and Mabee, 1996; Renn et al., 2006), both mechanisms of bone formation appear to depend upon wild-type *mow*^{w60} function.

Our mapping data (see Fig. 5.2) suggested that *mow*^{w60} and the *hi3357* retroviral insertion mutation (Amsterdam et al., 2004; Golling et al., 2002; Nissen et al., 2006) could be alleles, so we performed similar gross morphological and histological analyses of *hi3357.* Indeed, we found that *hi3357* shared the same morphological and histological defects as observed in *mow*^{w60} (Fig. 5.1 I-L). Crossing a heterozygous *mow*^{w60} male to a heterozygous *hi3357* female gave 21 of 98 (21.4%) animals with the mutant phenotype (Fig. 5.1 M,N), as expected if *mow*^{w60} and *hi3357* fail to complement and thus disrupt the same gene. Dissections and flat-mounts of pharyngeal cartilages suggested that chondrogenic cells had reached their appropriate anatomical locations and condensed normally in *hi3357* mutants (Fig. 5.1 O,P). These results show that the *mow*^{w60} and *hi3357* mutations interrupt a gene essential for deposition of the Alcian blue-positive extracellular matrix of cartilage and normal bone morphogenesis in zebrafish.

mow^{w60} *and* hi3357 *disrupt* uxs1

To identify the molecular genetic defect responsible for the *mow*^{w60} phenotype, we generated an F2 mapping population. A genome scan using simple sequence repeats (SSRs) (Knapik et al., 1998; Shimoda et al., 1999) and bulk segregant analysis, showed that z3124, z9112 and z6663 on linkage group nine (LG9) were linked to the mutant locus. The genotyping of 480 individuals from the F2 mapping panel mapped *mow* about 20cM from z3124 (Fig. 5.2 A). A search of ZFIN (http://zfin.org) for mutants near *mow* identified the pro-viral insertion *hi3357,*

Figure 5.2. Mapping, cloning, and sequencing of *mow*. A. LG9 from the HS panel above (Woods et al., 2005) and the interval mapped on the *mow* mapping panel below. Distances are in centiMorgans (cM). B. Phylogenetic tree of proteins retrieved by BLASTP search of NCBI, aligned by Clustal-X, and analyzed by neighbor joining showed that the zebrafish protein disrupted by *mow*^{w60} falls in the Uxs1 clade, well-separated from the next most closely related protein, TGDS. Sequences and sequence identities listed in Methods. C. Nucleotide sequence comparison of *uxs1* from homozygous *mow*^{w60} mutants to wild-type siblings and reference sequence NM_173242 revealed a G-to-A replacement at nucleotide position 1283 (arrow). D. Amino acid sequence alignment of the portion of Uxs1 corresponding to the nucleotides shown in part C for zebrafish (WT), zebrafish (*mow*^{w60}) with the arginine (R) to histidine (H) substitution at position 233, human, chicken, frog *Xenopus tropicalis*, sea squirt *Ciona intestinalis* (translated from genomic sequence, CINT1.95:scaffold_505), fruit fly *Drosophila melanogaster* (CG7979- PA), fungus *Cryptococcus neoformans* (AAM22494) and mustard plant *Arabidopsis thaliana* (NP_190920). Unless otherwise noted, sequences were the same as those used in panel B. E. Structure of the zebrafish *uxs1* gene, showing the location of the *hi3357* viral insert in exon 1 (Amsterdam et al., 2004; Golling et al., 2002; Nissen et al., 2006) and the position of the *mow*^{w60} nucleotide substitution in exon 9.


A
LG9
hoxd4a
dlx2
Z15449
AI588511
AA658595
pour2
hha
Z4673
Z3124
AW342969
taco3
wnt10a
Z4168
fn1
0.0
3.0
5.7
10.9
15.9
25.7
28.1
30.5
33.0
35.9
38.7
41.8
44.6
uxs1
mow
z3124
0.7cM
B
0.1
1000 UXS1 human
987 Uxs1 mouse
970 Uxs1 chicken
1000 Uxs1 frog
Uxs1 pufferfish
955 Uxs1 zebrafish
523 Uxs1 fruit fly
Uxs1 beetle
1000 Uxs1 bee
765
1000
892
Uxs1p fungus
UXS6 mustard
Uxs1 rice
Uxs1 bacterium
1000
997 TGDS human
1000 Tgds mouse
1000 Tgds chicken
997 Tgds frog
1000 Tgds zebrafish
1000 MUM4 mustard
Tgds rice
549 Tgds bacterium
C
1260
1270
1280
1290
NM_173242 CGGGCCTGTTATGATGAAGGAAAGCGTGTTGCTG
wt sibling CGGGCCTGTTATGATGAAGGAAAGCGTGTTGCTG
mow CGGGCCTGTTATGATGAAGGAAAGCATGTTGCTG
D
233
zebrafish (wt) R A C Y D E G K R V A E
zebrafish (mow) R A C Y D E G K H V A E
human R A C Y D E G K R V A E
chicken R A C Y D E G K R V A E
frog R A C Y D E G K R V A E
sea squirt R A C Y D E G K R V A E
fruit fly R A C Y D E G K R V S E
fungus R A C Y D E G K R V A E
mustard plant R S C Y D E G K R T A E
E
hi3357
mow
exons: 0.1kb
introns: 10kb
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

the preliminary characterization of which showed decreased Alcian blue staining (Amsterdam et al., 2004; Golling et al., 2002; Nissen et al., 2006) (see our Fig. 5.1 I-L). To see if *mow* maps near *hi3357,* we used BLAST to search the zebrafish genome for sequence flanking *hi3357* and identified on Zv4_scaffold346 an SSR from intron-7 of *uxs1*, the predicted gene near the insertion. We genotyped our *mow* mapping panel using mapping primers flanking this SSR, and learned that this SSR mapped less than 0.5cM from *mow*, as expected if mow^{w60} and *hi3357* disrupt the same gene.

Sequences flanking the *hi3357* insert are similar to the human *UDP xylose synthase 1* (*UXS1*) gene (Amsterdam et al., 2004; Golling et al., 2002). Evidence that this sequence is the zebrafish ortholog of *UXS1* comes from phylogenetic analysis and conserved syntenies. A neighbor-joining phylogenetic tree of amino acid sequences from organisms as diverse as humans and prokaryotes (Fig. 5.2 B) shows that the protein disrupted by mow^{w60} and *hi3357* falls in the Uxs1 clade of the tree, well-separated from the next most similar clade, Tgds (TDP-glucose 4,6-dehydratase), which itself has a zebrafish ortholog (NP_956111). Analysis of conserved syntenies confirms orthologies: the sequence lying to the immediate right of *uxs1* on zebrafish LG9 is zgc:112443, whose reciprocal best BLAST hit in the human genome is *ECRG4*, which lies adjacent to *UXS1* on human chromosome 2. The proximity of these neighbors (*UXS1* and *ECRG4*) has been conserved for 450 million years since the divergence of human and zebrafish lineages (Hedges, 2002).

To identify the induced molecular change in mow^{w60}, we sequenced overlapping fragments amplified across *uxs1* from mutant cDNA template and found eight nucleotide differences from the reference sequence NM_173242: four differences are present in ESTs of wild-type zebrafish and thus represent naturally occurring polymorphisms; one causes a synonymous change; two

are in the 5' untranslated region; and one changes a guanine (G) to adenosine (A) at nucleotide position 1283 in the highly conserved exon 9 (Fig. 5.2 C,E). The G1283A change replaces arginine 233 with histidine (R233H) in a ten amino acid portion of the Uxs1 protein that is nearly invariant among all sequenced animals and fungi (Fig. 5.2 D). In summary, our genetic mapping data and gene sequencing data support the inference from the complementation study (Fig. 5.1 M,N) that *mow*^{w60} and *hi3357* disrupt the same gene, *uxs1*.

Uxs1 is one of the most strongly conserved non-mitochondrial proteins

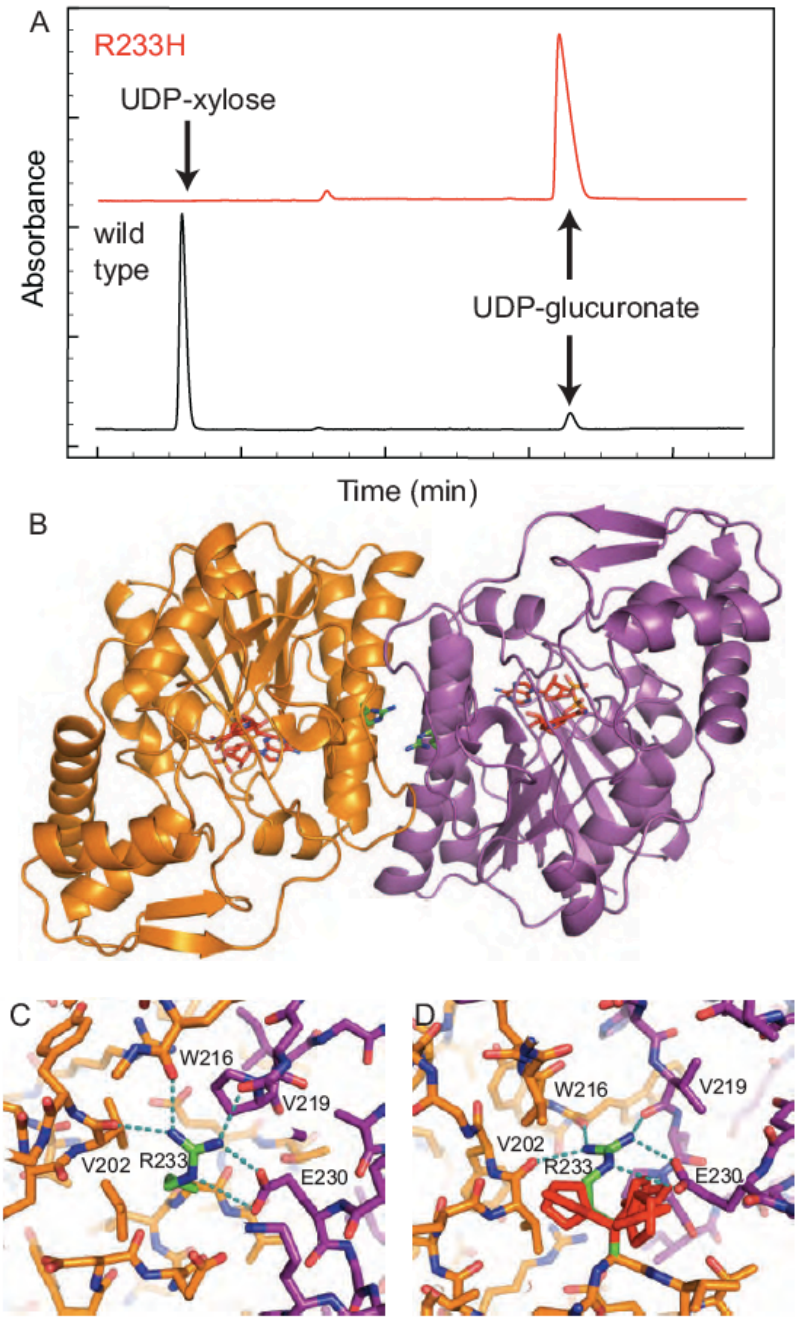
Sequence alignments showed remarkably strong conservation of Uxs1 sequence across all three domains of life. Compared to the Uxs1 of the prokaryote bacterium *Rhodospirillum rubrum*, eukaryotes have 55% to 63% identity (vertebrates, such as human, mouse, chicken, frog, and fish, had 57% amino acid identity; insects, such as fruit fly and beetle, 59%, a fungus (*Cryptococcus neoformans*) 55%; and plants, such as rice and mustard, 62-63%). An archaebacterium, for example *Methanocaldococcus vulcanius*, has 51%, while other bacteria have about 50% or more identity (*Rhizobium leguminosarum*, 54%, *Escherichia coli, 68%, Thermotoga neapolitana*, 49%). To see how Uxs1 ranks on a list of conserved proteins, we compared the human proteome to the proteome of the bacterium *R. rubrum* by BLAST analysis, ranked the results by E-value, and examined the literature for sub-cellular localization of each hit, ruling out mitochondrial proteins encoded by nuclear or mitochondrial genomes. Results showed that Uxs1 (E-value = 2.10E-100) ranked fifth on the list after GPI, IDH1, ADH5, and GMDS, which had E-value scores of 2.80E-179 to 2.20E-117 and % identities of 61% to 66%. We conclude that the amino acid sequence of Uxs1 has been one of the most strongly conserved proteins for over 3 billion years of biological evolution.

Molecular function and structure of mutant Uxs1

Because it disrupts sequences directly upstream of translation initiation, *uxs1*^{hi3357} is likely a null allele (Fig. 5.2 E). Given the phenotypic similarity of *mow*^{w60} to *uxs1*^{hi3357}, we wanted to determine whether *mow*^{w60} is also a null allele and to try to understand why Uxs1 is so strongly constrained for life on Earth. Zebrafish Uxs1 shares 86% sequence identity to the human enzyme, with most sequence diversity residing in the amino terminal transmembrane domain. The soluble domains of Uxs1 that include arginine 233 share 96% sequence identity between zebrafish and human. To determine whether the arginine-to-histidine replacement observed in *mow*^{w60} changes a site essential for enzyme activity, we replaced the equivalent arginine with a histidine in the human enzyme (R236H; hereafter all residues are referred to with zebrafish numbering, i.e. R233H). We then assayed the wild-type and mutant enzymes for activity. Results showed that the mutant enzyme did not convert the substrate UDP-glucuronic acid to the product, UDP-xylose, but the wild-type enzyme made this conversion readily (Fig. 5.3 A). This result shows that the substitution of arginine by histidine at this position is sufficient to eliminate the enzymatic function of Uxs1 protein.

Using the unpublished, publicly available crystal structure of human UXS1 (The Structural Genomics Consortium, PDB entry 2B69), we modeled the zebrafish R233H mutation. The crystal structure of UXS1 revealed a homodimer with each subunit containing a molecule of nicotinamide adenine dinucleotide (NAD) bound in the active site (Fig. 5.3 B). Arg233 is buried in the dimer interface about 16 Å distant from NAD, and it appears to stabilize the dimer interface. The guanidinium of Arg233 makes hydrogen bonds with the main-chain carbonyl oxygen atoms of Trp216 and Val202 (Fig. 5.3 C). Arg233 also interacts with residues from the other subunit by making a salt-bridge with Glu230 and a hydrogen bond to the main-chain

Figure 5.3. Effects of the *mow*^{w60} mutation on Uxs1 function and structure. A. Capillary zone electrophoresis chromatogram of enzyme reactions catalyzed by wild-type human UXS1 (lower black trace) or mutated UXS1 corresponding to the zebrafish *mow*^{w60} (R233H, upper red trace). Peaks represent absorbance at 260nm and are on the same scale, but offset vertically for ease of comparison using Plot (http://plot.micw.eu). UXS1 with the *mow*^{w60} mutation produced no detectable enzyme product, UDP-xylose, after an 18 hour incubation with substrate UDP-glucuronate at 37°C but the wild-type enzyme converted nearly all of the substrate into product. B. Ribbon drawing of the crystal structure of dimerized human UXS1 (PDB entry 2B69), with different monomers colored orange and purple. NAD (red and blue) and the side chains corresponding to zebrafish R233 (green and blue) are depicted as sticks. C, D. Conserved hydrogen bonding and salt bridge interactions (dashed pale blue lines) revealed the structural consequences of the zebrafish R233H substitution (D) at the dimer interface, compared to wild-type (C). The histidine (red sticks in D) was modeled in several of its common rotomeric states to illustrate the unfavorable contacts it introduces and its inability to satisfy the electrostatic interactions of the R233 guanidinium. Depiction of R233 (green sticks) provided as a frame of reference. Numbering corresponds to zebrafish residues. Figure generated using Pymol (DeLano, 2002).


A
R233H
UDP-xylose
Absorbance
wild
type
UDP-glucuronate
Time (min)
B
C
W216
V219
V202
R233
E230
D
W216
V219
V202
R233
E230

carbonyl oxygen of Val219.

Our modeling predicts that mutating Arg233 to histidine would disrupt the dimer interface because a histidine cannot satisfy the salt-bridge with Glu230 without introducing unfavorable contacts throughout the interface. Dimerization of the R233H mutant would bury the two acidic residues (one from each subunit) without a basic residue to neutralize the charge. Burying an uncompensated acid is energetically costly and will destabilize protein structure (Dao-pin et al., 1991). In addition, histidine cannot satisfy the extensive hydrogen bonding network supported by the guanidinium of Arg233 (Fig. 5.3 D). The carbonyl oxygens of Trp216 and Val219 of the other subunit are 3.3 Å apart, which would result in unfavorable electrostatic interactions in the absence of Arg233. Finally, histidine is not isosteric to the well-packed Arg233, and its bulky imidazole ring in place of the alkyl portion of the arginine side chain would introduce unfavorable contacts in the interface. Crystallographic analysis of an arginine-to-alanine substitution at this position in human UXS1 indicates local secondary structure disordering at the dimer interface (SJP and ZAW, *manuscript in preparation*). This evidence shows that the zebrafish R233H mutation should disrupt the Uxs1 dimer, thereby abolishing enzyme activity.

Expression of uxs1 *in zebrafish development*

If the phenotype of *mow*^{w60} is caused by disruption of *uxs1*, then *uxs1* should be expressed in craniofacial regions at or before the developmental defect becomes apparent. RT-PCR and whole-mount *in situ* hybridization experiments showed that one-cell embryos have maternal *uxs1* transcript (Fig. 5.4 A,B). Under the RT-PCR conditions used, *uxs1* signal decreased during epiboly and increased during segmentation stages (Fig. 5.4 A), as if maternal mRNA gradually disappeared and was replaced by zygotic transcript. At 24 hpf, *in situ* hybridization on sectioned

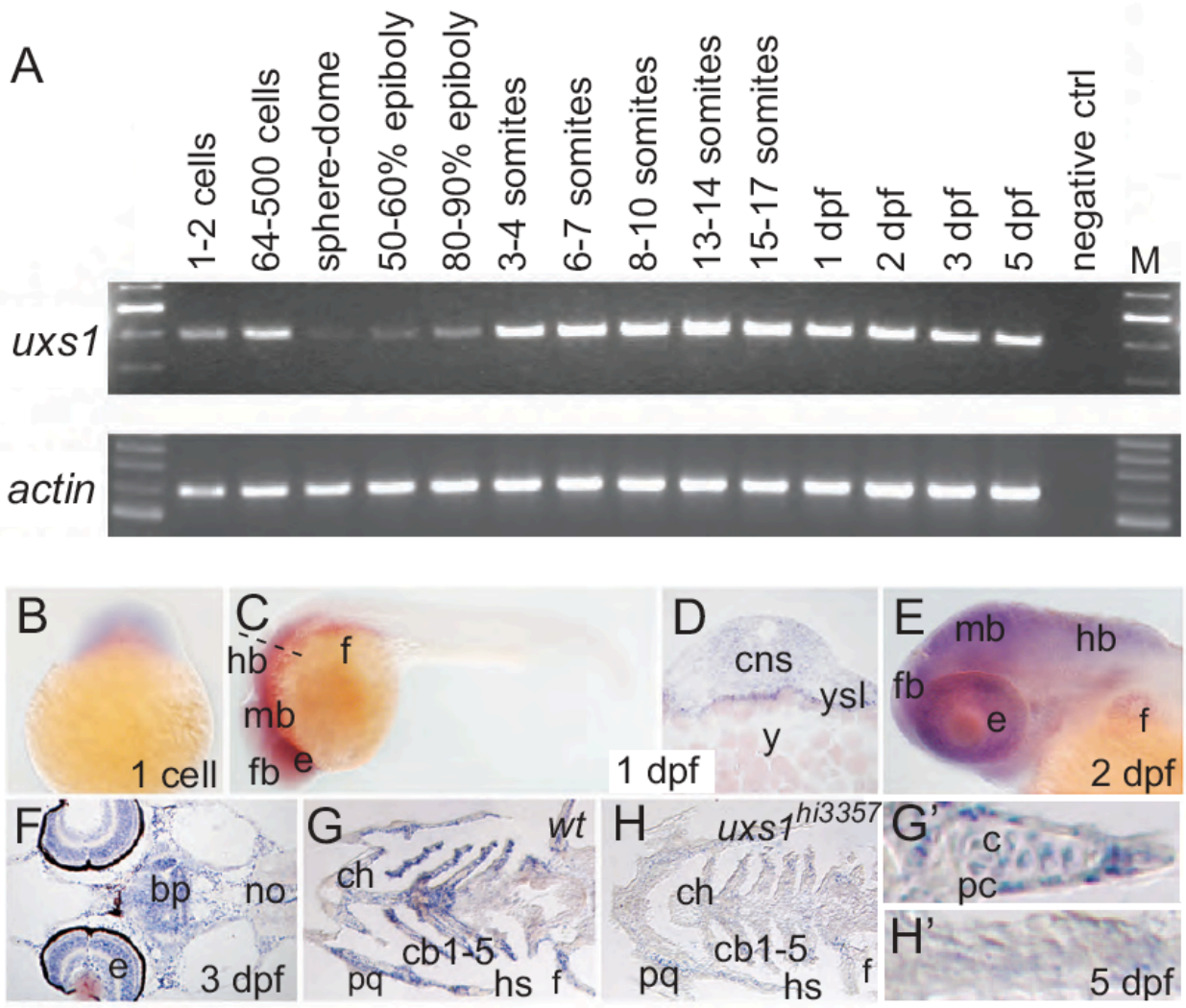


Figure 5.4. *uxs1* expression during zebrafish embryogenesis. A. RT-PCR for *uxs1* transcript in animals of indicated ages, along with β-*actin* positive controls. Maternal *uxs1* mRNA was detected at the 1-2 cell stage and detection decreased at sphere-dome stage. Zygotic *uxs1* expression appeared to increase gradually and was maintained at least through 5 dpf. B. Whole-mount *in situ* hybridization of a one-cell embryo revealed transcript in the fertilized egg. C,D. Whole-mount (C) and section (D) of 24 hpf embryos illustrated general expression of *uxs1* in brain and craniofacial mesenchyme, as well as in the yolk syncytial layer. The dashed line in C indicates the plane of section in D. E. Lateral view of whole-mount 2 dpf embryo showed widespread *uxs1* expression in the craniofacial region. F-H, horizontal sections of 3 dpf (F) and 5 dpf (G,H) animals. Expression of *uxs1* became localized to layers of the retina, brain, and cartilages of the pharyngeal arches. Levels of *uxs1* transcript were severely reduced or absent in pharyngeal regions of *uxs1*^{hi3357} embryos. High magnification of 5 dpf ceratohyals shows *uxs1* expression in both chondrocytes (c) and perichondral cells (pc) of wild types (G'), but low transcript levels in *uxs1*^{hi3357} embryos (H'). Abbreviations: bp, basal plate; c, chondrocyte; cb1-5, ceratobranchials 1-5; ch, ceratohyal; e, eye; f, fin bud; fb, forebrain; hb, hindbrain; hs, hyosymplectic; mb, midbrain; no, notochord; pc, perichondrium; pq, palatoquadrate; y, yolk; ysl, yolk syncytial layer.

and whole-mounted embryos revealed *uxs1* transcript in ventral craniofacial domains and along the yolk-endoderm boundary (Fig. 5.4 C,D). At 2 and 3 dpf, *uxs1* transcripts were broadly distributed throughout the head, appearing in the pharyngeal arches, eye, fin bud, neurocranium, notochord, and brain (Fig. 5.4 E,F). At 5 dpf, *uxs1* expression localized to the developing pharyngeal arch cartilages in both chondrocytes and perichondrium (Fig. 5.4 G,G'). We also observed that *uxs1* transcript is diminished in *uxs1*^{hi3357} larvae at 5dpf (Fig. 5.4 H,H'), suggesting that transcript containing the viral insert is unstable. These results show that *uxs1* transcript is localized in a pattern consistent with the mutant craniofacial phenotype, and that transcript instability may contribute to the loss of *uxs1* function in *uxs1*^{hi3357} larvae.

uxs1 *mutants show altered morphogenesis of cartilage, perichondrium, and dermal bone*

To identify cellular mechanisms responsible for the disruption of skeletal development in *uxs1* mutants, we crossed *uxs1* mutant alleles into the transgenic line *Tg(fli1:EGFP)y1*, which expresses green fluorescent protein (GFP) in developing skeletogenic neural crest cells (Lawson and Weinstein, 2002). Dual-channel confocal microscopy of vital Alizarin red-stained wild-type and mutant larvae of the *Tg(fli1:EGFP)y1*;*uxs1*^{hi3357} and *Tg(fli1:EGFP)y1;mow*^{w60} strains highlighted both crest-derived chondrocytes and bone calcification centers in living animals. In 4 dpf wild types, cells had already begun to stack in craniofacial cartilages (Fig. 5.5 A,E). In 4 dpf *uxs1* mutants, however, chondrocytes did not intercalate with their neighbors and stack normally (Fig. 5.5 C,G). In 7 dpf wild-type larvae, chondrocytes maintained their intercalated and stacked organization, and the bone collar had started to form around the central portion of the ceratohyal (Fig. 5.5 B,F). Additionally, Alizarin red staining revealed the growth of dermal bones, such as the dentary, which by 7 dpf had proceeded laterally from its location in 4 dpf larvae. In 7 dpf

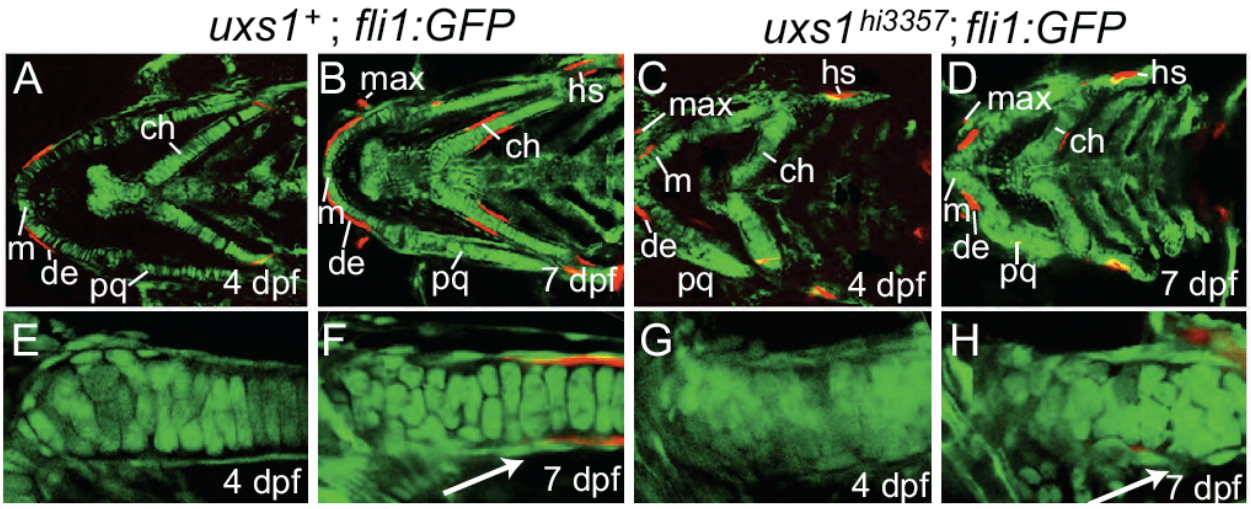


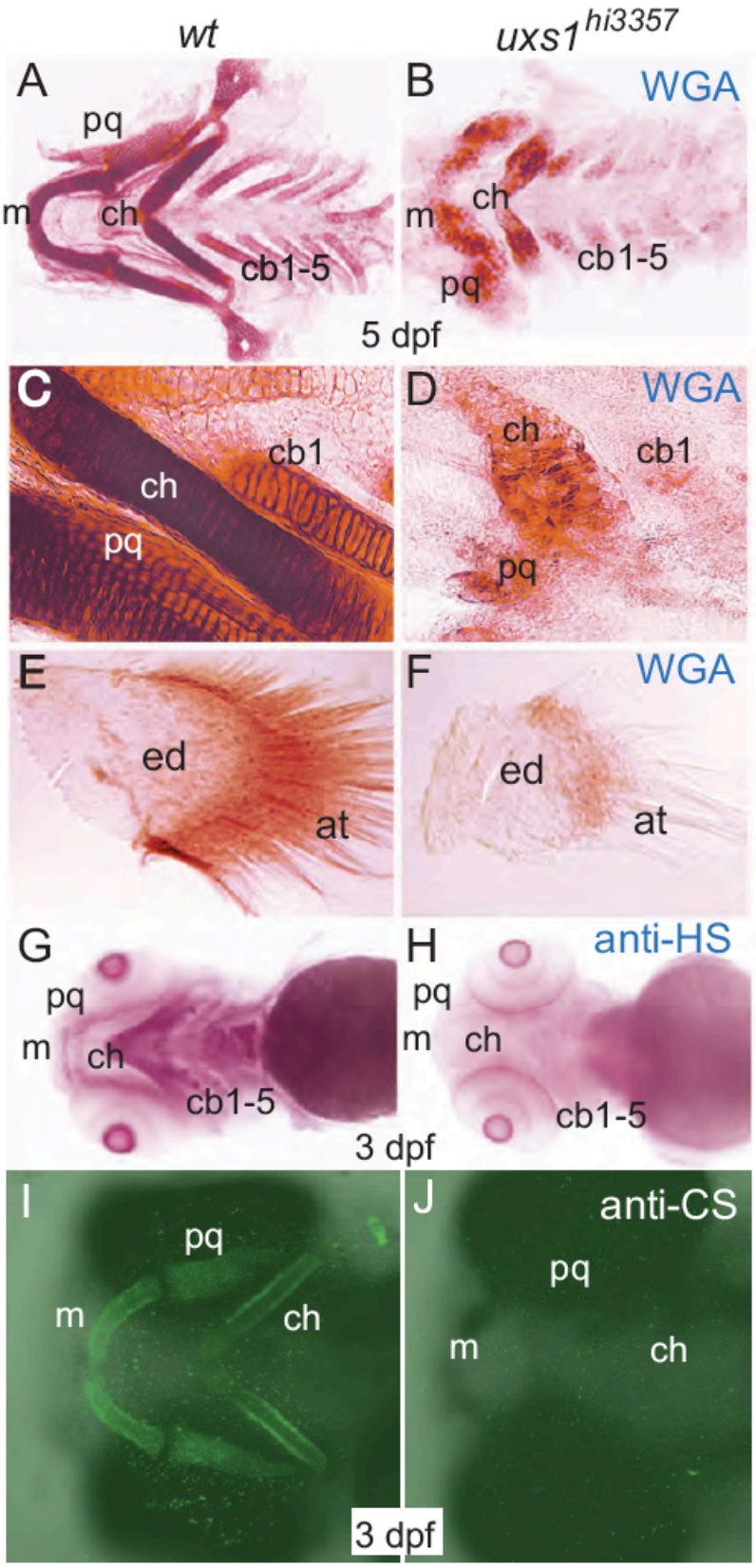
Figure 5.5. Cellular visualization of cartilage and bone morphologies in wild-type and *uxs1* mutant larvae. A-H, Optical sections of live Alizarin red-stained *Tg(fli1:EGFP)y1* larvae, ventral views, at 4 dpf and the same individuals at 7 dpf., E-H focus on the ceratohyal. In wild types (A,B,E,F), chondrocytes stacked and were lined with a flattened layer of perichondral cells (white arrow in F). Ossification centers stained with Alizarin red, reflecting perichondral bone formation in the ceratohyal and hyosymplectic and intramembranous ossification in the dentary and maxilla. In homozygous *uxs1*^{hi3357} animals (C,D,G,H), however, chondrocytes were disorganized, the perichondral sheath did not align properly (white arrow in H), and Alizarin red-positive ossification centers (dentary, maxilla, and ceratohyal) were severely reduced in perichondral and intramembranous sites. Abbreviations: ch, ceratohyal; de, dentary; hs, hyosymplectic; m, Meckel's cartilage; max, maxilla; pq, palatoquadrate.

uxs1 mutants, however, chondrocytes remained rounded, with little intercalation, and were spatially disorganized (Fig. 5.5 D,H). Likewise, the mutant dentary showed defective morphogenesis, failing to extend as far laterally as in wild-type animals (compare Fig. 5.5 B to 5.5 D). The perichondrium, which depends on signaling from chondrocytes, was not well organized in mutants, lacking the continuous flattened layer of cells seen in wild types (compare Fig. 5.5 F to 5.5 H, arrows). At both 4 and 7 dpf, the ceratohyal showed no clear difference in cell number between mutants and wild types, although mutant cartilage elements appeared shorter, suggesting a failure in elongation. These data demonstrate that *uxs1* functions in many aspects of skeletal morphogenesis, helping chondrocytes intercalate and stack, driving cartilage elongation, and directing perichondral and dermal bone morphology.

Proteoglycan levels are reduced in uxs1 *mutants*

We next sought to identify molecular alterations in the mutant ECM that might provide a mechanism for observed defects in skeletal histogenesis and morphogenesis. To investigate whether *uxs1* is required for normal proteoglycan biosynthesis, we used biotinylated-succinylated wheat germ agglutinin (WGA) to stain N-acetylglucosamine, a sugar found in the GAG chains of cartilage proteoglycans (Lang et al., 2006). In wild types at 5 dpf, WGA localized to cartilages of the pharyngeal arches (Fig. 5.6 A). At higher magnification, WGA staining appeared in organized layers between cells in the palatoquadrate and in ceratobranchial-1 (Fig. 5.6 C). WGA-positive material also appeared in intracellular foci (data not shown), possibly due to N-acetylglucosamine within the Golgi complex of chondrocytes. In homozygous *uxs1*^{hi3357} and *mow*^{w60} mutants, however, WGA was reduced or absent in cartilages of the pharyngeal arches, including the ceratohyal and Meckel's cartilage (Fig. 5.6 B, and data not

Figure 5.6. Proteoglycan detection in wild-type and *uxs1* mutant skeletons. A-F, whole-mount wheat germ agglutinin (WGA) staining to visualize N-acetylglucosamine, ventral views. G-J, whole-mount immunostaining against heparan sulfate (G,H) and chondroitin sulfate (I,J) proteoglycans, ventral views. Dissected pharyngeal cartilages revealed reduced WGA staining in *uxs1*^{hi3357} mutants (B,D), compared to wild-type siblings (A,C) at 5 dpf. Higher magnification of ceratohyal regions also showed that WGA-positive material was not deposited normally in mutants (D), compared to organized deposition in wild types (C). Dissected pectoral fins showed that both endoskeletal disc and actinotrichia had less WGA staining and fewer actinotrichia in *uxs1*^{hi3357} mutants (F), compared to wild-type siblings (E) at 5 dpf. Immunodetection of heparan sulfate demonstrated that HSPGs were localized to pharyngeal domains in wild type (G), but HSPGs were not detectable in homozygous *uxs1*^{hi3357} animals (H). Similarly, immunodetection of chondroitin sulfate was abundant in wild-type cartilages (I), but was absent in *uxs1* mutants (J). Abbreviations: at, actinotrichia; cb1-5, ceratobranchials 1-5; ch, ceratohyal; ed, endoskeletal disc; m, Meckel's cartilage; pq, palatoquadrate.


wt
uxs1hi3357
A
pq
m
ch
cb1-5
B
WGA
ch
m
cb1-5
pq
5 dpf
C
cb1
ch
pq
D
WGA
ch
cb1
pq
E
ed
at
F
WGA
ed
at
G
pq
m
ch
cb1-5
H
anti-HS
pq
m
ch
cb1-5
3 dpf
I
pq
m
ch
J
anti-CS
pq
m
ch
3 dpf

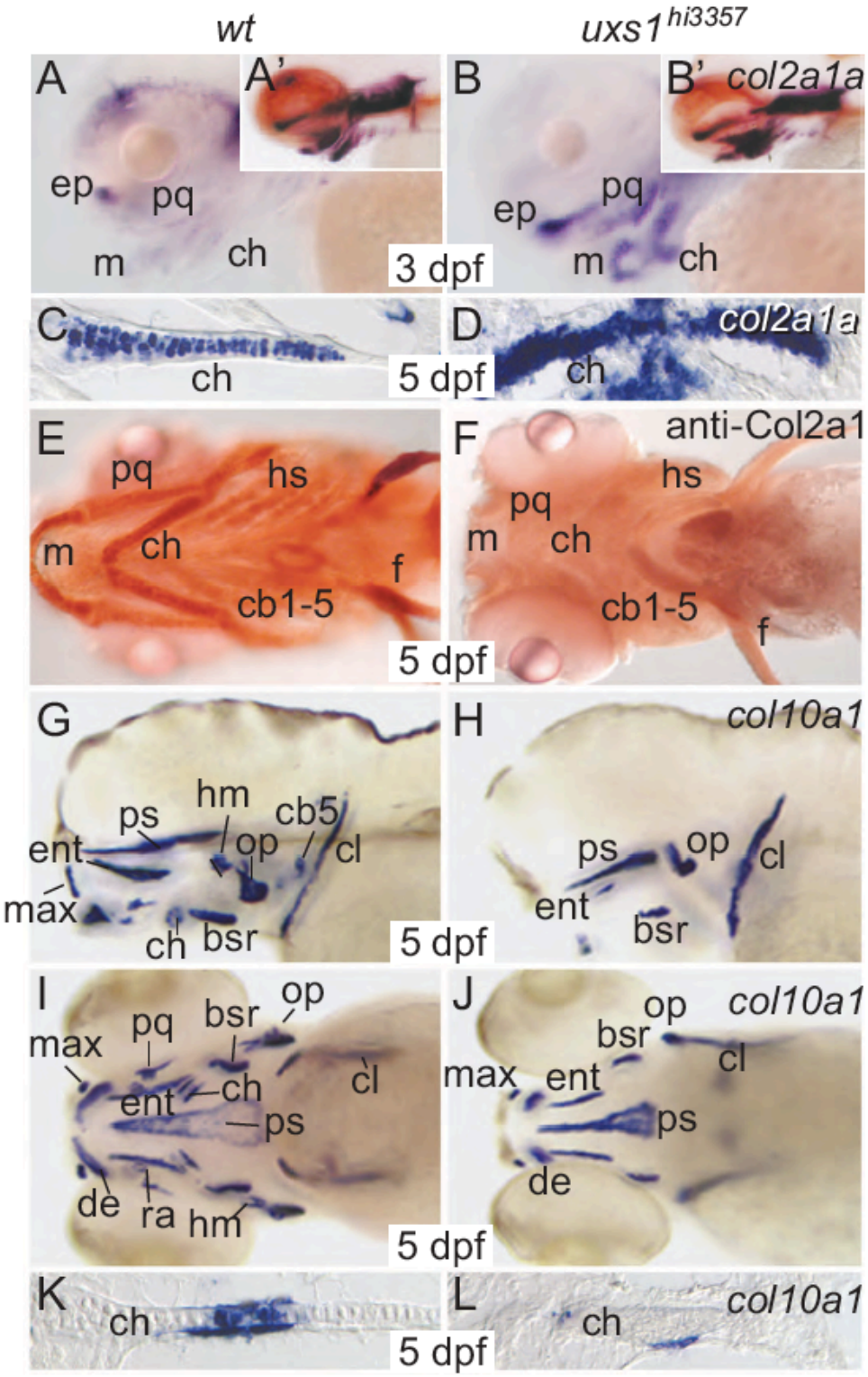
shown). Moreover, cartilage morphologies were disorganized, and WGA staining was irregularly dispersed between limited numbers of cells in mutants (Fig. 5.6 D). In the appendicular skeleton of wild types, the endoskeletal disc stained with WGA and abundant actinotrichia formed a filamentous fin tip (Fig. 5.6 E). Homozygous *uxs1* mutants, however, showed little WGA staining in their diminutive endoskeletal discs and their actinotrichia were reduced in number and stained poorly (Fig. 5.6 F).

If *uxs1* mutants fail to make UDP-xylose for the tetrasaccharide linker, then they should not produce heparan sulfate or chondroitin sulfate proteoglycans (HSPGs, CSPGs). In wild types at 3 dpf, heparan sulfate antibody (α-HS) identified HSPGs in pharyngeal arches and pectoral fin (Fig. 5.6 G). In homozygous *uxs1* mutant heads, however, staining disappeared (Fig. 5.6 H), showing that *uxs1* activity is essential for the formation of HSPGs. Similarly, immunodetection of CSPGs was abundant in wild-type cartilage at 3 dpf, but *uxs1* mutants failed to show any evidence of CSPGs (Fig 5.6 I,J). Together, these results show that *uxs1* function is required for proteoglycan expression and localization in pharyngeal cartilages, and is necessary for normal development of pectoral appendages.

Uxs1 activity regulates expression of skeletal collagens

Surprisingly, we discovered that proper expression of collagenous ECM components relied on Uxs1-dependent proteoglycan production. First, we investigated the expression of *col2a1a*, which encodes a major collagen of cartilage (Hamerman, 1989; Yan et al., 1995). Compared to wild-type siblings, homozygous *uxs1* mutant embryos at 3 dpf had increased levels of *col2a1a* transcript in craniofacial cartilage elements (Fig. 5.7 A,B). *In situ* hybridization of sections of 5 dpf larvae confirmed increased *col2a1a* transcript levels in developing chondrocytes of *uxs1*

Figure 5.7. Collagen detection in wild-type and *uxs1* mutant skeletons. A-D whole-mount (A,B) and horizontal section (C,D) *in situ* hybridization for *col2a1a* gene expression; E,F, whole-mount immunostaining for Col2 protein; G-L, whole-mount (G-J) and horizontal section (K,L) *in situ* hybridization for *col10a1* gene expression. Expression of *col2a1a* increased in developing cartilage of *uxs1*^{hi3357} mutants in lateral views of the head at 3 dpf (B) and ceratohyal sections at 5 dpf (D), compared to wild-type siblings (A,C). Longer substrate developing times demonstrated that *col2a1a* levels are high in cartilage of both wild type and mutant heads at 3 dpf (A',B') In contrast, although Col2a1 protein was easily detected in wild type cartilages (E) in ventral view at 5 dpf, it was not detected in mutant cartilages (F). Lateral (G,H) and ventral (I,J) whole-mount views showed that domains of *col10a1* gene expression were greatly reduced in regions of endochondral and intramembranous skeletal elements in 5 dpf *uxs1* mutants (H,J), compared to wild types (G,I). *In situ* hybridization on histological sections of the ceratohyal at 5 dpf illustrated reduced perichondral staining of *col10a1* in mutants (L), compared with wild types (K). Also, chondrocyte expression of *col10a1* was absent in mutants, although wild-type ceratohyal chondrocytes strongly expressed *col10a1*. Abbreviations: bsr, branchiostegal ray; cb1-5, ceratobranchials 1-5; ch, ceratohyal; cl, cleithrum; de, dentary; ent, entopterygoid; ep, ethmoid plate; f, fin; hm, hyomandibular; hs, hyosymplectic; m, Meckel's cartilage; max, maxilla; op, opercle; pq, palatoquadrate; ps, parasphenoid.


wt
uxs1 hi3357
A
A'
B
B' col2a1a
ep
pq
m
ch
3 dpf
C
ch
5 dpf
D
col2a1a
ch
E
pq
hs
m
ch
f
cb1-5
5 dpf
F
anti-Col2a1
hs
pq
m
ch
cb1-5
f
G
hm
ps
cb5
ent
op
cl
max
ch
bsr
5 dpf
H
col10a1
ps
op
cl
ent
bsr
I
op
bsr
pq
max
cl
ch
ent
ps
de
ra
hm
5 dpf
J
col10a1
op
bsr
cl
max
ent
ps
de
K
ch
5 dpf
L
col10a1
ch

mutant ceratohyals (Fig. 5.7 C,D). Differences between wild types and *uxs1* mutants were also apparent upon analysis of Col2a1 protein. Immunohistochemistry revealed that Col2a1 was present in all craniofacial cartilages and in the fin in wild types at 5 dpf (Fig. 5.7 E). In contrast, Col2a1 protein was almost completely absent from craniofacial cartilages of homozygous *mow*^{w60} and *uxs1*^{hi3357} mutants and was greatly reduced in the fin (Fig. 5.7 F, and data not shown). These results show that *uxs1* normally inhibits the accumulation of *col2a1a* transcripts, but, somewhat paradoxically, promotes the deposition and/or maintenance of stainable Col2a1 protein in the extracellular matrix.

Type X collagen (Col10a1) is a marker of both osteoblasts and maturing chondrocytes in teleosts (Avaron et al., 2006; Clement et al., 2008; Simoes et al., 2006). In wild types, whole-mount *in situ* hybridization at 5 dpf showed that *col10a1* is expressed in osteoblasts of both the endochondral (e.g., ceratohyal, hyomandibular) and intramembranous (e.g., opercle, dentary) skeletons (Fig. 5.7 G,I). In homozygous *uxs1* larvae at 5 dpf, however, *col10a1* expression occurred only in dermal bones, and these expression domains were generally smaller than normal (Fig. 5.7 H,J). Endochondral ossification appeared delayed or absent in *uxs1* mutants at 5 dpf, since perichondral *col10a1* expression is missing from the mutant ceratohyal, but is obvious in wild types at this stage. Because chondrocyte maturation also reflects the degree to which endochondral ossification has progressed, we performed *in situ* hybridization on sections to analyze chondrocyte *col10a1* expression. At 5 dpf, maturing chondrocytes in the mid-diaphyseal region of wild-type ceratohyals expressed *col10a1*, whereas *uxs1* mutants failed to express this marker of chondrocyte maturation (Fig. 5.7 K,L). Collectively, these results demonstrate that *uxs1* function is essential for the induction and/or maintenance of skeletal collagens, such as

Col2a1 and *col10a1*, and suggest that the maturation of ceratohyal chondrocytes to the *col10a1*-expressing stage relies upon Uxs1-dependent proteoglycans.

Uxs1 is upstream of the 'master regulators' of cartilage and bone

The transcription factors Sox9 and Runx2 have been thought of as master regulators of cartilage and bone, respectively (Akiyama, 2008; Eames et al., 2003; Eames and Helms, 2004; Eames et al., 2004; Otto et al., 1997; Yoshida and Komori, 2005). To determine whether the molecular defects in skeletal histogenesis that we discovered in *uxs1* mutants result from Uxs1-dependent expression of these important regulators, we compared the expression patterns of *sox9a, sox9b, runx2a,* and *runx2b* in wild-type and *uxs1* mutant larvae in histological sections.

Ceratohyal cartilages in wild-type zebrafish express *sox9a* in chondrocytes at 3 dpf, but *sox9a* becomes down-regulated at 5 dpf in the mid-diaphyseal region as chondrocytes mature (Fig. 5.8 A,C). In contrast, the ceratohyal of *uxs1* mutants over-expresses *sox9a* and fails to down- regulate *sox9a* on the normal schedule (Fig. 5.8 B,D). We did not detect any differences in *sox9a* expression between wild types and *uxs1* mutants at 48 hpf (data not shown). Therefore, up- regulation of *sox9* in mutants only occurred after chondrogenic cells in wild types began to secrete abundant proteoglycans at 3 dpf, as detected by antibodies to chondroitin sulfate (data not shown). Chondrocytes of the wild-type ceratohyal do not express *sox9b* at 72 hpf, although *sox9b* expression is high in developing chondrocytes at 48 hpf (Fig. 5.8 E, data not shown; (Yan et al., 2005). In *uxs1* mutants, however, weak *sox9b* expression persists inappropriately in ceratohyal chondrocytes at 3 dpf (Fig. 5.8 F). In wild types, *runx2a* and *runx2b* are expressed in the perichondrium of the ceratohyal in 3 dpf larvae (Fig. 5.8 G,I). As would be expected from the previous defects in perichondral bone of *uxs1* mutants, *runx2a* and *runx2b* expression in cells

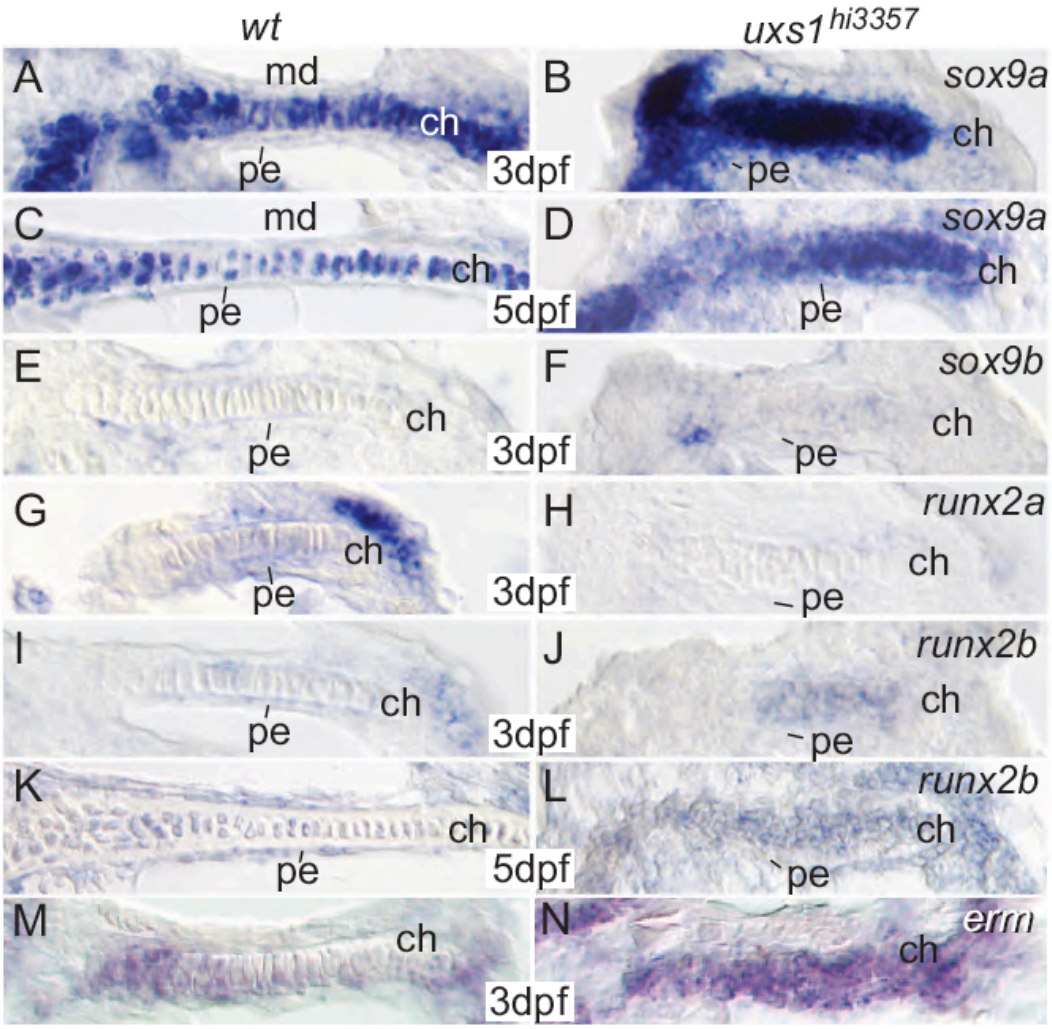


Figure 5.8. Detection of molecular regulators of skeletogenesis in *uxs1* mutant cartilage. A-N, *in situ* hybridization on horizontal sections through the ceratohyal for *sox9a* (A-D), *sox9b* (E,F), *runx2a* (G,H), *runx2b* (I-L), and *erm* (M,N). Wild-type chondrocytes in the mid-diaphyseal region showed decreased *sox9a* expression from 3 dpf (A) to 5 dpf (C) as they matured. Not only did chondrocytes of *uxs1*^{hi3357} mutants fail to show this down-regulation over time (B,D), but in addition, *sox9a* expression overall was much higher in mutants compared to wild types. Expression of *sox9b* was absent in wild-type chondrocytes at 3 dpf (E), but transcripts were detected in *uxs1* mutants (F). *runx2a* expression was obvious in perichondrium of wild types at 3 dpf (G), but was absent in *uxs1* mutants (H). *runx2b* expression was found in perichondrium of wild types at 3 dpf (I) and 5 dpf (K), but was not easily detected in perichondrium of mutants at these timepoints (J,L). In addition, chondrocyte expression of *runx2b* was much higher in *uxs1* mutants compared to wild types at 3 and 5 dpf. Expression of the FGF-responsive gene *erm* was found in just a few wild-type chondrocytes at 3 dpf (M), whereas *erm* transcripts were at high levels in all *uxs1* mutant chondrocytes (N). Abbreviations: ch, ceratohyal; md, mid- diaphyseal region; pe, perichondrium.

surrounding condensed ceratohyal chondrocytes was decreased (Fig. 5.8 H,J). In addition, *runx2b* expression in chondrocytes of the ceratohyal was up-regulated in *uxs1* mutants compared to wild types at 3 and 5 dpf (Fig. 5.8 I-L). Normal expression of *runx2a* does not occur at high levels in chondrocytes, and we saw no ectopic *runx2a* expression in mutant chondrocytes.

Because FGF signaling drives *sox9* expression in chondrocytes (Coumoul and Deng, 2003; de Crombrugghe et al., 2000; Eames and Schneider, 2008; Itoh and Ornitz, 2004), and because proteoglycans are known to affect FGF signaling (Ornitz, 2000; Pellegrini, 2001), we sought to determine whether the *uxs1* mutation increased *sox9* expression through increased FGF signaling. While transcripts of the FGF-responsive genes *sprouty4* and *pea3* were not detected in chondrocytes of the ceratohyal at 3 dpf (data not shown), *erm* expression was found in few chondrocytes in wild types (Fig. 5.8 M). Supporting the notion that FGF signaling is up-regulated in early cartilages of *uxs1* mutants, *erm* expression was detected at high levels in all mutant chondrocytes (Fig. 5.8 N). From these data, we conclude that *uxs1* normally acts as a negative regulator of *sox9a*, *sox9b*, and *runx2b* in chondrocytes, perhaps through modulation of FGF signaling, and that *uxs1* function is required to turn on *runx2a* and *runx2b* in the perichondrium.

uxs1 *mutants and Hedgehog signaling*

Aberrant Hedgehog (Hh) signaling could explain many aspects of the described skeletal phenotype in *uxs1* mutants because Hh signaling drives proper histogenesis and morphogenesis of cartilage and perichondrium during endochondral ossification (Colnot et al., 2005; Cortes et al., 2009; Kronenberg, 2003; St-Jacques et al., 1999). In addition, proteoglycans are important for growth factor signaling, including that of the Hh family (Koziel et al., 2004; Lin, 2004). In

normal development, *Indian hedgehog* (*Ihh*) is expressed in maturing chondrocytes, while expression of Patched (Ptc), a Hh receptor and target of Hh signaling, is restricted to proliferating chondrocytes and perichondrium (Avaron et al., 2006; Iwasaki et al., 1997; Nakase et al., 2001; Nakashima and de Crombrugghe, 2003; St-Jacques et al., 1999; Vortkamp et al., 1996). In wild-type ceratohyals at 5 dpf, *ihha* and *ihhb* were expressed in maturing chondrocytes located at the mid-diaphyseal region (Fig. 5.9 A,C). In homozygous *uxs1* mutants, expression of *ihha* and *ihhb* was detected in chondrocytes of the ceratohyal, although it was unclear whether the overall expression pattern within the skeletal element was normal, due to the disrupted organization of the mutant cartilage (Fig. 5.9 B,D). Similar patterns were observed at 4 and 6 dpf (data not shown). Given *ihh* expression in mutant cartilages, we next assayed for evidence of Hh signaling, which should be apparent in the expression of *ptc1* and *ptc2* (Goodrich et al., 1996; Lewis et al., 1999). In wild-type 5 dpf ceratohyals, *ptc1* and *ptc2* expression were relatively high in perichondrium and low in chondrocytes (Fig. 5.9 E,G). Homozygous *uxs1* mutants showed low levels of chondrocyte staining with the *ptc1* and *ptc2* probes, but expression in the perichondrium and surrounding tissues appeared to be up- regulated compared to wild types (Fig. 5.9 F,H). Additional markers of Hh signaling, such as *gli1, gli2, gli3* and *pthrp1 and pthrp2* genes (Katoh and Katoh, 2009), were analyzed. Most of these genes showed weak expression In cartilage or perichondrium at this timepoint in both wild-type and mutant ceratohyals (data not shown), although *gli3* transcripts were at slightly higher levels in cells surrounding *uxs1* mutant cartilage than those observed in wild-type siblings (Fig. 5.9 I,J). From these data, we conclude that Hedgehog signaling appeared to be functional and elevated during endochondral ossification in *uxs1* mutants, suggesting that normal Uxs1 activity inhibits Hh signaling.

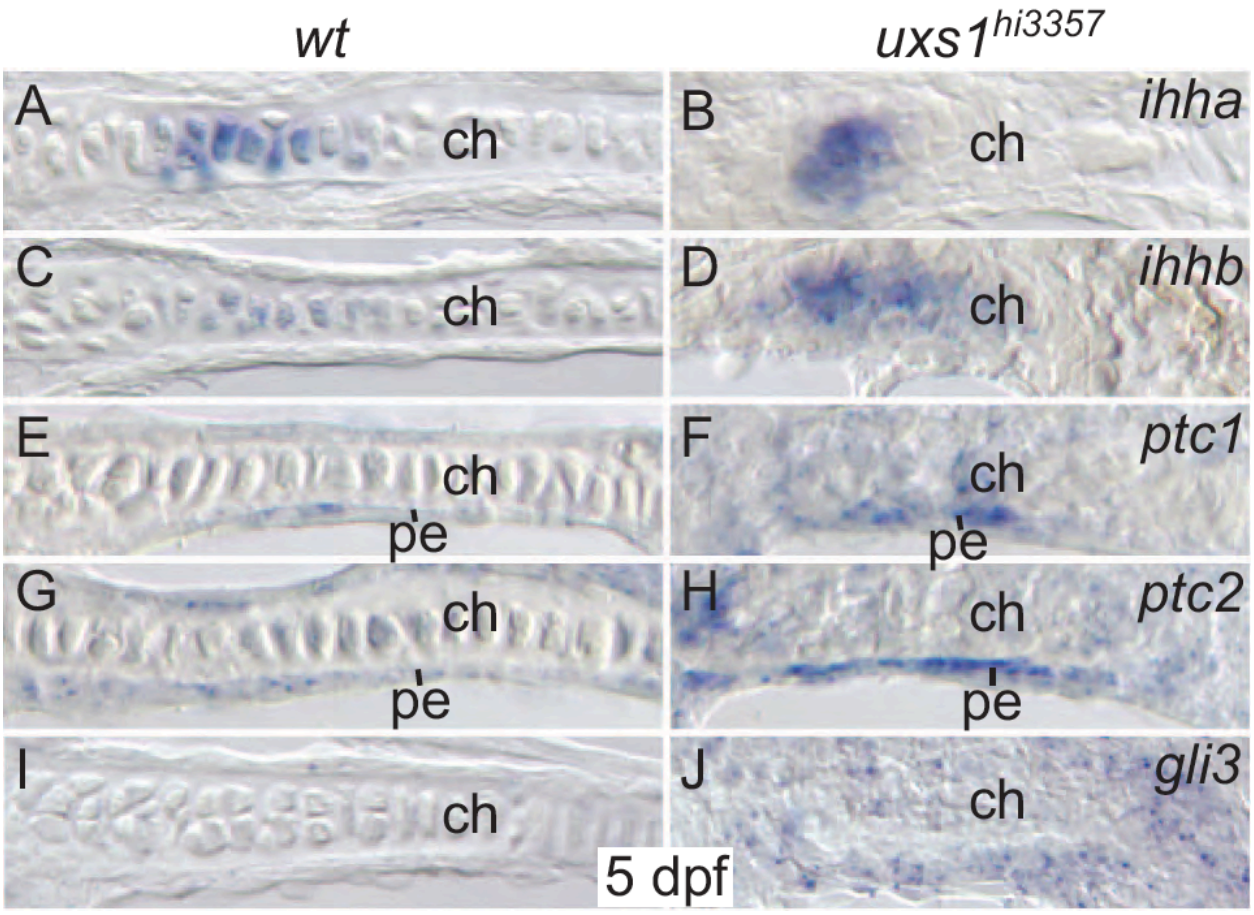


Figure 5.9. Markers of Hedgehog signaling during *uxs1* mutant endochondral ossification. A- J, *in situ* hybridization on horizontal sections through 5 dpf ceratohyal for *ihha* (A,B), *ihhb* (C,D), *ptc1* (E,F), *ptc2* (G,H), and *gli3* (I,J). Wild-type expression of *ihha* (A) and *ihhb* (B) appeared similar to that observed in *uxs1*^{hi3357} mutant chondrocytes (B,D). Compared to wild-type perichondrium (E,G,I), *uxs1* mutant perichondrium demonstrated increased expression of *ptc1* (F), *ptc2* (H), and *gli3* (J).

Discussion

A unique model for analysis of Uxs1 function in vertebrates

Uxs1 is required for many functions of proteoglycans because this enzyme converts UDP-glucuronic acid to UDP-xylose, which is the initial sugar added to the core protein of many proteoglycans (Bar-Peled et al., 2001). Our phylogenetic analyses revealed Uxs1 to be one of the most broadly and strongly conserved non-mitochondrial proteins, with 57% amino acid identity shared between the bacterial and human enzymes. Despite its central biological importance, however, mutational analysis of Uxs1 function has been limited to three experimental systems, only one of which is a multicellular organism. A mutant pathogenic fungus *Cryptococcus neoformans* shows that Uxs1 is required for fungal virulence through its action in biosynthesis of the capsule (Moyrand et al., 2002). Also, a Chinese hamster ovary *Uxs1*-mutant cell line (*pgsI-208*) was recently identified, providing insight into subcellular trafficking of UDP-xylose (Bakker et al., 2009). Finally, the *sqv-1* mutants of *C. elegans* demonstrate the requirement of Uxs1 in vulval morphogenesis and early zygotic cytokinesis (Hwang and Horvitz, 2002). With respect to vertebrates, however, zebrafish mutants offer both to reveal unique Uxs1 functions during skeletal histogenesis and morphogenesis, and to elucidate novel roles for proteoglycans.

We report here the functional consequences of two mutant alleles in zebrafish *uxs1*. The zebrafish mutation *hi3357* contains an insertion into a gene with reciprocal best BLAST relationship to human *UXS1* (Golling et al., 2002; Nissen et al., 2006), and the *mow*^{w60} mutation revealed a single disruptive nucleotide substitution in *uxs1* that changes an arginine to a histidine (R233H) in the Uxs1 protein. Arginine at this position has been conserved in all sequenced multicellular eukaryotes (plants, fungi, animals) for at least 1.6 billion years (Wang et al., 1999), as would be expected if it were important for enzymatic function. To determine whether the

R233H mutation alters Uxs1 function, we introduced the same sequence change into the human *UXS1* gene and found that it abolished enzymatic activity. Similar to *uxs1*^{hi3357}, which disrupts exon 1 and decreases transcript stability or abundance, we conclude that *mow*^{w60} is a null activity allele. Modeling the *mow*^{w60} mutation with the crystal structure of human UXS1, we found that histidine cannot satisfy the salt-bridge with Glu230 or participate in the hydrogen bonding network supported by Arg233, both of which would introduce destabilizing contacts in the well-packed dimer interface and can explain the observed loss of enzyme activity.

Uxs1 function is essential for normal proteoglycan deposition

With insufficient UDP-xylose, proteoglycan synthesis should diminish, leading to defective extracellular matrix. Our results show that in the absence of Uxs1 function, negatively charged acidic mucopolysaccharides (Scott, 1996) and proteoglycans do not accumulate normally in the extracellular matrix of embryonic cartilages. We observed loss of Alcian blue staining for sulfated glycosaminoglycans, disruption of WGA staining for N-acetylglucosamine, a sugar in the GAG chains of cartilage proteoglycans (Lang et al., 2006), and absence of immunoreactivity to both heparan sulfate proteoglycans (HSPG), a major proteoglycan of many cell types, including skeletal cells, and chondroitin sulfate proteoglycans (CSPGs), the most abundant cartilage proteoglycan. The small amount of WGA staining we observed in *uxs1* mutants is likely either due to the synthesis of the proteoglycan keratan sulfate, which does not use UDP-xylose as a sugar linker (Knudson and Knudson, 2001), or due to Uxs1 protein, *uxs1* transcript, and/or UDP-xylose deposited maternally.

Uxs1 drives skeletal morphogenesis

Our histological stains and visualization of GFP-labeled cells revealed that Uxs1 function is required for normal morphogenesis of cartilage, perichondrium, and bone. Because prechondrogenic condensations appear on schedule and in correct anatomical locations in *uxs1* mutants, zygotic expression of *uxs1* is not essential for the migration of neural crest cells or for the condensation of the resulting mesenchyme into pre-cartilage elements. When wild-type embryos secrete large quantities of ECM, however, *uxs1* mutants begin to display dramatic morphological deficits in both cartilage and perichondrium. Chondrocytes of *uxs1* mutants persist as small, rounded cells and fail to intercalate and stack. Because both of these cell movements appear to be required for overall cartilage morphogenesis (Kimmel et al., 1998), *uxs1* mutants produce shorter, thicker cartilage elements. In the context that zygotic expression of *uxs1* is necessary for proper morphogenetic behaviors of chondrocytes, it is significant that mutant embryos have concomitant defects in morphogenesis of perichondral cells. This developmental coincidence may reflect a common origin for precursors of these two cell populations (Verreijdt et al., 2002) or may illustrate another dimension of the well-documented communication between chondrocytes and perichondrium (Karsenty and Wagner, 2002).

Disruption of normal domains of *col10a1* expression and Alizarin red staining in *uxs1* mutants shows that Uxs1 activity is required for normal bone formation along both endochondral and intramembranous pathways. While perichondral dysmorphogenesis can explain bone malformation along the endochondral pathway, our data demonstrate that proper growth of bones forming by the intramembranous pathway, such as the dentary, also requires Uxs1 function. While intramembranous ossification employs proteoglycans, other mutants in the proteoglycan synthesis pathway (discussed below) do not show a strong phenotype in this class of bones.

Perhaps these genetic models have no obvious dermal bone phenotype because they do not affect both heparan and chondroitin sulfate production simultaneously (discussed below), as occurs in *uxs1* mutants. More generally, the fact that morphological defects in cartilage, perichondrium, and bone of *uxs1* mutants become apparent at the time that skeletogenic cells are secreting abundant ECM suggests that Uxs1-dependent components of the extracellular matrix play a significant role in complex morphogenetic processes. These effects may be mediated either directly by changing the extracellular environment upon which morphogenetic movements may rely or indirectly by proteoglycan influence on growth factor signaling.

Other vertebrate proteoglycan synthesis mutants

In the pathway of GAG biosynthesis, the substrate of Uxs1 (i.e., UDP-glucuronic acid) is the product of UDP-glucose dehydrogenase (Ugdh). In zebrafish, disruption of *ugdh* by mutation (*jekyll*) or antisense knockdown causes a decrease in Alcian blue staining in the pharyngeal cartilages that is less severe than that of *uxs1* (Busch-Nentwich et al., 2004; Neuhauss et al., 1996; Walsh and Stainier, 2001). If $ugdh^{jekyll}$ were a null allele, it would be expected to have a phenotypic effect as strong as that of *uxs1*. Thus, the $ugdh^{jekyll}$ Ile331-to-Asp substitution, which occupies a nonpolar pocket of the enzyme (Walsh and Stainier, 2001), is likely to be either a hypomorph or zebrafish has mechanisms other than Ugdh to supply UDP-glucuronic acid.

Mutations in genes that encode enzymes downstream of Uxs1 in GAG synthesis also display disrupted skeletogenesis. Exostosins (Ext's) attach GAG sugars to growing proteoglycan chains. Full loss of function mutants for Ext1 and Ext2 in mouse are early embryonic lethal, precluding skeletal analyses. However, mouse models of heterozygous Ext loss of function have defects in skeletal homeostasis and development (Clement et al., 2008; Lin et al., 2000; Stickens

et al., 2005). Additionally, numerous recent studies illustrate the importance of GAG sulfation in proper skeletal development. For example, Slc35b2 (formerly called Papst1) is necessary for GAG sulfation. Zebrafish mutations in *ext2, extl3,* and *slc35b2* result in embryos with short, wide pharyngeal skeletal elements exhibiting Alcian blue staining similar to, but generally less severe than, those of *uxs1* mutants (Clement et al., 2008; Karlstrom et al., 1996; Lee et al., 2004; Schilling et al., 1996). Where examined, these mutants also show reduced bone formation and aberrant bone morphogenesis, including perichondral bone. While the perichondrium appeared to be unaffected morphologically in *ext2* and *slc35b2* mutant zebrafish, our visualization of GFP-labelled cells in *uxs1* mutants clearly reveal morphological defects in formation of the perichondrium, offering a new model system in which to study molecular mechanisms of perichondral development. The similarities of *ugdh*, *uxs1*, *ext2*, and *extl3* mutants indicate that the proteoglycan biosynthetic pathway is necessary for proper histogenesis and morphogenesis of craniofacial cartilage and bone in zebrafish, but that the *uxs1* mutations described here have the most severe phenotypes. In the mouse, numerous genes in the sulfation pathway of CSPGs have been demonstrated to be important for proper ECM production in developing cartilage. Mutations in *Jaws* (Sohaskey et al., 2008), *Gpapp* (Frederick et al., 2008), and *Papss2* (Cortes et al., 2009) all result in chondrodysplasias, and are characterized specifically by under-sulfated CSPGs.

A novel role for proteoglycans during cartilage histogenesis

We hypothesize that defects in skeletal histogenesis, revealed from transcriptional profiles of such genes as *col2a1a* and *col10a1* in *uxs1* mutant cartilages, can be explained molecularly by Uxs1-dependent changes to Sox9 and Runx2 activity. Up-regulation of *col2a1a* in developing

mutant chondrocytes was associated with increased transcript levels of *sox9a* and *sox9b.* These data can be explained by the fact that Sox9 regulates *Col2a1* transcription in tetrapods and teleosts (Lefebvre et al., 1997; Yan et al., 2002; Yan et al., 2005). On the other hand, *uxs1* mutant chondrocytes failed to express the maturation marker *col10a1* on the appropriate timetable, even though *runx2b* was up-regulated in these cells. Since Runx2 is a transcriptional activator of *Col10a1* expression in tetrapods (Iwamoto et al., 2003; Komori, 2005; Yoshida et al., 2004; Zheng et al., 2003), perhaps a repressor of Runx2 protein activity is present in mutant chondrocytes. Indeed, over-expression of Sox9 in chondrocytes undergoing endochondral ossification abolishes *Col10a1* expression (Eames et al., 2004), and Sox9 protein inhibits activity of Runx2 through a direct molecular interaction (Zhou et al., 2006). Therefore, elevated expression of *sox9a* and *sox9b* in mutant chondrocytes may obviate the effects of increased *runx2b* levels, assuming the *sox9* mRNAs are translated to increased protein levels in *uxs1* mutant chondrocytes. We conclude that the molecular mechanism by which histogenic programs are altered in *uxs1* mutant chondrocytes is mediated, at least in part, through aberrant regulation of Sox9 levels and/or activity.

Why is *sox9* gene expression increased in *uxs1*-deficient chondrogenic cells? If Uxs1 promotes differentiation of *sox9*-expressing pre-chondrocytes, then in the absence of Uxs1 function, these cells would arrest and accumulate, effectively increasing *sox9* expression. However, our analyses of transgenic animals revealed no dramatic changes to the number of cells in a pre-cartilage element. Our developmental time series suggests a novel role for proteoglycans during cartilage histogenesis: they feed back onto expression of chondrogenic genes, such as *sox9*. Up-regulation of *sox9* only occurred after chondrogenic cells began to secrete abundant proteoglycans, as detected by antibodies to chondroitin sulfate, at 3 dpf. Levels

of *sox9* transcripts were similar between wild type and *uxs1* mutant chondrogenic condensations at 48 hpf, for example (data not shown). Therefore, either proteoglycan- dependent cell signaling (discussed below), or some other integrin-like mechanism by which cells directly sense proteoglycans in the ECM, may normally repress *sox9* expression during early stages of chondrocyte differentiation (Fig. 5.10). While the answer to this interesting cell biological question remains unclear, other zebrafish mutants of proteoglycan synthesis and secretion show similarly increased *sox9a* levels in differentiating chondrocytes (Clement et al., 2008; Lang et al., 2006), suggesting a novel feedback mechanism from ECM during early stages of chondrogenesis.

Alterations to cell signaling in uxs1 *mutants*

In *Drosophila* and vertebrates, proteoglycans can act as co-receptors in signal transduction pathways (Iozzo, 1998; Lin, 2004; Lin et al., 1999). We hypothesize that disruption of the ECM, as shown by the aberrant distribution of Col2a1, N-acetylglucosamine, CSPG, and HSPG in *uxs1* mutants, leads to abnormal proteoglycan-facilitated developmental signaling between chondrocytes and perichondrium (Goldring et al., 2006; Olsen et al., 2000; Provot and Schipani, 2005). HSPGs, for example, modify Hedgehog (Hh) signaling, which is an important mediator between chondrocytes and perichondral cells during development (Colnot et al., 2005; Cortes et al., 2009; Kronenberg, 2003; St-Jacques et al., 1999). In *Drosophila*, the HSPG protein core mutant *dlp* is required for proper Hh signaling (Lum et al., 2003). Also, HSPGs may act as co-factors of ADAM-dependent release and spread of Shh (Dierker et al, 2009). On the other hand, our finding that *ptc* genes are up-regulated in the perichondrium of *uxs1* mutants is in agreement with vertebrate studies of proteoglycan synthesis mutants. For example, *Ext1* mutant mice show increased Ihh activity and range of Ihh signaling (Hilton et al., 2005; Koziel et al., 2004). These

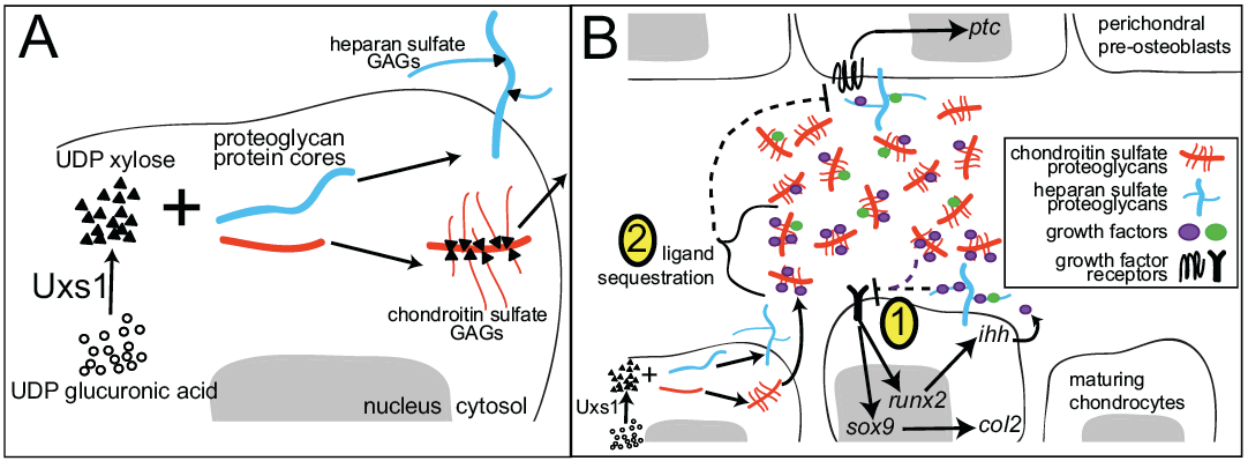


Figure 5.10. The role of Uxs1 and the extracellular matrix in skeletogenesis. A. Uxs1 converts UDP-glucuronic acid (open circles) to UDP-xylose (closed triangles), which serves as the linker sugar between the protein core (thick red and blue lines) and subsequent glycosaminoglycan deposition (thin red and blue lines) for heparan sulfate and chondroitin sulfate proteoglycans. Endoplasmic reticulum and secretory organelles are omitted for simplicity. B. Model for signaling roles of proteoglycans as they mediate interactions between chondrocytes and perichondrium during endochondral ossification. 1) Throughout chondrocyte development, proteoglycans negatively regulate growth factor signaling, such as FGF, that serves to promote *sox9* expression, which would otherwise drive *col2* transcription. In a similar fashion in maturing chondrocytes, proteoglycans inhibit *runx2* expression, which would otherwise drive expression of markers of chondrocyte maturation, such as *ihh*. 2) When Ihh is produced by maturing chondrocytes, proteoglycans negatively regulate its action on perichondral pre-osteoblasts by ligand sequestration.

data inform a model whereby proteoglycans inhibit intensity and range of Hh signaling, perhaps by ligand sequestration (Fig. 5.10). On the other hand, under-sulfation of CSPGs decreased the range of Ihh diffusion within cartilage matrix (Cortes et al, 2009). In this light, proteoglycans may exert different effects on short-range signaling of Ihh to perichondral cells versus long-range Ihh signaling to periarticular chondrocytes. In fact, these differences may reflect reliance upon HSPG- versus CSPG-mediated Hh signaling.

Many of the defects described here in *uxs1* skeletogenesis can be explained by proteoglycan-dependent alterations to growth factor signaling as chondrogenic cells embed themselves in abundant ECM. HSPGs promote a functional signaling complex between Fibroblast growth factors (FGFs) and their receptors (FGFRs) (Ornitz, 2000; Pellegrini, 2001) and HSPG sulfation is required for the interaction of FGFs with their receptors (Nakato and Kimata, 2002). *Fgfr2* is expressed in chondrogenic condensations, and FGF signaling functions upstream of *Sox9* during chondrogenesis (Coumoul and Deng, 2003; de Crombrugghe et al., 2000; Eames and Schneider, 2008; Itoh and Ornitz, 2004; Ornitz and Itoh, 2001). Interestingly, we found increased expression of the FGF-responsive gene *erm* in chondrocytes of *uxs1* mutants, suggesting that cartilage proteoglycans, such as CSPGs, may normally inhibit FGF signaling. Taken together, proteoglycan-dependent alterations to FGF signaling may explain our finding that *sox9a* and *sox9b* are up-regulated in *uxs1* mutant chondrocytes (Fig. 5.10).

uxs1 *and disease*

Besides the role of *uxs1* in skeletogenesis, defects in proteoglycan-mediated pathways can disrupt other aspects of human health. Chondroitin sulfate proteoglycan helps protect neurons against the build up of β-amyloid protein, an indicator of Alzheimer's disease progression

(Miyata et al., 2007). Proteoglycans are important for tumorigenesis, and the blocking of heparan sulfate synthesis may inhibit metastasis (Belting, 2002). Given the central role of UDP-xylose in proteoglycan biosynthesis, the *mow*^{w60} zebrafish that we report here is the first thoroughly investigated animal model deficient in *uxs1*, and will be a useful tool for studying the role of proteoglycans in health and disease.

Acknowledgements

Thanks to Ruth Bremiller and Amanda Rapp for assistance with histology and animal care, and to Jeremy Wegner and Yasuko Honjo for *erm, gli1, gli2,* and *gli3* probes. G.A.S. submitted a portion of this work as an undergraduate thesis to the Clark Honors College, University of Oregon, and offers special thanks to Nathan Tublitz and Joseph Fracchia. B.F.E. would like to thank Charles Kimmel for his support. This work was supported by grant numbers 5 F32 DE016778-03 (B.F.E.), R01 DE13834 (C.B.K.), and 5R01RR020833 and P01 HD22486 (J.H.P.) from the National Institutes of Health. The contents of this study are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH.

References

Akiyama, H., 2008. Control of chondrogenesis by the transcription factor Sox9. Mod Rheumatol. 18**,** 213-9.

Ala-Kokko, L., Baldwin, C. T., Moskowitz, R. W., Prockop, D. J., 1990. Single base mutation in the type II procollagen gene (COL2A1) as a cause of primary osteoarthritis associated with a mild chondrodysplasia. Proc Natl Acad Sci U S A. 87**,** 6565-8.

Amsterdam, A., Nissen, R. M., Sun, Z., Swindell, E. C., Farrington, S., Hopkins, N., 2004. Identification of 315 genes essential for early zebrafish development. Proc Natl Acad Sci U S A. 101**,** 12792-7.

Avaron, F., Hoffman, L., Guay, D., Akimenko, M. A., 2006. Characterization of two new zebrafish members of the hedgehog family: atypical expression of a zebrafish indian hedgehog gene in skeletal elements of both endochondral and dermal origins. Dev Dyn. 235**,** 478-89.

Bakker, H., Oka, T., Ashikov, A., Yadav, A., Berger, M., Rana, N. A., Bai, X., Jigami, Y., Haltiwanger, R. S., Esko, J. D., Gerardy-Schahn, R., 2009. Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase. J Biol Chem. 284**,** 2576-83.

Bar-Peled, M., Griffith, C. L., Doering, T. L., 2001. Functional cloning and characterization of a UDP- glucuronic acid decarboxylase: the pathogenic fungus Cryptococcus neoformans elucidates UDP-xylose synthesis. Proc Natl Acad Sci U S A. 98**,** 12003-8.

Belting, M., Borsig, L., Fuster, M.M., Brown, J.R., Persson, L., Fransson, L.A. and Esko, J.D., 2002. Tumor attenuation by combined heparan sulfate and polyamine depletion. Proc Natl Acad Sci U S A. 99**,** 371-376.

Birney, E., Andrews, T. D., Bevan, P., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cuff, J., Curwen, V., Cutts, T., Down, T., Eyras, E., Fernandez-Suarez, X. M., Gane, P., Gibbins, B., Gilbert, J., Hammond, M., Hotz, H. R., Iyer, V., Jekosch, K., Kahari, A., Kasprzyk, A., Keefe, D., Keenan, S., Lehvaslaiho, H., McVicker, G., Melsopp, C., Meidl, P., Mongin, E., Pettett, R., Potter, S., Proctor, G., Rae, M., Searle, S., Slater, G., Smedley, D., Smith, J., Spooner, W., Stabenau, A., Stalker, J., Storey, R., Ureta-Vidal, A., Woodwark, K. C., Cameron, G., Durbin, R., Cox, A., Hubbard, T., Clamp, M., 2004. An overview of Ensembl. Genome Res. 14**,** 925-8.

Busch-Nentwich, E., Sollner, C., Roehl, H., Nicolson, T., 2004. The deafness gene dfna5 is crucial for ugdh expression and HA production in the developing ear in zebrafish. Development. 131**,** 943-51.

Clement, A., Wiweger, M., von der Hardt, S., Rusch, M. A., Selleck, S. B., Chien, C. B., Roehl, H. H., 2008. Regulation of zebrafish skeletogenesis by ext2/dackel and papst1/pinscher. PLoS Genet. 4**,** e1000136.

Colnot, C., de la Fuente, L., Huang, S., Hu, D., Lu, C., St-Jacques, B., Helms, J. A., 2005. Indian hedgehog synchronizes skeletal angiogenesis and perichondrial maturation with cartilage development. Development. 132**,** 1057-67.

Cortes, M., Baria, A. T., Schwartz, N. B., 2009. Sulfation of chondroitin sulfate proteoglycans is necessary for proper Indian hedgehog signaling in the developing growth plate. Development. 136**,** 1697-706.

Coumoul, X., Deng, C. X., 2003. Roles of FGF receptors in mammalian development and congenital diseases. Birth Defects Res C Embryo Today. 69**,** 286-304.

Cubbage, C. C., Mabee, P. M., 1996. Development of the Cranium and Paird Fins in the Zebrafish *Danio rerio* (Ostariophysi, Cyprinidae). Journal of Morphology. 229**,** 121-160.

Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., Matthews, B. W., 1991. Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. Biochemistry. 30**,** 11521-9.

de Crombrugghe, B., Lefebvre, V., Behringer, R. R., Bi, W., Murakami, S., Huang, W., 2000. Transcriptional mechanisms of chondrocyte differentiation. Matrix Biol. 19**,** 389-94.

DeLano, W. L., 2002. The PyMOL User's Manual. DeLano Scientific, Palo Alto, CA, USA.

Eames, B. F., de la Fuente, L., Helms, J. A., 2003. Molecular ontogeny of the skeleton. Birth Defects Res C Embryo Today. 69**,** 93-101. Eames, B. F., Helms, J. A., 2004. Conserved molecular program regulating cranial and appendicular skeletogenesis. Dev Dyn. 231**,** 4-13.

Eames, B. F., Schneider, R. A., 2008. The genesis of cartilage size and shape during development and evolution. Development. 135**,** 3947-58.

Eames, B. F., Sharpe, P. T., Helms, J. A., 2004. Hierarchy revealed in the specification of three skeletal fates by Sox9 and Runx2. Dev Biol. 274**,** 188-200.

Frederick, J. P., Tafari, A. T., Wu, S. M., Megosh, L. C., Chiou, S. T., Irving, R. P., York, J. D., 2008. A role for a lithium-inhibited Golgi nucleotidase in skeletal development and sulfation. Proc Natl Acad Sci U S A. 105**,** 11605-12.

Goldring, M. B., Tsuchimochi, K., Ijiri, K., 2006. The control of chondrogenesis. J Cell Biochem. 97**,** 33-44.

Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S., Lin, S. Y., Nissen, R. M., Hopkins, N., 2002. Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. Nat Genet. 31**,** 135-40.

Goodrich, L. V., Johnson, R. L., Milenkovic, L., McMahon, J. A., Scott, M. P., 1996. Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. Genes Dev. 10**,** 301-12.

Hamerman, D., 1989. The biology of osteoarthritis. N. Engl. J. Med. 320**,** 1322-1330.

Hedges, S. B., 2002. The origin and evolution of model organisms. Nat Rev Genet. 3**,** 838-49.

Henikoff, S., Henikoff, J. G., 1992. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 89**,** 10915-9.

Hilton, M. J., Gutierrez, L., Martinez, D. A., Wells, D. E., 2005. EXT1 regulates chondrocyte proliferation and differentiation during endochondral bone development. Bone. 36**,** 379-86.

Holt, C. E., Dickson, B. J., 2005. Sugar codes for axons? Neuron. 46**,** 169-72.

Iozzo, R. V., 1998. Matrix proteoglycans: from molecular design to cellular function. Annu Rev Biochem. 67**,** 609-52.

Ishimizu, T., Sano, K., Uchida, T., Teshima, H., Omichi, K., Hojo, H., Nakahara, Y., Hase, S., 2007. Purification and substrate specificity of UDP-D-xylose:beta-D-glucoside alpha-1,3-17 D-xylosyltransferase involved in the biosynthesis of the Xyl alpha1-3Xyl alpha1-3Glc beta1-O-Ser on epidermal growth factor-like domains. J Biochem. 141**,** 593-600.

Itoh, N., Ornitz, D. M., 2004. Evolution of the Fgf and Fgfr gene families. Trends Genet. 20**,** 563-9. Iwamoto, M., Kitagaki, J., Tamamura, Y., Gentili, C., Koyama, E., Enomoto, H., Komori, T., Pacifici, M., Enomoto-Iwamoto, M., 2003. Runx2 expression and action in chondrocytes are regulated by retinoid signaling and parathyroid hormone-related peptide (PTHrP). Osteoarthritis Cartilage. 11**,** 6-15.

Iwasaki, M., Le, A. X., Helms, J. A., 1997. Expression of indian hedgehog, bone morphogenetic protein 6 and gli during skeletal morphogenesis. Mech Dev. 69**,** 197-202.

Izvolsky, K. I., Shoykhet, D., Yang, Y., Yu, Q., Nugent, M. A., Cardoso, W. V., 2003. Heparan sulfate-FGF10 interactions during lung morphogenesis. Dev Biol. 258**,** 185-200.

Jowett, T., Yan, Y. L., 1996. Double fluorescent in situ hybridization to zebrafish embryos. Trends Genet. 12**,** 387-9.

Karlstrom, R. O., Trowe, T., Klostermann, S., Baier, H., Brand, M., Crawford, A. D., Grunewald, B., Haffter, P., Hoffmann, H., Meyer, S. U., Muller, B. K., Richter, S., van Eeden, F. J., Nusslein Volhard, C., Bonhoeffer, F., 1996. Zebrafish mutations affecting retinotectal axon pathfinding. Development. 123**,** 427-38.

Karsenty, G., Wagner, E. F., 2002. Reaching a genetic and molecular understanding of skeletal development. Dev Cell. 2**,** 389-406.

Kasprzyk, A., Keefe, D., Smedley, D., London, D., Spooner, W., Melsopp, C., Hammond, M., Rocca-Serra, P., Cox, T., Birney, E., 2004. EnsMart: a generic system for fast and flexible access to biological data. Genome Res. 14**,** 160-9.

Kearns, A. E., Vertel, B. M., Schwartz, N. B., 1993. Topography of glycosylation and UDP-xylose production. J Biol Chem. 268**,** 11097-104.

Kimmel, C. B., Miller, C. T., Kruze, G., Ullmann, B., BreMiller, R. A., Larison, K. D., Snyder, H. C., 1998. The shaping of pharyngeal cartilages during early development of the zebrafish. Dev Biol. 203**,** 245-63.

Kirn-Safran, C. B., Gomes, R. R., Brown, A. J., Carson, D. D., 2004. Heparan sulfate proteoglycans: coordinators of multiple signaling pathways during chondrogenesis. Birth Defects Res C Embryo Today. 72**,** 69-88.

Kizawa, H., Kou, I., Iida, A., Sudo, A., Miyamoto, Y., Fukuda, A., Mabuchi, A., Kotani, A., Kawakami, A., Yamamoto, S., Uchida, A., Nakamura, K., Notoya, K., Nakamura, Y., Ikegawa, S., 2005. An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. Nat Genet. 37**,** 138-44.

Kjellen, L., Lindahl, U., 1991. Proteoglycans: structures and interactions. Annu Rev Biochem. 60**,** 443-75.

Knapik, E. W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M. C., Jacob, H. J., 1998. A microsatellite genetic linkage map for zebrafish (Danio rerio). Nat Genet. 18**,** 338-43.

Knowlton, R. G., Katzenstein, P. L., Moskowitz, R. W., Weaver, E. J., Malemud, C. J., Pathria, M. N., Jimenez, S. A., Prockop, D. J., 1990. Genetic linkage of a polymorphism in the type II procollagen gene (COL2A1) to primary osteoarthritis associated with mild chondrodysplasia. N Engl J Med. 322**,** 526-30.

Knudson, C. B., Knudson, W., 2001. Cartilage proteoglycans. Semin Cell Dev Biol. 12**,** 69-78.

Komori, T., 2005. Regulation of skeletal development by the Runx family of transcription factors. J Cell Biochem. 95**,** 445-53.

Koziel, L., Kunath, M., Kelly, O. G., Vortkamp, A., 2004. Ext1-dependent heparan sulfate regulates the range of Ihh signaling during endochondral ossification. Dev Cell. 6**,** 801-13.

Kronenberg, H. M., 2003. Developmental regulation of the growth plate. Nature. 423**,** 332-6.

Krovel, A. V., Olsen, L. C., 2004. Sexual dimorphic expression pattern of a splice variant of zebrafish vasa during gonadal development. Developmental Biology. 271**,** 190-197.

Lander, A. D., Selleck, S. B., 2000. The elusive functions of proteoglycans: in vivo veritas. J Cell Biol. 148**,** 227-32.

Lang, M. R., Lapierre, L. A., Frotscher, M., Goldenring, J. R., Knapik, E. W., 2006. Secretory COPII coat component Sec23a is essential for craniofacial chondrocyte maturation. Nat Genet. 38**,** 1198-203.

Lawson, N. D., Weinstein, B. M., 2002. In vivo imaging of embryonic vascular development using transgenic zebrafish. Dev Biol. 248**,** 307-18.

Lee, J. S., von der Hardt, S., Rusch, M. A., Stringer, S. E., Stickney, H. L., Talbot, W. S., Geisler, R., Nusslein-Volhard, C., Selleck, S. B., Chien, C. B., Roehl, H., 2004. Axon sorting in the optic tract requires HSPG synthesis by ext2 (dackel) and extl3 (boxer). Neuron. 44**,** 947-60.

Lefebvre, V., Huang, W., Harley, V. R., Goodfellow, P. N., de Crombrugghe, B., 1997. SOX9 is a potent activator of the chondrocyte-specific enhancer of the pro alpha1(II) collagen gene. Mol. Cell Biol. 17**,** 2336-2346.

Lewis, K. E., Concordet, J. P., Ingham, P. W., 1999. Characterisation of a second patched gene in the zebrafish Danio rerio and the differential response of patched genes to Hedgehog signaling. Dev Biol. 208**,** 14-29.

Lin, X., 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. Development. 131**,** 6009-21.

Lin, X., Buff, E. M., Perrimon, N., Michelson, A. M., 1999. Heparan sulfate proteoglycans are essential for FGF receptor signaling during Drosophila embryonic development. Development. 126**,** 3715-23.

Lin, X., Wei, G., Shi, Z., Dryer, L., Esko, J. D., Wells, D. E., Matzuk, M. M., 2000. Disruption of gastrulation and heparan sulfate biosynthesis in EXT1-deficient mice. Dev Biol. 224**,** 299-311.

Lum, L., Yao, S., Mozer, B., Rovescalli, A., Von Kessler, D., Nirenberg, M., Beachy, P. A., 2003. Identification of Hedgehog pathway components by RNAi in Drosophila cultured cells. Science. 299**,** 2039-45.

Miyata, S., Nishimura, Y., Nakashima, T., 2007. Perineuronal nets protect against amyloid beta-protein neurotoxicity in cultured cortical neurons. Brain Res.

Moyrand, F., Klaproth, B., Himmelreich, U., Dromer, F., Janbon, G., 2002. Isolation and characterization of capsule structure mutant strains of Cryptococcus neoformans. Mol Microbiol. 45**,** 837-49.

Nakase, T., Miyaji, T., Kuriyama, K., Tamai, N., Horiki, M., Tomita, T., Myoui, A., Shimada, K., Yoshikawa, H., 2001. Immunohistochemical detection of parathyroid hormone-related peptide, Indian hedgehog, and patched in the process of endochondral ossification in the human. Histochem Cell Biol. 116**,** 277-84.

Nakashima, K., de Crombrugghe, B., 2003. Transcriptional mechanisms in osteoblast differentiation and bone formation. Trends Genet. 19**,** 458-66.

Nakato, H., Kimata, K., 2002. Heparan sulfate fine structure and specificity of proteoglycan functions. Biochim Biophys Acta. 1573**,** 312-8.

Neuhauss, S. C., Solnica-Krezel, L., Schier, A. F., Zwartkruis, F., Stemple, D. L., Malicki, J., Abdelilah, S., Stainier, D. Y., Driever, W., 1996. Mutations affecting craniofacial development in zebrafish. Development. 123**,** 357-67.

Nissen, R. M., Amsterdam, A., Hopkins, N., 2006. A zebrafish screen for craniofacial mutants identifies wdr68 as a highly conserved gene required for endothelin-1 expression. BMC Dev Biol. 6**,** 28.

Olsen, B. R., Reginato, A. M., Wang, W., 2000. Bone development. Annu Rev Cell Dev Biol. 16**,** 191-220.

Ornitz, D. M., 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. Bioessays. 22**,** 108-12.

Ornitz, D. M., Itoh, N., 2001. Fibroblast growth factors. Genome Biol. 2**,** REVIEWS3005.

Otto, F., Thornell, A. P., Crompton, T., Denzel, A., Gilmour, K. C., Rosewell, I. R., Stamp, G. W., Beddington, R. S., Mundlos, S., Olsen, B. R., Selby, P. B., Owen, M. J., 1997. Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. Cell. 89**,** 765-71.

Pellegrini, L., 2001. Role of heparan sulfate in fibroblast growth factor signalling: a structural view. Curr Opin Struct Biol. 11**,** 629-34.

Piotrowski, T., Schilling, T. F., Brand, M., Jiang, Y. J., Heisenberg, C. P., Beuchle, D., Grandel, H., van Eeden, F. J., Furutani-Seiki, M., Granato, M., Haffter, P., Hammerschmidt, M., Kane, D. A., Kelsh, R. N., Mullins, M. C., Odenthal, J., Warga, R. M., Nusslein-Volhard, C., 1996. Jaw and branchial arch mutants in zebrafish II: anterior arches and cartilage differentiation. Development. 123**,** 345-56.

Provot, S., Schipani, E., 2005. Molecular mechanisms of endochondral bone development. Biochem Biophys Res Commun. 328**,** 658-65.

Prydz, K., Dalen, K. T., 2000. Synthesis and sorting of proteoglycans. J Cell Sci. 113 Pt 2**,** 193-205.

Renn, J., Winkler, C., Schartl, M., Fischer, R., Goerlich, R., 2006. Zebrafish and medaka as models for bone research including implications regarding space-related issues. Protoplasma. 229**,** 209-14.

Rothschild, B. M., Panza, R. K., 2007. Lack of bone stiffness/strength contribution to osteoarthritis--evidence for primary role of cartilage damage. Rheumatology (Oxford). 46**,** 246-9.

Schilling, T. F., Piotrowski, T., Grandel, H., Brand, M., Heisenberg, C. P., Jiang, Y. J., Beuchle, D., Hammerschmidt, M., Kane, D. A., Mullins, M. C., van Eeden, F. J., Kelsh, R. N., Furutani-Seiki, M., Granato, M., Haffter, P., Odenthal, J., Warga, R. M., Trowe, T., Nusslein-Volhard, C., 1996. Jaw and branchial arch mutants in zebrafish I: branchial arches. Development. 123**,** 329-44.

Scott, J. E., 1996. Alcian blue. Now you see it, now you don't. Eur J Oral Sci. 104**,** 2-9. Shimoda, N., Knapik, E. W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H., Fishman, M. C., 1999. Zebrafish genetic map with 2000 microsatellite markers. Genomics. 58**,** 219-32.

Simoes, B., Conceicao, N., Viegas, C. S., Pinto, J. P., Gavaia, P. J., Hurst, L. D., Kelsh, R. N., Cancela, M. L., 2006. Identification of a promoter element within the zebrafish colXalpha1 gene responsive to runx2 isoforms Osf2/Cbfa1 and til-1 but not to pebp2alphaA2. Calcif Tissue Int. 79**,** 230-44.

Sohaskey, M. L., Yu, J., Diaz, M. A., Plaas, A. H., Harland, R. M., 2008. JAWS coordinates chondrogenesis and synovial joint positioning. Development. 135**,** 2215-20.

St-Jacques, B., Hammerschmidt, M., McMahon, A. P., 1999. Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. Genes Dev. 13**,** 2072-86.

Stickens, D., Zak, B. M., Rougier, N., Esko, J. D., Werb, Z., 2005. Mice deficient in Ext2 lack heparan sulfate and develop exostoses. Development. 132**,** 5055-68.

van den Berg, S., Lofdahl, P. A., Hard, T., Berglund, H., 2006. Improved solubility of TEV protease by directed evolution. J Biotechnol. 121**,** 291-8.

Velleman, S. G., 2000. The role of the extracellular matrix in skeletal development. Poult Sci. 79**,** 985-9.

Velleman, S. G., Clark, S. H., 1992. The cartilage proteoglycan deficient mutation, nanomelia, contains a DNA polymorphism in the proteoglycan core protein gene that is genetically linked to the nanomelia phenotype. Matrix. 12**,** 66-72.

Verreijdt, L., Vandervennet, E., Sire, J. Y., Huysseune, A., 2002. Developmental differences between cranial bones in the zebrafish (Danio rerio): some preliminary light and TEM observations. Connect Tissue Res. 43**,** 109-12.

Vertel, B. M., Walters, L. M., Flay, N., Kearns, A. E., Schwartz, N. B., 1993. Xylosylation is an endoplasmic reticulum to Golgi event. J Biol Chem. 268**,** 11105-12.

Vortkamp, A., Lee, K., Lanske, B., Segre, G. V., Kronenberg, H. M., Tabin, C. J., 1996. Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. Science. 273**,** 613-22.

Walker, M., Kimmel, C., 2007. A two-color acid-free cartilage and bone stain for zebrafish larvae. Biotech Histochem. 82**,** 23-8.

Walsh, E. C., Stainier, D. Y., 2001. UDP-glucose dehydrogenase required for cardiac valve formation in zebrafish. Science. 293**,** 1670-3.

Wang, D. Y., Kumar, S., Hedges, S. B., 1999. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc Biol Sci. 266**,** 163-71.

Woods, I. G., Wilson, C., Friedlander, B., Chang, P., Reyes, D. K., Nix, R., Kelly, P. D., Chu, F., Postlethwait, J. H., Talbot, W. S., 2005. The zebrafish gene map defines ancestral vertebrate chromosomes. Genome Res. 15**,** 1307-14.

Yan, Y. L., Hatta, K., Riggleman, B., Postlethwait, J. H., 1995. Expression of a type II collagen gene in the zebrafish embryonic axis. Dev Dyn. 203**,** 363-76.

Yan, Y. L., Miller, C. T., Nissen, R. M., Singer, A., Liu, D., Kirn, A., Draper, B., Willoughby, J., Morcos, P. A., Amsterdam, A., Chung, B. C., Westerfield, M., Haffter, P., Hopkins, N., Kimmel, C., Postlethwait, J. H., 2002. A zebrafish sox9 gene required for cartilage morphogenesis. Development. 129**,** 5065-79.

Yan, Y. L., Willoughby, J., Liu, D., Crump, J. G., Wilson, C., Miller, C. T., Singer, A., Kimmel, C., Westerfield, M., Postlethwait, J. H., 2005. A pair of Sox: distinct and overlapping functions of zebrafish sox9 co-orthologs in craniofacial and pectoral fin development. Development. 132**,** 1069-83.

Yoshida, C. A., Komori, T., 2005. Role of Runx proteins in chondrogenesis. Crit Rev Eukaryot Gene Expr. 15**,** 243-54.

Yoshida, C. A., Yamamoto, H., Fujita, T., Furuichi, T., Ito, K., Inoue, K., Yamana, K., Zanma, A., Takada, K., Ito, Y., Komori, T., 2004. Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. Genes Dev. 18**,** 952-63.

Zheng, Q., Zhou, G., Morello, R., Chen, Y., Garcia-Rojas, X., Lee, B., 2003. Type X collagen gene regulation by Runx2 contributes directly to its hypertrophic chondrocyte-specific expression in vivo. J Cell Biol. 162**,** 833-42.

Zhou, G., Zheng, Q., Engin, F., Munivez, E., Chen, Y., Sebald, E., Krakow, D., Lee, B., 2006. Dominance of SOX9 function over RUNX2 during skeletogenesis. Proc Natl Acad Sci U S A. 103**,** 19004-9.

APPENDIX B

ROLE OF PACKING DEFECTS IN THE EVOLUTION OF ALLOSTERY AND INDUCED FIT IN HUMAN UDP-GLUCOSE DEHYDROGENASE[4]

[4] Reprinted with permission from Kadirvelraj, R., Sennett, N. C., Polizzi, S. J., Weitzel, S., and Wood, Z. A. *Biochemistry 50*, 5780-5789. Copyright (2011) American Chemical Society

S.J. Polizzi contributed to the ultracentrifugation analysis of UDP-glucose dehydrogenase.

Abstract

Allosteric feedback inhibition is the mechanism by which metabolic end products regulate their own biosynthesis by binding to an upstream enzyme. Despite its importance in controlling metabolism, there are relatively few allosteric mechanisms understood in detail. This is because allostery does not have an identifiable structural motif, making the discovery of new allosteric enzymes a difficult process. The lack of a conserved motif implies that the evolution of each allosteric mechanism is unique. Here we describe an atypical allosteric mechanism in human UDP-α-D-glucose 6-dehydrogenase (hUGDH) based on an easily acquired and identifiable structural attribute: packing defects in the protein core. In contrast to classic allostery, the active and allosteric sites in hUGDH are present as a single, bifunctional site. Using two new crystal structures, we show that binding of the feedback inhibitor, UDP-α-D-xylose, elicits a distinct induced-fit response; a buried loop translates ~4 Å along and rotates ~180 degrees about the main-chain axis, requiring surrounding side-chains to repack. This allosteric transition is facilitated by packing defects, which negate the steric conformational restraints normally imposed by the protein core. Sedimentation velocity studies show that this repacking favors the formation of an inactive hexameric complex with unusual symmetry. We present evidence that hUGDH and the unrelated enzyme dCTP-deaminase have converged to very similar atypical allosteric mechanisms using the same adaptive strategy: the selection for packing defects. Thus, the selection for packing defects is a robust mechanism for the evolution of allostery and induced-fit.

Introduction

The nucleotide sugar UDP-α-D-glucuronic acid (UGA) is an essential substrate in two biochemical pathways with special relevance to cancer biology: (i) maintenance of the extracellular matrix through proteoglycan and hyaluronan biosynthesis and (ii) Phase II metabolism of drugs. The first pathway controls metastasis; disrupting proteoglycan or hyaluronan biosynthesis has been shown to attenuate tumor growth and progression (Belting, 2002; Sanderson, 2005; Fuster, 2005; Itano, 2008; Stern, 2008). The second pathway (Phase II metabolism) limits the effectiveness of chemotherapeutics by conjugating glucuronic acid to drugs. These glucuronides are rapidly eliminated from the body. In fact, some colon and lung cancers have been shown to use glucuronidation as an effective drug resistance mechanism (Cummings, 2002; Cummings, 2003; Cummings, 2006; Cummings, 2004; Cecchin, 2009; Oguri, 2004; Takahashi, 1997). Thus, understanding how UGA biosynthesis is regulated is an important goal in developing strategies to reduce metastasis, alter the pharmacokinetic properties of chemotherapeutics, and sensitize tumors that use glucuronidation as a drug resistance mechanism.

Human UDP-α-D-glucose 6-dehydrogenase (hUGDH) catalyzes the NAD^+ dependent oxidation of UDP-α-D-glucose (UDG) to UDP-α-D-glucuronic acid. In addition to its roles in maintaining the extracellular matrix and drug metabolism, UGA is also the substrate for the biosynthesis of UDP-α-D-xylose (UDX) (Fig. 6.1 A). UDX is an essential nucleotide-sugar in proteoglycan biosynthesis, and a feedback inhibitor that controls hUGDH activity (Neufeld, 1965; Bakker, 2009). Even though UDX was first recognized as a feedback inhibitor to hUGDH more than 45 years ago (Neufeld, 1965), the mechanism of inhibition was poorly understood. Based on the sigmoidal inhibition kinetics, it was initially believed that the feedback mechanism

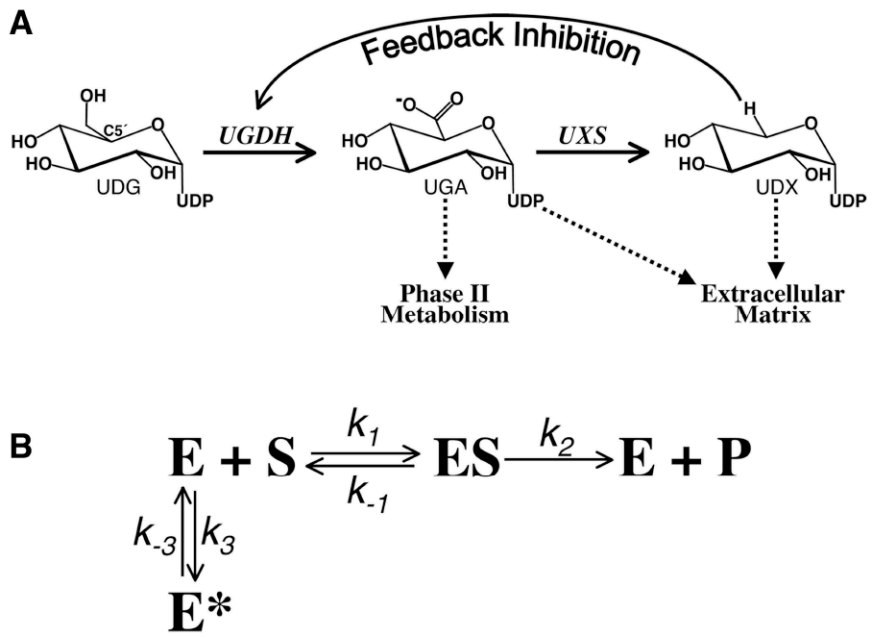


Figure 6.1. Regulation of UDP-glucose dehydrogenase. (A) UDP-glucose dehydrogenase (UGDH) oxidizes UDP-glucose (UDG) to UDP-glucuronic acid (UGA), which is then decarboxylated by UDP-xylose synthase (UXS) to produce the feedback inhibitor UDP-xylose (UDX). The C5′ atom of UDG is labeled. Dashed lines indicate the metabolic pathways in which UGA and UDX are essential substrates. (B) Enzyme hysteresis occurs when the transition between the inactive (E^*) and active (E) forms of an enzyme is not at equilibrium with the formation and turnover of the Michaelis complex (ES). At low substrate concentrations, k_3[E] competes with k_1[E][S] to reduce the apparent reaction rate. At high substrate concentrations, the rate at which the Michaelis complex forms (k_1[E][S]) is faster than the inactivation step (k_3[E]), effectively trapping the enzyme in the active state.

of hUGDH would be allosteric (Neufeld, 1965). In later studies, Gainey *et al.* identified hysteresis in the kinetics of UDX binding, and proposed an alternative model for the cooperativity observed in the feedback inhibition of hUGDH (Gainey, 1975). Hysteretic enzymes exhibit a slow conformational change between active and inactive states (Ainslie, 1972; Frieden, 1979). This slow transition can produce a phenomenon known as kinetic cooperativity (Ainslie, 1972; Frieden, 1979). Briefly, the rate at which the Michaelis complex forms depends on the concentration of substrate (k_1[E][S]) (Fig. 6.1 B). At low substrate concentrations, the slow inactivation step (k_3[E]) can compete with the slow formation of the Michaelis complex and reduce the apparent rate of the reaction. At higher substrate concentrations, the rapid formation of the Michaelis complex effectively traps the enzyme in the active state. Since the concentration of active enzyme is increasing with substrate, the resulting substrate saturation curves will be sigmoidal due to a kinetic, not allosteric process. Later work by Dickinson suggested that hysteresis in hUGDH was due to the slow dissociation of an inactive oligomeric complex to form the active species (Dickinson, 1988). Thus, if UDX binding stabilizes or induces the formation of the inactive oligomer, feedback inhibition could exhibit kinetic cooperativity.

The aim of this paper is to determine the feedback inhibition mechanism of hUGDH. Here we show that UDX inhibition of hUGDH involves an atypical form of allostery; the active and allosteric sites in hUGDH are present as a single, bifunctional site. Binding of the feedback inhibitor UDX induces a remarkable repacking of the protein core that converts the enzyme into an inactive, low symmetry complex. We also present evidence that hUGDH and an unrelated enzyme, dCTP-deaminase, both converged to similar, atypical allosteric mechanisms by selecting for large packing defects, such as cavities, voids and pockets, (Richards, 1977) near pre-existing, smaller defects. These packing defects are key to the distinct induced-fit

conformations that accommodate the binding of the substrate or the allosteric inhibitor. Thus, packing defects may represent hot spots for the evolution of allosteric networks and induced-fit.

Materials and Methods

Protein Expression and Purification

The sequence for hUGDH was synthesized (Geneart) and cloned into a pet15b plasmid vector modified to contain a tobacco etch virus (TEV) cleavage site prior to the Met start codon of the gene. Rosetta (DE3) pLysS cells containing the construct were grown at 37 °C to an OD_{A600} of 0.9 to 1.0 (four L) and then induced overnight at 20 °C with 0.15 mM Isopropyl-β-D-thiogalactoside. Cells were sonicated in 100 mL of 50 mM phosphate pH 7.8, 300 mM NaCl. Recombinant hUGDH was purified using Talon® metal affinity resin, and the His-tag removed with codon-optimized TEV protease (van den Berg, 2006). The purified protein was dialyzed into its storage buffer, 25 mM Tris pH 8, 50 mM NaCl and 2 mM β-mercaptoethanol and concentrated to 12 mg/mL.

Crystallization, Structure Solution and Analysis

Protein containing 5 mM of UDP-xylose (Carbosource, CCRC, at the University of Georgia) and either 5 mM NAD^+ (Crystal form **1**), or 5 mM of NADH (Crystal form **2**) was used for crystallization screens. Crystals were grown at 26 °C using the hanging drop vapor diffusion method with 2 μl drops mixed 1:1 (protein:reservoir). The reservoir contained: 6% heptanediol, 10% PEG 3350 and 100 mM HEPES pH 7.5 for hUGDH:UDX:NAD^+ for Crystal form **1** (monoclinic $P2_1$); or 15% hexanediol, 10% PEG 3350 and 100 mM HEPES pH 7 for hUGDH:UDX:NADH for Crystal form **2** (orthorhombic $P2_12_12_1$). Crystals grew as triangular

plates within ~7-10 days from both conditions. Crystals were cryoprotected with 20% glycerol containing 2 mM of the appropriate ligands and plunged into liquid nitrogen prior to data collection. Diffraction data were collected at the SER-CAT 22-ID beamline at the Argonne National Lab using a MAR 300 mm CCD detector, a wavelength of 1 Å and an oscillation step of 1.0° for Crystal form **1** and 0.5° for Crystal form **2**. Data were processed with XDS (Kabsch, 2010) and 5% of the data were set aside for cross-validation. Crystal parameters, data collection and refinement statistics are summarized in Table 6.1.

The structures were solved by molecular replacement using PHASER (McCoy, 2007) and chain A of PDB entry 2Q3E as a search model (bound ligands and water were removed prior to the search). Both structures (Crystal forms **1** and **2**) contained a complete, broken-hexamer in the asymmetric unit. Initial model refinement involved iterative cycles of manual rebuilding of the peptide using COOT (Emsley, 2004) and refinement with REFMAC5 (Vagin, 2004). This was followed by the location and refinement of bound ligands and ordered water molecules. Loop residues 382-388 were disordered in both structures. Protein domain motion was analyzed using the program DynDom (Hayward, 1998). Packing defects were identified using CASTp (Dundas, 2006) with a probe radius 1.4 Å. All figures were generated with Pymol (http://pymol.souceforge.net) unless otherwise specified.

Analytical Ultracentrifugation

Sedimentation velocity studies were performed using an Optima XLA analytical ultracentrifuge (Beckman Coulter). For all samples, 9.1 μM hUGDH was dialyzed into buffer (25mM Tris pH 8.0, 175 mM NaCl) with or without 79 μM ligand (UDG or UDX). For analysis, samples were loaded into cells with 12 mm centerpieces and equilibrated for ~1 hour at 20 °C.

Cells were loaded into an An60 Ti rotor and run at 50,000 rpm. Data was collected at a wavelength of 280 nm using a radial step size of 0.003 cm. The partial specific volume of 0.7384 mL/g was calculated from the amino acid sequence. The buffer density of 1.00616 g/mL and the viscosity of 0.010253 Poise were calculated using the program SEDNTERP (Laue, 1992). All data were analyzed using the program SEDFIT (Brown, 2006). Continuous sedimentation coefficient distribution analyses were restrained by maximum entropy regularization at a confidence interval P=0.68. The baseline, meniscus, frictional coefficient and systematic time-invariant and radial-invariant noise were fitted. The rmsd values for all reported experiments were 0.009 OD or less.

Results and Discussion

UDX inhibition reduces the symmetry of the hUGDH hexamer.

We have solved the crystal structures of two different UDX-inhibited hUGDH complexes: hUGDH:UDX:NAD^+ and hUGDH:UDX:NADH (Table 6.1). The hUGDH:UDX:NAD^+ complex crystallized in the spacegroup $P2_1$ and was refined to 2.5 Å resolution. The hUGDH:UDX:NADH complex crystallized in the higher symmetry spacegroup $P2_12_12_1$, but only diffracted to 3.1 Å resolution. The root mean squared deviation (r.m.s.d.) between the two crystal forms is 0.28 Å for 458 corresponding Cα atoms. The N- (residues 1-212) and C-terminal domains (residues 323-466) of hUGDH each contain a Rossmann fold for binding NAD^+ and UDG, respectively (Fig. 6.2 A). The terminal domains are connected by an α-helical domain (residues 213-322) that forms a dimer-building interface and buries ~2600 $Å^2$ of surface per monomer. The structure of the UGDH dimer is highly conserved. Two crystal structures of *Streptococcus pyogenes* UGDH have been described in literature (Campbell, 2000):

Table 6.1. Data collection and refinement statistics (molecular replacement).

	Crystal form **1**	Crystal form **2**
Data collection		
PDB entry	3PTZ	3PRJ
Space group	$P2_1$	$P2_12_12_1$
Unit cell dimensions (a,b,c,α,β,γ)	89.1, 196.7, 111.7, 90.0, 111.9, 90.0	111.8, 160.6, 205.9, 90.0, 90.0, 90.0
R_{meas}	9.2 (48)[a]	8.5 (42.6)
$R_{mrgd\text{-}F}$	11.8 (52.2)	10.3 (44.2)
$I / \sigma I$	15.9 (3.2)	17 (3.8)
Completeness (%)	96.5 (90.3)	96.9 (92.5)
Redundancy	3.8 (3.4)	4.1 (3.5)
Refinement		
Resolution (Å)	2.5	3.1
No. reflections	113593	62511
R_{work} / R_{free}	0.19/0.26	0.23/0.27
No. atoms	22754	22213
Protein	21502	21597
Ligand	468	468
Water	784	148
***B*-factors**		
Protein	31.7	54.7
Ligand	32	50.7
Water	28.8	27.8
Stereochemical Ideality		
Bond lengths (Å)	0.018	0.014
Bond angles (°)	1.7	1.4
ϕ,ψ Most favored (%)	88.3	88.9
ϕ,ψ Allowed (%)	11.6	10.9
ϕ,ψ Generously Allowed (%)	0.1	0.2

[a]Values in parentheses are for highest-resolution shell.

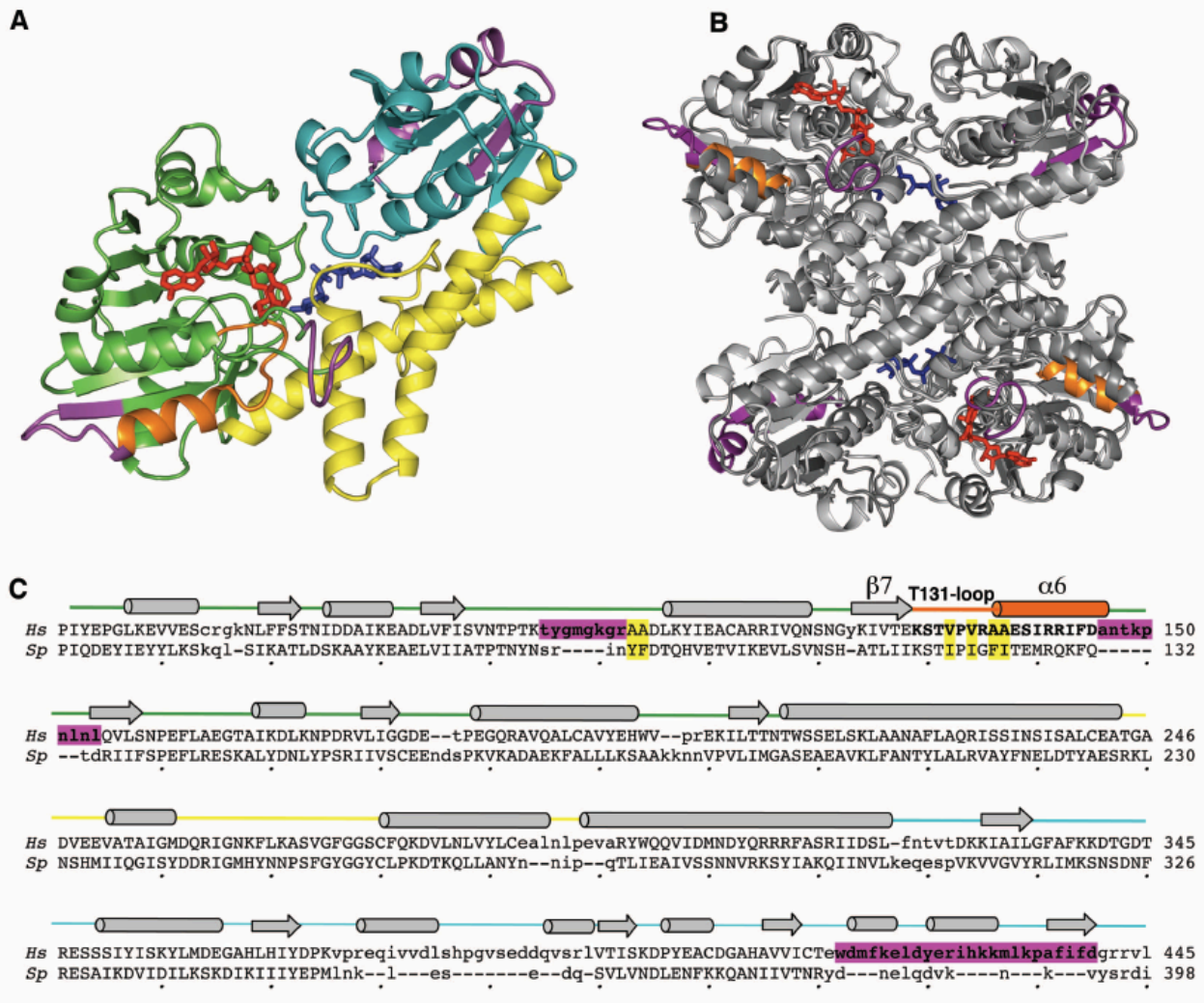


Figure 6.2. Human and bacterial UGDH are highly conserved. (A) Cartoon rendering of hUGDH depicting the N-terminal domain (green), the α-helical dimer interface (yellow) and the C-terminal domain (teal). Inserted elements that make up the hexamer-building interface are colored purple (see text for details). The Thr131-loop and α6 helix are orange. UDX (blue) and NADH (red) are depicted as sticks. (B) The conserved structure of the UGDH dimer. The human UGDH:UDX:NAD^+ complex is colored light grey, and the bacterial UGDH:UDX:NAD^+ complex (PDB entry IDLI) is dark grey. The Thr131-loop, the α6 helix and the inserted elements in the human enzyme are colored as above. (C) Structure based sequence alignment of human (residues 50-445) and *S. pyogenes* (residues 43-398) UGDH. α-helices (cylinders) and β-sheets (arrows) are shown with the color of the line connecting the structural elements identifying the N-terminal (green), dimerization (yellow) and C-terminal (teal) domains respectively. Inserted elements that make up the hexamer-building interface in hUGDH are highlighted in purple. The large-to-small amino acid substitutions surrounding the Thr131-loop in the human enzyme are highlighted in yellow (see text for details). The Thr131-loop and α6 helix are colored orange and shown in bold letters.

*Sp*UGDH:UDX:NAD^+ and *Sp*UGDH:UGA:NADH. Despite a low sequence identity of 23%, the entire hUGDH and *Sp*UGDH dimers overlay with an r.m.s.d of 1.8 Å for 732 corresponding Cα atoms (Fig. 6.2 B,C).

In addition to our new structures, the Structural Genomics Consortium (Oxford, England) has deposited four crystal structures of hUGDH: two mixed substrate/product complexes (hUGDH:UDG:NADH (PDB entry 2Q3E) and hUGDH:UGA:NAD^+ (PDB entry 2QG4)); hUGDH in complex with a covalent thioester intermediate (PDB entry 3KHU); and an apo form with an active site mutation (PDB entry 3ITK). In contrast to our inhibited structures, these four hUGDH structures reveal a hexamer displaying point group symmetry 32 (or a trimer of dimers) hereafter referred to as the 'high-symmetry hexamer' (Fig. 6.3 A). The hexamer-building interface in hUGDH is formed from a series of inserted loops not conserved in *Sp*UGDH (Fig. 6.2). In contrast, both of our UDX-inhibited structures reveal a low symmetry complex (point group symmetry 2) that we call the 'broken-hexamer' (Fig. 6.3 B). The broken-hexamer conformation is related to the high-symmetry hexamer by a 10.5° rigid-body rotation about an axis centered in the hexamer-building interface (Fig. 6.3 C). Previous studies have shown that hUGDH forms a hexamer in solution (Sommer, 2004), but our results suggest two distinct hexameric complexes are possible. To resolve this apparent ambiguity we investigated the hUGDH complexes using sedimentation velocity. These studies reveal that apo-UGDH is a complex distribution in solution, consisting predominantly of hexamers and dimers, with a small amount of monomers and tetramers also present (Fig. 6.3 D). Adding 79 μM of substrate UDG ($K_m \cong$ 10-20 μM (Huh, 2004; Sommer, 2004)) has little or no effect on the distribution (Fig. 6.3 D). These results show that the high-symmetry hexamer is a relatively weak complex and that the high protein concentrations during crystallization stabilized the structure. In contrast, 79 μM

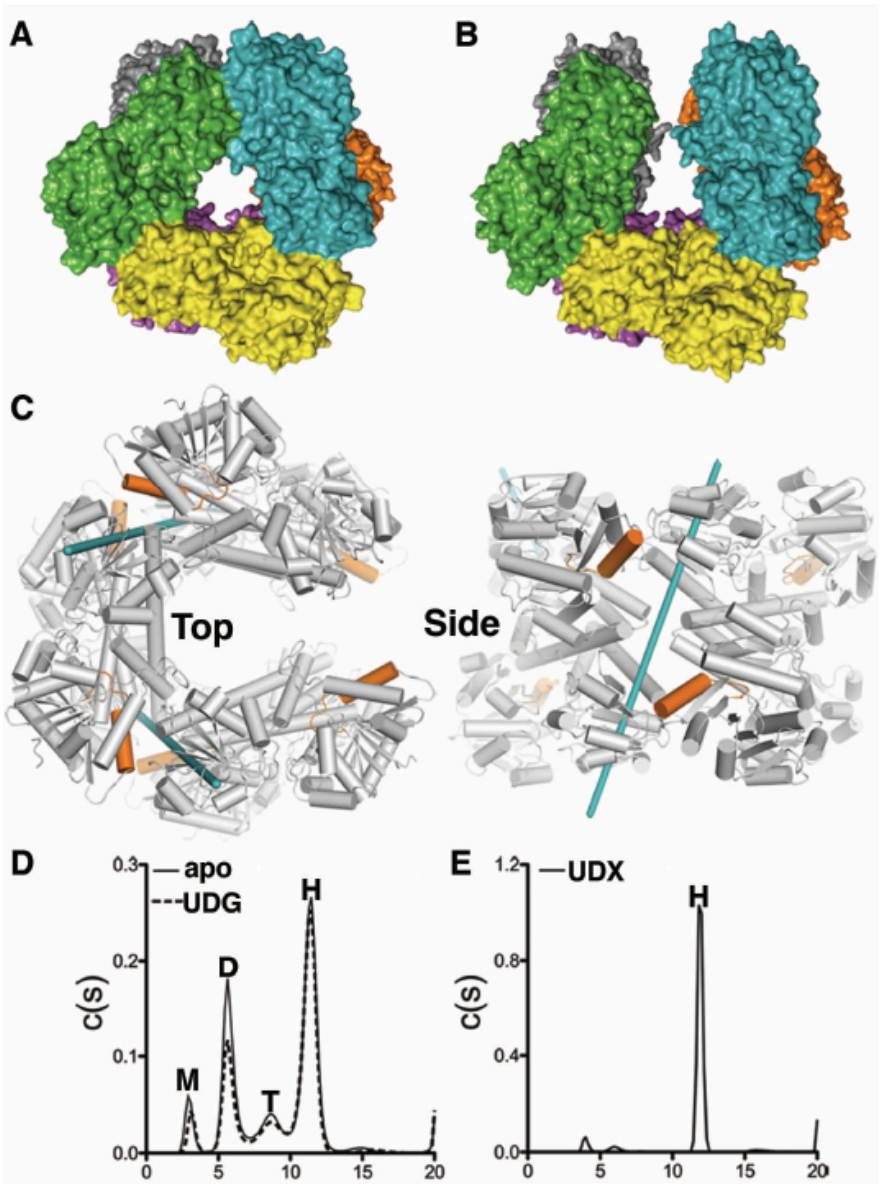


Figure 6.3. Feedback inhibition disrupts the oligomeric structure of hUGDH. (A) Solvent accessible surface of the high-symmetry conformation of UDG bound hUGDH hexamer (PDB entry 2Q3E) with each monomer colored differently. (B) The broken-hexamer form of UDX-inhibited hUGDH. (C) Top and side views of the broken-hexamer with rigid-body rotation axis depicted as teal rods passing through the hexamer-building interface. The Thr131-loop and α6 helix are colored orange (see text). (D) Sedimentation velocity results of apo and UDG bound hUGDH (solid and dashed lines, respectively). The c(s) distribution reveals a heterogeneous mixture of sedimenting species: monomer (M, 3.1 S), dimer (D, 5.6 S), tetramer (T, 8.7 S) and hexamer (H, 11.5 S). (E) UDX stabilizes a hexameric species at 12.7 S.

of inhibitor UDX ($K_i \cong 10$ μM (Neufeld, 1965)) has a profound effect on the enzyme conformation, and stabilizes a hexameric species with a sedimentation coefficient of 12.7 S (Fig. 6.3 E). This is slightly more than the S values of the apo and UDG bound hexamers (11.5 S) and suggests a difference in conformational states. Alternatively, the smaller S value of the apo and UDG bound hexamers is also consistent with the existence of a rapid equilibrium with the lower molecular weight species. Our observation of the same structure in two different space groups is strong evidence that UDX-inhibited hUGDH forms the broken-hexamer conformation in solution, based on the following: (i) UDX stabilizes the hexameric complex; (ii) crystal nucleation selects for low energy conformations (Zhang, 1995); and (iii) crystal lattice forces are too weak to deform a protein structure (Zhang, 1995; Dickerson, 1994).

The Thr131-loop is the allosteric switch that stabilizes the broken-hexamer.

UDX lacks the C5′-hydroxymethyl found in the substrate, UDG (Fig. 6.1 A). To determine how the small difference between UDX and UDG disrupts the hexamer, we compared our structures to the 2 Å resolution high-symmetry hexamer complex hUGDH:UDG:NADH (PDB entry 2Q3E). The monomers of the high-symmetry and broken-hexamer complexes superimpose with an r.m.s.d. of 0.7 Å for 456 corresponding Cα atoms. In the high-symmetry hexamer, the NAD cofactor and UDG substrate bind in extended conformations in the active site, positioning the nicotinamide and the pyranose at the junction of the two terminal domains (Fig. 6.2 A). The active site structure is conserved with *Sp*UGDH, which has already been described in detail (Campbell, 2000). Specifically, the C5′-hydroxymethyl (C6′-OH) makes hydrogen bonds with Asn224 and the catalytic base Lys220, and places the C6′ atom of UDG within ~3.4 Å of the S_γ of Cys276, the active site nucleophile (Fig. 6.4 A). Thr131 donates a hydrogen bond to the

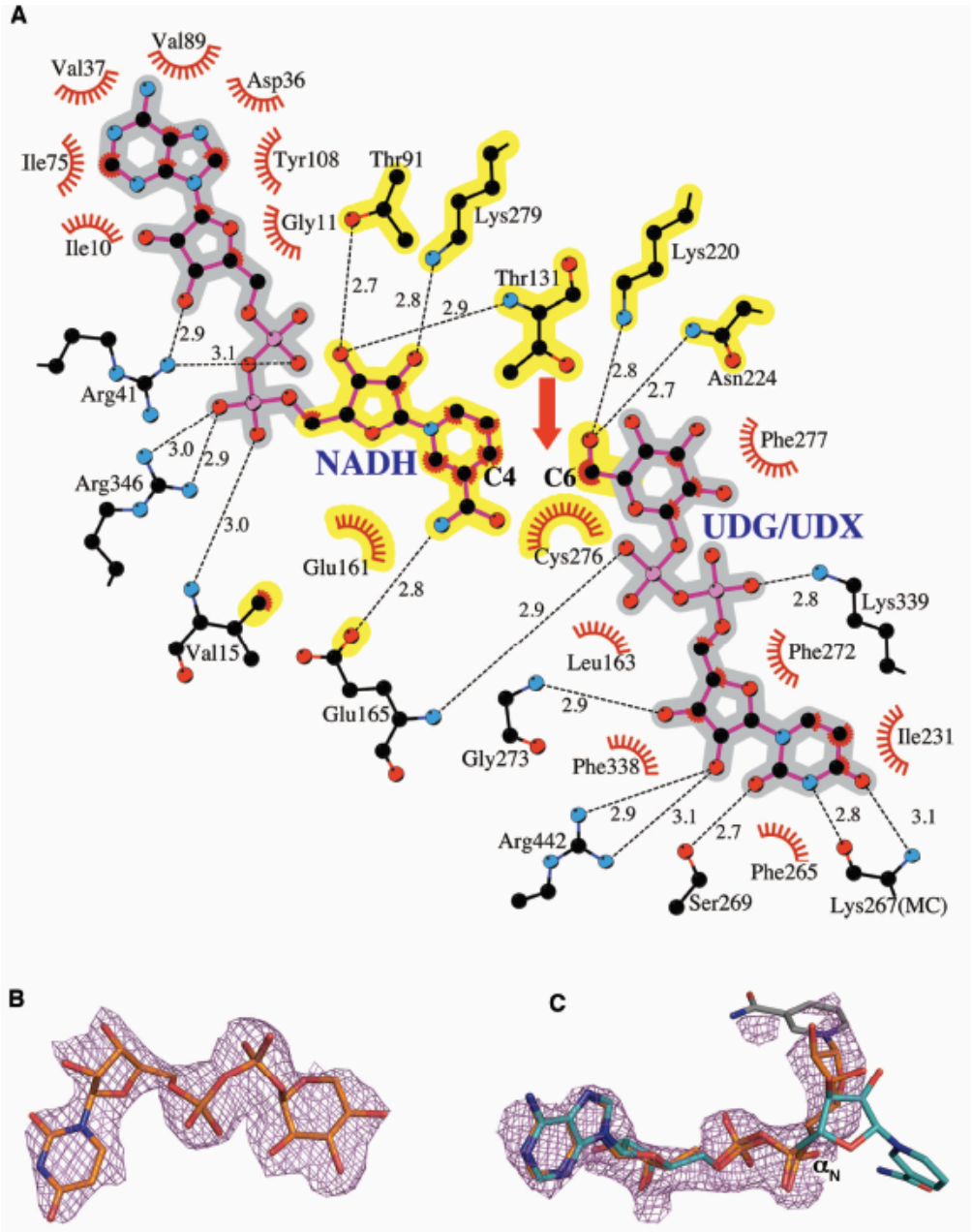


Figure 6.4. Substrate and inhibitor interactions in the active site of hUGDH. (A) This ligand plot depicts the hydrogen bonds (dashed lines) and packing interactions (red, feathered lines) between substrate and cofactor in the active site of the enzyme. Interactions that are lost in the inhibited state are shaded yellow. The lack of a C5′ hydroxymethyl (labeled C6) induces the Thr131 to move into the active site (red arrow), blocking the nicotinamide ring from binding in a productive conformation. **(**B**)** An F_o-F_c difference map calculated following molecular replacement and an initial round of restrained refinement, but prior to modeling of the UDX molecule. The 2.5 Å resolution map is contoured at 3σ, and the ligand is from the final refined coordinates (PDB entry 3PTZ). (C) Same difference density as above, but for the NAD^+ ligand (teal sticks), illustrating the disorder of the nicotinamide ring (grey sticks). The model for NADH bound in the high-symmetry hexamer (PDB entry 2Q3E) is superimposed to illustrate the rotation of the α_N torsion angle (O3-P_N-O5′$_N$-C5′$_N$).

ribose of the nicotinamide, positioning the C4 atom of the pyridine near the C6′ atom of the pyranose (Fig. 6.4 A). The most significant changes between the active and inhibited complexes are localized to the protein core between the active site and the hexamer-building interface. The feedback inhibited structure shows that UDX binds to all six active sites in the broken hexamer (Fig. 6.4 B). The absence of the C5′-hydroxymethyl in UDX would be expected to create a void in the active site. Instead, Thr131 has moved ~4 Å to pack against the C5′ atom of the xylose (Fig. 6.4 A and 6.5 A,B). This inhibitor-specific induced-fit response prevents the cofactor from binding in a productive conformation; the α_N torsion angle of the nicotinamide (O3-P_N-O5′$_N$-C5′$_N$) rotates from -83° in the active conformation to 68° in the inhibited structure (Fig. 6.4 C, 6.5 B). This conformation flips the nicotinamide into a solvent-exposed pocket where it is disordered, presumably due to poor packing.

Thr131 is part of a buried loop (residues 129-135, hereafter referred to as the Thr131-loop) connected to the α6 helix that straddles the rotation axes in the hexamer-building interface (Fig. 6.3 C). The repacking of the Thr131-loop causes the α6 helix to tilt by 10.6°, creating a surface that compliments the rigid body rotations that form the broken-hexamer conformation (Fig. 6.3 C and Fig. 6.5 A,B). The rotations bury ~172 Å^2 more surface in each hexamer-building interface (776 ± 29 Å^2 per monomer), consistent with our observation that the UDX-bound hexamer is more stable than the high-symmetry complex (Fig. 6.3 D,E). Thus, the Thr131-loop is the molecular switch that allosterically changes the affinity of the hexamer-building interface in response to the UDX binding. Rather than ratcheting the hexamer into the broken conformation, it is likely that the weak high-symmetry complex dissociates, allowing the interface to repack into the more stable, low symmetry conformation. There are no significant conformational differences between the exposed and buried hexamer-building interfaces in the

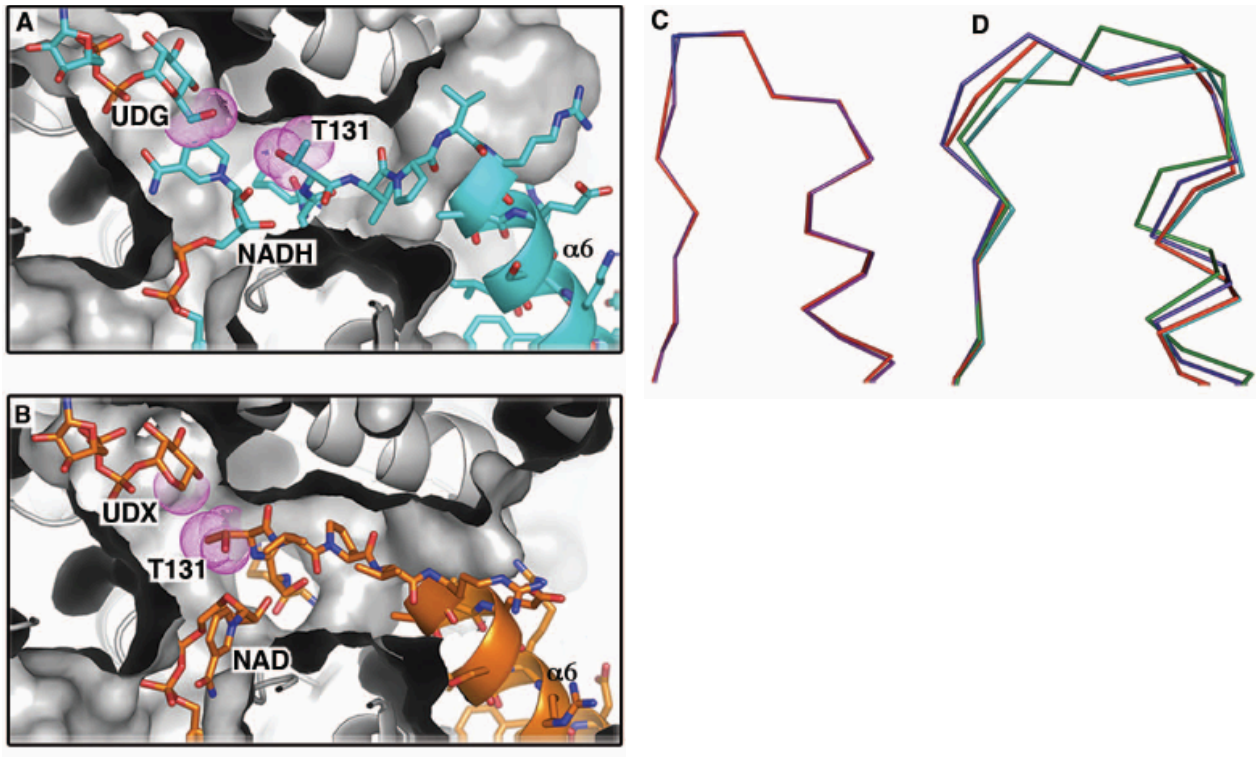


Figure 6.5. The Thr131-loop is key to the allosteric transition of hUGDH. (A) Cutaway of the molecular surface (grey) of the high-symmetry hexamer (PDB entry 2Q3E) depicting the 'active' conformation of the Thr131-loop and α6 helix (teal sticks and cartoon). The buried Thr131-loop was excluded from the surface rendering, giving the appearance that it is in a tunnel passing through the protein core. UDG and NADH (sticks) are bound in the active site. The van der Waals radii (purple dots) are shown for the C5′-hydroxymethyl of UDG and the side chain of Thr131. (B) Same view as above, but with UDX and NAD^+ bound in the broken-hexamer. In response to the smaller van der Waals radii of the C5′ atom (purple dots) of UDX, the Thr131-loop (orange) has moved into the active site, causing the α6 helix to tilt in the hexamer-building interface. In this conformation, Thr131 prevents the nicotinamide from binding in a productive conformation. (C) Superimposed C-alpha traces showing the structural variation of the Thr131-loop from the broken-hexamer complexes: red for the 2.5 Å and purple for the 3.1 Å resolution complexes, respectively. (D) Superimposed C-alpha traces of the most structurally divergent Thr-131 loops from the high-symmetry hexamers: teal (2Q3E), purple (3KHU), red (2QG4) and green (3ITK).

broken-hexamer complex. It appears that the asymmetry of the complex is due to the repacking of the core, which produces a rotated interface incompatible with the formation of the high-symmetry hexamer.

Additional evidence linking the structure of the Thr131-loop to the stability of the complex comes from a comparison of the available structures. A superposition of the most structurally divergent monomers in the 2.5 Å and 3.1 Å resolution broken-hexamer complexes shows that the corresponding Cα atoms of residues 131-134 have an r.m.s.d. of 0.34 Å (Fig. 6.5 C). This variation in loop conformation is less than what would be expected from the coordinate error of the models (Luzzati coordinate error 0.31 Å and 0.61 Å for the 2.5 Å and 3.1 Å resolution structures, respectively). In contrast, the Thr131-loops in the high-symmetry hexamers show a much larger variation; a superposition of PDB entry 2Q3E onto entries 3KHU, 2QG4 and 3ITK yields r.m.s.d.'s of 0.89 Å, 1.41 Å and 2.79 Å, respectively, for residues 131-134 (Fig. 6.5 D). The 3KHU and 2QG4 structures contain a reaction intermediate and the product UGA, respectively. The variability in the structure of the Thr131-loop in these structures likely reflects the conformational changes that would be expected to occur during the catalytic cycle. In the apo-structure (PDB ID 3ITK), the Thr131 is mutated to alanine, and the loss of the threonine Cγ packing contributions to the protein core likely adds to the conformational differences (r.m.s.d of 2.79 Å). These observations imply that the high-symmetry complexes have a more flexible hexamer-building interface, consistent with the sedimentation velocity studies showing a weaker complex in solution (Fig. 6.3 D,E).

Packing defects facilitate the allosteric transition of the Thr131-loop.

The induced-fit response of hUGDH to UDX requires a remarkable rearrangement of buried residues in the protein core (Fig. 6.6, Table 6.2): (i) the Cα atoms of the Thr131-loop translate through the core by ~4 Å, shifting the register of the buried residues; (ii) the ϕ,ψ torsion angles of the Thr131-loop rotate, flipping Val132 and Pro133 by ~180° around the axis of main-chain; and (iii) eight buried residues surrounding the loop change rotameric states. This repacking is facilitated by packing defects in the protein core. The structure of the high-symmetry hexamer (2Q3E) shows that five large packing defects (totaling 389 ± 10 $Å^3$) in the protein core are in contact with at least one residue in the Thr131-loop (Fig. 6.7 A). In the broken-hexamer conformation, the repacking of the core is reflected by changes in the volume (321 ± 32 $Å^3$), shape and location of the packing defects (Fig. 6.7 B). Several water molecules are buried in the cavities, providing a means to satisfy the main-chain hydrogen bonding requirements of the Thr131-loop in both conformations (Fig. 6.S1 A,B). In the high-symmetry conformation, two of the main-chain hydrogen bonds of the buried Thr131-loop are satisfied by one of the buried waters (Fig. 6.S1 A). In the broken-hexamer conformation, the water structure rearranges so that two buried waters participate in three hydrogen bonds with the peptide backbone (Fig. 6.S1 B).

How did hUGDH evolve the packing degeneracy needed to facilitate allosteric communication via the Thr131-loop? The equivalent Thr131-loop in the bacterial *Sp*UGDH structures adopts the same conformation seen in the dimers of the high-symmetry hUGDH hexamers (Fig. 6.7 C). Unlike the human enzyme, there is no induced-fit response upon UDX binding in *Sp*UGDH. The lack of a structural response is expected, since *S. pyogenes* does not make UDX and the activity of *Sp*UGDH is likely to be regulated by tyrosine phosphorylation, as

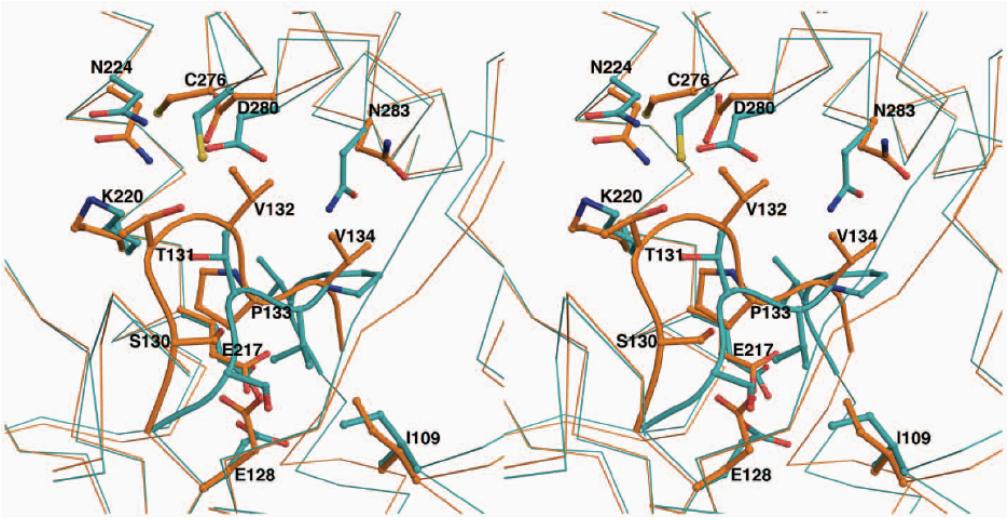


Figure 6.6. The allosteric transition requires the core to repack. Stereo view depicting the core packing degeneracy surrounding the Thr131-loop in the high-symmetry hexamer (teal) superimposed onto the broken-hexamer (orange). The buried Thr131-loop undergoes a screw-like translation of ~4 Å, flipping Val132 and Pro133 by ~180° around the axis of main-chain. The eight buried residues that change rotameric states (Table 6.2) are shown.

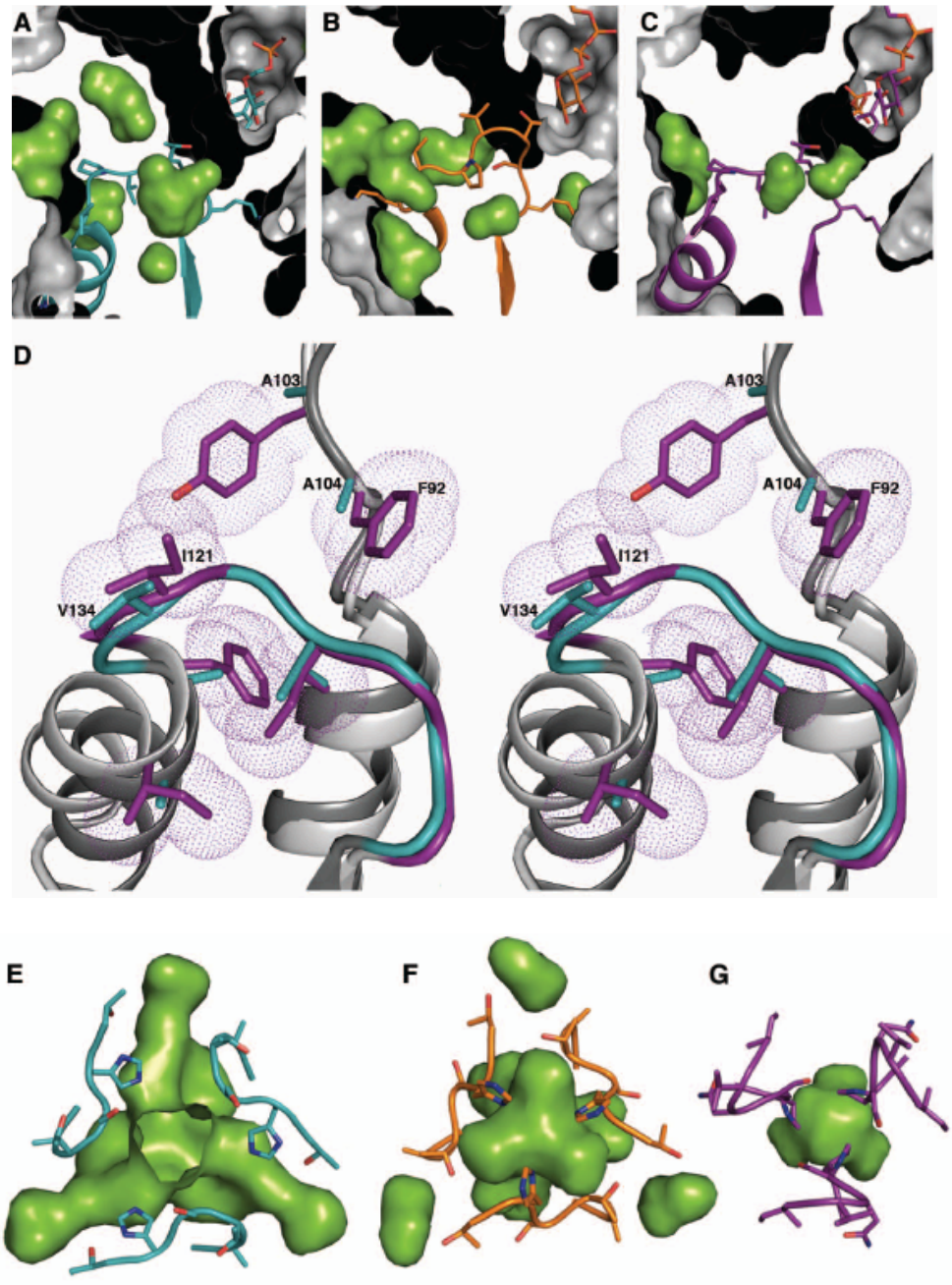
Table 6.2. Changes in buried dihedrals that contribute to the packing degeneracy.

Side Chain Rotameric Shifts				
Residue	ΔChi1[a]	ΔChi2	ΔChi3	ΔChi4
Ile 109	15°	101°		
Glu 128	28°	91°	105°	
Glu 217	11°	33°	2°	
Lys 220	19°	8°	34°	105°
Asn 224	23°	42°		
Cys 276	117°			
Asp 280	107°	69°		
Asn 283	94°	74°		

Main Chain Torsional Changes		
Residue	ΔPhi	ΔPsi
Thr 131	16°	157°
Val 132	3°	63°
Pro 133	37°	36°
Val 134	57°	39°

[a] dihedral angles are absolute differences between chains A in hUGDH:UDG:NADH (PDB entry 2Q3E) and hUGDH:UDX:NAD^+.

Figure 6.7. Packing defects facilitate alternate core arrangements in hUGDH and dCTP deaminase. (A) Cutaway of the molecular surface of the high-symmetry hUGDH structure (PDB entry 2Q3E) showing the packing defects in the core (cavities and deep pockets, green surface) surrounding the buried Thr131 loop (teal). UDG (teal, sticks) is depicted in the active site (upper right of panel). Cavity surfaces were calculated using a 1.4 Å probe. (B) Same orientation as above, depicting the changes in the packing defects that occur in the UDX bound broken-hexamer conformation. (C) Same orientation as above, illustrating the smaller packing defects surrounding Thr131-equivalent loop in *Sp*UGDH:UDX (PDB entry 1DLI). Due to the smaller cavity size, the surfaces were generated using a 1.3 Å probe radius for clarity. (D) Stereo view illustrating the six large-to-small substitutions that have occurred during the evolution of packing degeneracy. hUGDH (teal, PDB entry 2Q3E) is superimposed onto *Sp*UGDH:UDX (purple, PDB entry 1DLI), (Fig. 6.S1 C). Purple dots depict the van der Waals radii of the large hydrophobic residues in the bacterial structure. (E) Cavities (green surfaces) in the interface and core of allosteric *M. tuberculosis* dCTP-DA (PDB entry 2QLP, teal) shown looking down the three-fold axis of the active enzyme. The A115-loops in the active site of the three subunits are shown. (F) Same view as above, except showing the changes in packing defects in inhibited *M. tuberculosis* dCTP-DA (PDB entry 2QXX, orange). (G) Same view as above, but depicting the smaller cavities in the homologous, non-allosteric *M. tuberculosis* dUTPase (PDB entry 1SLH, purple). All cavity surfaces were generated using a 1.4 Å probe.


A
B
C
D
A103
A104
F92
I121
V134
A103
A104
F92
I121
V134
E
F
G

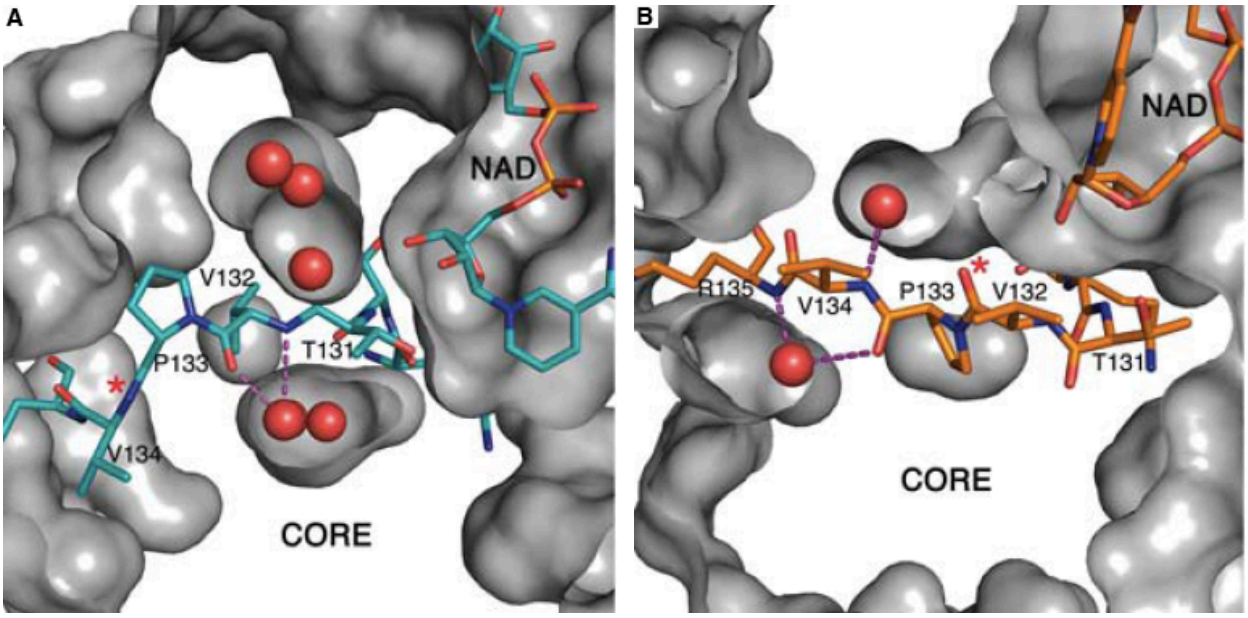


Figure 6.S1. Buried water hydrogen bonding to main-chain of the Thr131-loop. Cutaway of a molecular surface showing hydrogen bonding (dashed lines) between the main-chain of the Thr131-loop and buried water molecules (red spheres) in the cavities of hUGDH. (A) Chain A of 2Q3E. The asterisk identifies the buried amide of V143, which is not participating in a hydrogen bond, but is near a deep pocket on the surface. Structural fluctuations may allow this pocket to widen, allowing the buried amide access to bulk solvent. Waters are W7, W224, W230, W387, W889 and W1004. (B) Chain A in the broken-hexamer conformation. The asterisk identifies the buried peptide carbonyl of V132, which appears to be unsatisfied, but is exposed to a large cavity in the core. It may be that disordered water shields the acceptor. Waters are W8, W25, W277 and W373 in chain A of 3PTZ.

in other bacteria (Lacour, 2008). A comparison of the protein cores surrounding the equivalent Thr131-loops in hUGDH and *Sp*UGDH reveals a dramatic structural divergence with respect to atomic packing density. The core surrounding the Thr131-loop in *Sp*UGDH buries a relatively small cavity (64 $Å^3$) (Fig. 6.7 C). In contrast, the evolution of the human enzyme has selected for six large-to-small cavity-forming substitutions surrounding the Thr131-loop (Fig. 6.7 D and 6.2 C). These six substitutions represent a loss of hydrophobic residues and have left significant packing defects in the core. Engineering cavities using large-to-small substitutions has been shown to destabilize protein structure through a loss of packing interactions equivalent to ~22 cal mol^{-1} $Å^{-3}$ (Eriksson, 1992; Joh, 2009). This implies that the cavities near the Thr131-loop in hUGDH represent a loss of ~8 kcal mol^{-1} in packing interactions relative to *Sp*UGDH. Since the unfolded states of globular proteins are typically 4 to 15 kcal mol^{-1} less stable than that of the folded conformations (Kumar, 2006), the severe packing defects in hUGDH must have been selected for during evolution. It is worth noting that the slightly smaller volume of the defects in the broken-hexamer conformation (321 ± 32 $Å^3$) implies a favorable gain of ~1.6 kcal mol^{-1} in packing energy, and may contribute to the observed stability of the broken-hexamer conformation.

We would expect the packing defects in the hUGDH structures to be reflected in the B-factors of the Thr131-loop. Instead, the Thr131-loop is well ordered in both the high- and low-symmetry hexamers, and is comparable to other buried regions of the protein. Other factors likely contribute to the ordering of the Thr131-loop, specifically: (i) the hydrogen bonding and packing interactions between Thr131 and the ligands in the active site (Fig 6.4 A and 6.5 A,B); and (ii) the packing of the α6 helix in the hexamer-building interface (Fig. 6.3 C). Since the two conformations of the Thr131-loop are ordered, it is reasonable to assume that both conformations

are energetically stable. This is consistent with the observed hysteresis in hUGDH activity (Dickinson, 1988); if the activation energy necessary to change the conformation of the Thr131-loop is relatively high, then a lag in progress curves could occur.

dCTP-deaminase evolved atypical allostery by selecting for packing defects .

The bifunctional active/allosteric site in hUGDH represents an unusual form of heterotropic allostery. A similar feedback inhibition mechanism has also been described in an unrelated class of enzymes, the dCTP deaminases (dCTP-DA) (Helt, 2008; Johansson, 2007). The dCTP-DA enzymes convert the nucleotide dCTP to dUTP, and the downstream metabolic product, dTTP, acts as a feedback inhibitor to regulate flux (Beck, 1975; Johansson, 2007). The feedback mechanism of dCTP-DA shares important features with hUGDH, specifically: (i) the substrate (dCTP) and inhibitor (dTTP) are structurally similar and compete for binding to the active site (Fig. 6.S2 A,B); and (ii) the enzyme adopts a distinct induced-fit conformation in response to inhibitor binding (Fig. 6.7 E,F). As in hUGDH, a chemical difference between the substrate and the inhibitor (the C5 methyl of dTTP) induces a buried loop (the A115-loop) to rearrange and change the conformation of the dCTP-DA complex (Fig. 6.7 E,F and 6.S2 C,D). Our analysis shows that this conformational change is facilitated by large packing defects in the core and subunit interfaces (Fig. 6.7 E,F and 6.S2 C,D). The dCTP-DAs are closely related to the non-allosteric dUTPases (Johansson, 2003; Johansson, 2007). In fact, the active sites in the dCTP-DAs from *M. tuberculosis* and *M. jannaschii* conserve both the dUTPase catalytic residues and activity (Johansson, 2003; Helt, 2008). By comparing the dUTPases with the dCTP-DAs, we observe that the latter enzymes have selected for large packing defects during evolution to accommodate the inhibitor specific induced-fit response (Fig. 6.7 E-G and 6.S2 C-E).

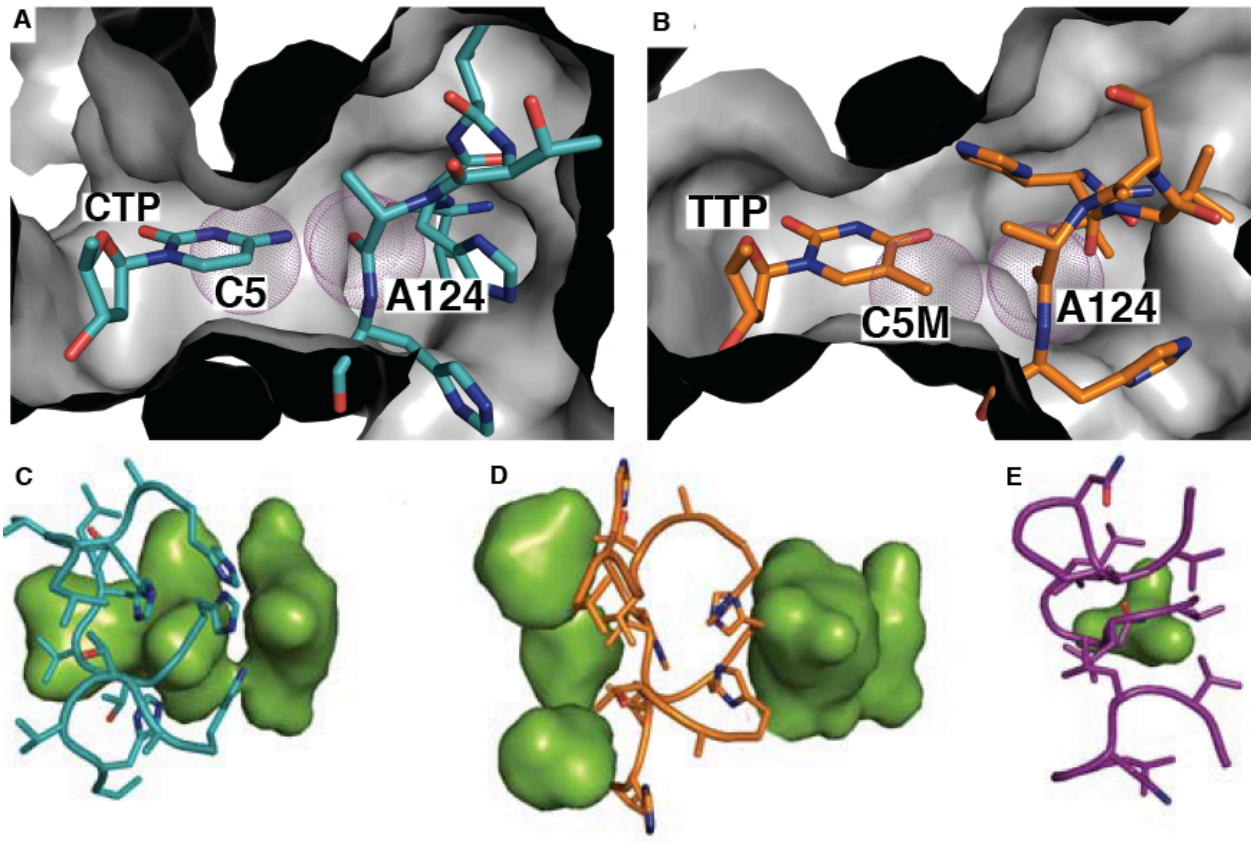


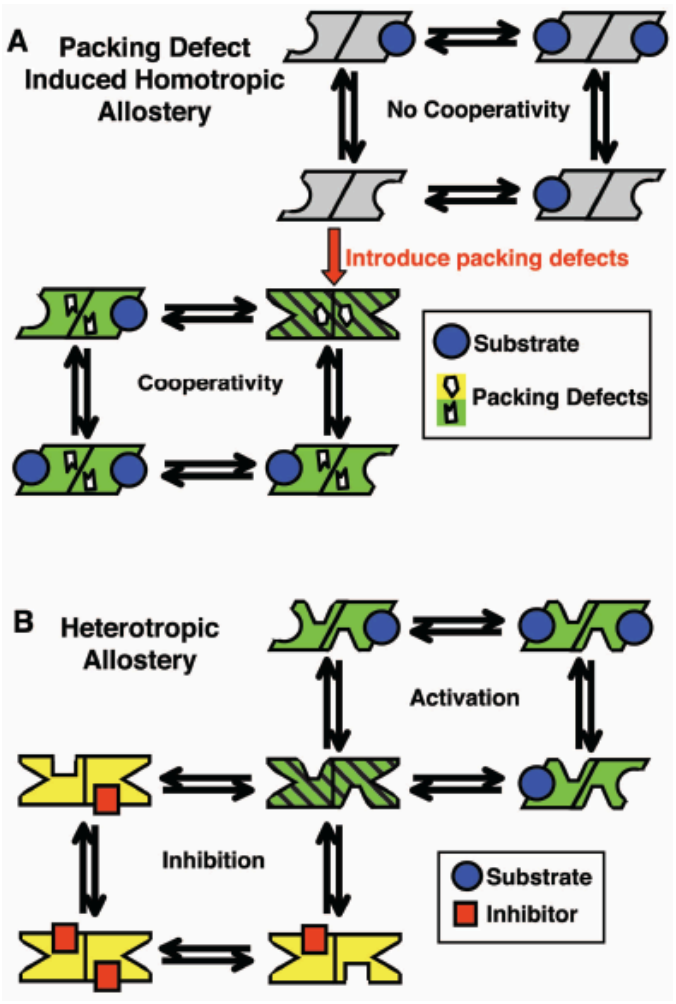
Figure 6.S2. The evolution of packing degeneracy in dCTP deaminase. Packing defects were selected for to facilitate core packing degeneracy. (A) Cutaway depicting the interaction between the dCTP substrate (CTP) and the A124-loop in the active site of *E. coli* dCTP-DA (1XS4). The van der Waals packing radii (purple dots) of the C5 atom of dCTP and the peptide carbonyl atom of A124. (B) Same view as above, except showing the conformational changes in *E. coli* dCTP-DA (2J4Q) that occur due to the C5-methyl (C5M) on the feedback inhibitor dTTP (TTP). (C) Cavities (green surfaces) in the interface and core of allosteric *E. coli* dCTP-DA (1XS4). The three-fold axis of the trimeric complex is horizontal and the A124-loops of the three subunits are shown. (D) Same view as above, except showing the changes in packing defects in inhibited *E. coli* dCTP-DA (2J4Q). (E) Same view as above, but depicting the smaller cavities in the homologous, non-allosteric *E. coli* dUTPase (1SYL). All cavity surfaces were generated using a 1.4 Å probe, except for the small cavity in 1SYL that was generated using a 1.3 Å probe.

The role of packing defects in the evolution of allostery and induced-fit.

Allostery occurs when an effector binding to one site of a protein changes the affinity of another, distant binding site. In the case of hUGDH, UDX binding to the active site increases the affinity between dimers to form the broken-hexamer (Fig. 6.3 E). But how does an allosteric interaction network evolve? The core of a natively folded protein is a relatively rigid structure, with atoms typically packed at densities approaching or surpassing that of close packed spheres (Richards, 1977). Previous studies have shown that engineering a packing defect into a non-allosteric enzyme can be sufficient to induce substrate (or homotropic) cooperativity (Fig. 6.8 A) (Kolodziej, 1996; Scrutton, 1992; Kuo, 1989). Packing defects add flexibility to the protein core (Bahar, 1997), thus introducing a defect can produce a more conformational diverse ensemble of structures (Gunasekaran, 2004). Homotropic allostery will occur only if a subset of the resulting ensemble adopts a conformation favoring the formation of the Michaelis complex (Gunasekaran, 2004). Substrate binding stabilizes protein structure and reduces flexibility (Ma, 1999). In an oligomeric complex, the substrate-induced stabilization of one subunit would be reflected in an ordering of the interface (Fig. 6.8 A). Thus, homotropic cooperativity is due to symmetry; ligand binding to one subunit would restrain the flexibility of an adjacent subunit to conformations resembling the Michaelis complex, increasing the affinity for the substrate.

In the case of classic heterotropic allostery, the effector and the substrate are chemically distinct molecules that bind to separate sites (Fig. 6.8 B). In this way, a metabolic end-product can control its own synthesis by binding to and inhibiting an earlier enzyme in the same biochemical pathway. Still, the evolution of a specialized allosteric binding site for a heterotropic effector would appear to be a complex process. In contrast, both hUGDH and dCTP-DA have evolved an atypical form of heterotropic allostery based on a single, bifunctional binding site: the

Figure 6.8. Allosteric regulation in enzymes. (A) Cartoon of a dimeric enzyme illustrating how homotropic allostery can be induced by packing defects. In the absence of packing defects (grey shading), the conformational state of the dimer favors the active, high substrate affinity state (right-leaning diagonal). Packing defects (white cutouts) add flexibility to the complex, producing a conformational ensemble (black hatching) with different substrate affinities (triangular binding site). Substrate binding (blue circle) to one subunit selects for the high substrate affinity conformation (green shading), explaining the observed homotropic cooperativity. (B) A similar cartoon depicting classic heterotropic feedback inhibition in enzymes. In the absence of the substrate, the enzyme is represented as a conformational ensemble (black hatching) with different affinities for substrate (blue circle) and inhibitor (red square). Substrate binding shifts the ensemble to the high substrate affinity state (green). Inhibitor binding selects for the low substrate affinity states (yellow). (C) *Allokleidic* regulation in enzymes. Due to packing defects (white cutouts) selected during evolution, the *allokleidic* enzyme is represented as a conformational ensemble (black hatching) with varying affinities for the substrate (blue circle) and inhibitor (red square). Substrate or inhibitor binding act as above, with the exception that both ligands compete for the same bifunctional *allokleidic* site.


A
Packing Defect Induced Homotropic Allostery
No Cooperativity
Introduce packing defects
Cooperativity
Substrate
Packing Defects
B
Heterotropic Allostery
Activation
Inhibition
Substrate
Inhibitor
C
Allokleidic Regulation
Activation
Inhibition
Substrate
Inhibitor
Packing Defects

active site can support catalysis or regulate enzyme activity in response to the chemical identity of the bound ligand (Fig. 6.5 A,B). When the substrate binds, the Thr131-loop is positioned to support catalysis. When the inhibitor binds, the Thr131-loop adjusts to conform to the chemical differences and switches the enzyme into an inactive state. Since the Thr131-loop directly connects the hexamer-building interface to the active site, it is possible to allosterically transfer information between adjacent molecules (Fig. 6.5 A,B). To explain specificity, Emil Fischer proposed that a substrate molecule fits into the active site much like a key into a lock. In keeping with the original nomenclature of Monod (Monod, 1963), and with deference to Fischer's model, we call this dual specificity active site an *allokleidic* (*allos* 'other' and *kleidi* 'key') site. Since most downstream metabolites are structurally similar to the substrates earlier in the pathway, the selection for an *allokleidic* mechanism offers an evolutionary advantage; relatively few adaptations to the active site would be required to bind a structurally similar effector. Our observation that two unrelated metabolic checkpoint enzymes (hUGDH and dCTP-DA) have converged to similar feedback mechanisms is good evidence that the evolution of an *allokleidic* active site offers a robust alternative to the selection for a separate allosteric site.

The *allokleidic* mechanism we have described is based on alternate core packing arrangements as part of an inhibitor specific induced-fit response. Still, the steric complementarity constraints of a densely packed core (Richards, 1977) make it difficult to envision alternate stable packing arrangements for buried amino acids (Fig. 6.6). So, how did hUGDH and dCTP-DA evolve core packing degeneracy? The non-allosteric homologs of both families contain relatively small packing defects in the same regions of the protein core where the allosteric members evolved *allokleidic* switches (Fig. 6.7 and 6.S2). Residues near existing packing defects are known to exhibit higher rates of sequence divergence relative to densely

packed residues (Liao, 2005). Together, these observations imply that the ancestral proteins in the UGDH and dCTP-DA families acquired the packing defects prior to the divergence of the non-allosteric and allosteric enzymes. Afterward, the allosteric members of the two families followed a relatively simple evolutionary strategy: the selection for cavities near existing packing defects. Packing defects are known to contribute to promiscuous binding (Tokuriki, 2009) and conformational flexibility (Bahar, 1997), two valuable adaptive traits in the evolution of induced-fit and allostery. Thus, the UGDH and dCTP-DA examples support our hypothesis that packing defects represent hot spots for the evolution of both allosteric networks and induced-fit mechanisms.

Acknowledgements

We thank Walt Baase (University of Oregon) for helpful discussions, William Lanzilotta (University of Georgia) and the SER-CAT beamline personnel (Advanced Photon Source, Argonne National Lab) for help with data collection.

References

Ainslie, G. R., Jr., Shill, J. P., and Neet, K. E. (1972) Transients and cooperativity. A slow transition model for relating transients and cooperative kinetics of enzymes, *J Biol Chem 247*, 7088-7096.

Bahar, I., Atilgan, A. R., and Erman, B. (1997) Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, *Fold Des 2*, 173-181.

Bakker, H., Oka, T., Ashikov, A., Yadav, A., Berger, M., Rana, N. A., Bai, X., Jigami, Y., Haltiwanger, R. S., Esko, J. D., and Gerardy-Schahn, R. (2009) Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase, *J Biol Chem 284*, 2576-2583.

Beck, C. F., Eisenhardt, A. R., and Neuhard, J. (1975) Deoxycytidine triphosphate deaminase of Salmonella typhimurium. Purification and characterization, *J Biol Chem 250*, 609-616.

Belting, M., Borsig, L., Fuster, M. M., Brown, J. R., Persson, L., Fransson, L. A., and Esko, J. D. (2002) Tumor attenuation by combined heparan sulfate and polyamine depletion, *Proc Natl Acad Sci U S A 99*, 371-376.

Brown, P. H., and Schuck, P. (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation, *Biophys J 90*, 4651-4661.

Campbell, R. E., Mosimann, S. C., van De Rijn, I., Tanner, M. E., and Strynadka, N. C. (2000) The first structure of UDP-glucose dehydrogenase reveals the catalytic residues necessary for the two-fold oxidation, *Biochemistry 39*, 7012-7023.

Cecchin, E., Innocenti, F., D'Andrea, M., Corona, G., De Mattia, E., Biason, P., Buonadonna, A., and Toffoli, G. (2009) Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan, *J Clin Oncol 27*, 2457-2465.

Cummings, J., Boyd, G., Ethell, B. T., Macpherson, J. S., Burchell, B., Smyth, J. F., and Jodrell, D. I. (2002) Enhanced clearance of topoisomerase I inhibitors from human colon cancer cells by glucuronidation, *Biochem Pharmacol 63*, 607-613.

Cummings, J., Ethell, B. T., Jardine, L., Boyd, G., Macpherson, J. S., Burchell, B., Smyth, J. F., and Jodrell, D. I. (2003) Glucuronidation as a mechanism of intrinsic drug resistance in human colon cancer: reversal of resistance by food additives, *Cancer Res 63*, 8443-8450.

Cummings, J., Zelcer, N., Allen, J. D., Yao, D., Boyd, G., Maliepaard, M., Friedberg, T. H., Smyth, J. F., and Jodrell, D. I. (2004) Glucuronidation as a mechanism of intrinsic drug resistance in colon cancer cells: contribution of drug transport proteins, *Biochem Pharmacol 67*, 31-39.

Cummings, J., Ethell, B. T., Jardine, L., and Burchell, B. (2006) Glucuronidation of SN-38 and NU/ICRF 505 in human colon cancer and adjacent normal colon, *Anticancer Res 26*, 2189-2196.

Dickerson, R. E., Goodsell, D. S., and Neidle, S. (1994) "...the tyranny of the lattice...", *Proc Natl Acad Sci U S A 91*, 3579-3583.

Dickinson, F. M. (1988) Studies on the unusual behaviour of bovine liver UDP-glucose dehydrogenase in assays at acid and neutral pH and on the presence of tightly bound nucleotide material in purified preparations of this enzyme, *Biochem J 255*, 775-780.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, *Nucleic Acids Res 34*, W116-118.

Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.

Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect, *Science 255*, 178-183.

Frieden, C. (1979) Slow transitions and hysteretic behavior in enzymes, *Annu Rev Biochem 48*, 471-489.

Fuster, M. M., and Esko, J. D. (2005) The sweet and sour of cancer: glycans as novel therapeutic targets, *Nat Rev Cancer 5*, 526-542.

Gainey, P. A., and Phelps, C. F. (1975) Interactions of urdine diphosphate glucose dehydrogenase with the inhibitor urdine diphosphate xylose, *Biochem J 145*, 129-134.

Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins?, *Proteins 57*, 433-443.

Hayward, S., and Berendsen, H. J. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme, *Proteins 30*, 144-154.

Helt, S. S., Thymark, M., Harris, P., Aagaard, C., Dietrich, J., Larsen, S., and Willemoes, M. (2008) Mechanism of dTTP inhibition of the bifunctional dCTP deaminase:dUTPase encoded by Mycobacterium tuberculosis, *J Mol Biol 376*, 554-569.

Huh, J. W., Yoon, H. Y., Lee, H. J., Choi, W. B., Yang, S. J., and Cho, S. W. (2004) Importance of Gly-13 for the coenzyme binding of human UDP-glucose dehydrogenase, *J Biol Chem 279*, 37491-37498.

Itano, N., and Kimata, K. (2008) Altered hyaluronan biosynthesis in cancer progression, *Semin Cancer Biol 18*, 268-274.

Joh, N. H., Oberai, A., Yang, D., Whitelegge, J. P., and Bowie, J. U. (2009) Similar energetic contributions of packing in the core of membrane and water-soluble proteins, *J Am Chem Soc 131*, 10846-10847.

Johansson, E., Bjornberg, O., Nyman, P. O., and Larsen, S. (2003) Structure of the bifunctional dCTP deaminase-dUTPase from Methanocaldococcus jannaschii and its relation to other homotrimeric dUTPases, *J Biol Chem 278*, 27916-27922.

Johansson, E., Thymark, M., Bynck, J. H., Fano, M., Larsen, S., and Willemoes, M. (2007) Regulation of dCTP deaminase from Escherichia coli by nonallosteric dTTP binding to an inactive form of the enzyme, *Febs J 274*, 4188-4198.

Kabsch, W. (2010) Xds, *Acta Crystallogr D Biol Crystallogr 66*, 125-132.

Kolodziej, A. F., Tan, T., and Koshland, D. E., Jr. (1996) Producing positive, negative, and no cooperativity by mutations at a single residue located at the subunit interface in the aspartate receptor of Salmonella typhimurium, *Biochemistry 35*, 14782-14792.

Kumar, M. D., and Gromiha, M. M. (2006) PINT: Protein-protein Interactions Thermodynamic Database, *Nucleic Acids Res 34*, D195-198.

Kuo, L. C., Zambidis, I., and Caron, C. (1989) Triggering of allostery in an enzyme by a point mutation: ornithine transcarbamoylase, *Science 245*, 522-524.

Lacour, S., Bechet, E., Cozzone, A. J., Mijakovic, I., and Grangeasse, C. (2008) Tyrosine phosphorylation of the UDP-glucose dehydrogenase of Escherichia coli is at the crossroads of colanic acid synthesis and polymyxin resistance, *PLoS One 3*, e3053.

Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins, In *Analytical ultracentrifugation in biochemistry and polymer science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), pp 90-125, The Royal Society of Chemistry, Cambridge.

Liao, H., Yeh, W., Chiang, D., Jernigan, R. L., and Lustig, B. (2005) Protein sequence entropy is closely related to packing density and hydrophobicity, *Protein Eng Des Sel 18*, 59-64.

Ma, B., Kumar, S., Tsai, C. J., and Nussinov, R. (1999) Folding funnels and binding mechanisms, *Protein Eng 12*, 713-720.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.

Monod, J., Changeux, J. P., and Jacob, F. (1963) Allosteric proteins and cellular control systems, *J Mol Biol 6*, 306-329.

Neufeld, E. F., and Hall, C. W. (1965) Inhibition of Udp-D-Glucose Dehydrogenase by Udp-D-Xylose: A Possible Regulatory Mechanism, *Biochem Biophys Res Commun 19*, 456-461.

Oguri, T., Takahashi, T., Miyazaki, M., Isobe, T., Kohno, N., Mackenzie, P. I., and Fujiwara, Y. (2004) UGT1A10 is responsible for SN-38 glucuronidation and its expression in human lung cancers, *Anticancer Res 24*, 2893-2896.

Richards, F. M. (1977) Areas, volumes, packing and protein structure, *Annu Rev Biophys Bioeng 6*, 151-176.

Sanderson, R. D., Yang, Y., Kelly, T., MacLeod, V., Dai, Y., and Theus, A. (2005) Enzymatic remodeling of heparan sulfate proteoglycans within the tumor microenvironment: growth regulation and the prospect of new cancer therapies, *J Cell Biochem 96*, 897-905.

Scrutton, N. S., Deonarain, M. P., Berry, A., and Perham, R. N. (1992) Cooperativity induced by a single mutation at the subunit interface of a dimeric enzyme: glutathione reductase, *Science 258*, 1140-1143.

Sommer, B. J., Barycki, J. J., and Simpson, M. A. (2004) Characterization of human UDP-glucose dehydrogenase. CYS-276 is required for the second of two successive oxidations, *J Biol Chem 279*, 23590-23596.

Stern, R. (2008) Association between cancer and "acid mucopolysaccharides": an old concept comes of age, finally, *Semin Cancer Biol 18*, 238-243.

Takahashi, T., Fujiwara, Y., Yamakido, M., Katoh, O., Watanabe, H., and Mackenzie, P. I. (1997) The role of glucuronidation in 7-ethyl-10-hydroxycamptothecin resistance in vitro, *Jpn J Cancer Res 88*, 1211-1217.

Tokuriki, N., and Tawfik, D. S. (2009) Protein dynamism and evolvability, *Science 324*, 203-207.

Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use, *Acta Crystallogr D Biol Crystallogr 60*, 2184-2195.

van den Berg, S., Lofdahl, P. A., Hard, T., and Berglund, H. (2006) Improved solubility of TEV protease by directed evolution, *J Biotechnol 121*, 291-298.

Zhang, X. J., Wozniak, J. A., and Matthews, B. W. (1995) Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme, *J Mol Biol 250*, 527-552.