

Bifunctional Substrate Activation via an Arginine Residue Drives Catalysis in Chalcone Isomerases

Jason R. Burke,[†] James J. La Clair,[†] Ryan N. Philippe,^{†,#} Anna Pabis,[‡] Joseph M. Jez,^{†,#} George A. Cortina,[§] Miriam Kaltenbach,^{||} Marianne E. Bowman,[†] Gordon V. Louie,[†] Katherine B. Woods,[†] Andrew T. Nelson,^{||} Dan S. Tawfik,^{||} Shina C.L. Kamerlin,[§] and Joseph P. Noel^{†,*}

[†]Howard Hughes Medical Institute, Jack H. Skirball Center for Chemical Biology and Proteomics, The Salk Institute for Biological Studies, La Jolla, CA 92037, United States

[#]Department of Chemistry - BMC, Uppsala University, BMC Box 576, S-751 23 Uppsala, Sweden

[§]Department of Molecular Physiology and Biomedical Engineering, University of Virginia, Charlottesville, VA, United States

^{||}Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot 76100, Israel

^{††}Department of Chemistry, University of Texas at Austin, Austin, TX 78712, United States

ABSTRACT: Chalcone isomerases are plant enzymes that perform enantioselective oxa-Michael cyclizations of 2'-hydroxy-chalcones into flavanones. An X-ray crystal structure of an enzyme-product complex and molecular dynamics simulations reveal an enzyme mechanism wherein the guanidinium ion of a conserved arginine positions the nucleophilic phenoxide and activates the electrophilic enone for cyclization through Brønsted and Lewis acid interactions. The reaction terminates by asymmetric protonation of the carbanion intermediate *syn* to the guanidinium. Interestingly, bifunctional guanidine- and urea-based chemical reagents, increasingly used for asymmetric organocatalytic applications, are synthetic counterparts to this natural system. Comparative protein crystal structures and molecular dynamics simulations further demonstrate how two active site water molecules coordinate a hydrogen bond network that enables expanded substrate reactivity for 6'-deoxychalcones in more recently evolved type-2 chalcone isomerases.

INTRODUCTION

During plant transitions from aquatic to terrestrial ecosystems approximately 450 MYA, critical chemo-adaptations emerged to attenuate exposure to UV radiation. In part, the biosynthesis of polyphenol-based sunscreens likely provided protection by efficiently absorbing UV light and mitigating associated damage from oxidative stress.¹ Over time, plants evolved diverse biological roles for flavonoids in guarding against pathogenesis and herbivory, facilitating phytohormone transport and microbial symbiosis, and modulating floral pigmentation.²

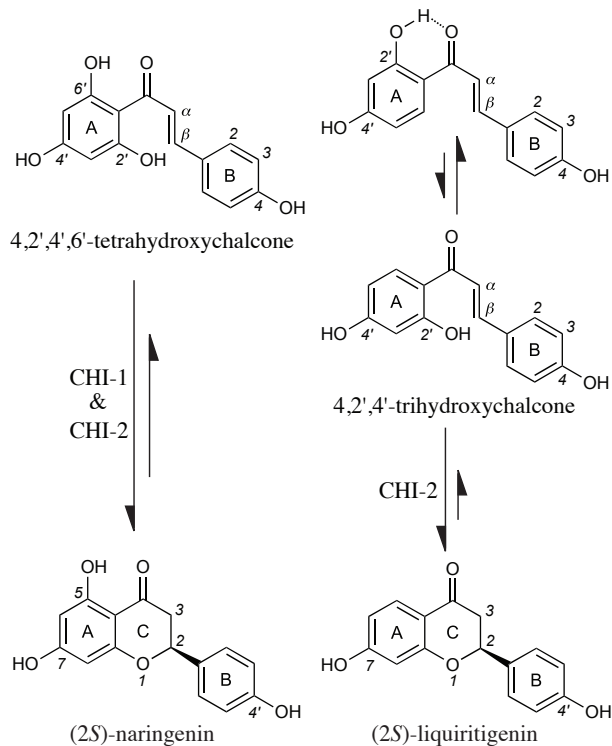
In plants, polycyclic flavanones arise via enzymatic intramolecular Michael reactions. While Michael-type, natural catalysts have been inferred that efficiently construct C-C and C-O bonds, highly enantioselective examples of these enzymes remain rare.³ Therefore, a better understanding of their catalytic principles is needed to guide efforts to engineer "Michaelase" biocatalysts.⁴ On the other hand, the Michael reaction is common in organic syntheses where it offers one of the most versatile and elegant means to combine fragments and install complex molecular frameworks.⁵ Recent advances in organocatalytic design have honed the use of guanidine- and urea-based templates for bifunctional asymmetric catalysts, which deploy hydrogen bonds and Lewis acids to activate and

position electrophilic and nucleophilic moieties.⁶ Here we show that the plant enzyme, chalcone isomerase (CHI), employs such a strategy to direct the enantioselective cyclization of flavanones, bridging organocatalytic rational design with biocatalytic evolutionary insights.

Two enzymatically distinct groups of CHIs guide the C-O bond forming oxa-Michael cyclization of flavanones in plants.⁷ Type-1 CHIs (CHI-1) are ubiquitous across the green plant lineage and catalyze the cyclization of 6'-hydroxychalcones, such as 4,2',4',6'-tetrahydroxy-chalcone, to form chiral flavanones such as (2*S*)-naringenin (Scheme 1). Type-2 CHIs (CHI-2) are more restricted throughout the plant kingdom and are most commonly found in Fabaceae (the Legume family). CHI-2s possess comparable activity to CHI-1s for 6'-hydroxychalcones; however, CHI-2s also possess high activity for the more narrowly distributed 6'-deoxychalcones. The 6'-deoxychalcones, such as 4,2',4'-trihydroxychalcone, contain a high energy intramolecular C=O--H-O bond, which prevents appreciable rates of auto-cyclization observed for 6'-hydroxychalcones.⁸

These conformationally restrictive hydrogen bonds in 6'-deoxychalcones require CHI-2 catalytic activity for facile disruption and ensuing cyclization reactions.^{7,8} Here, we identify specific amino acid residues in CHI-2s that order

Scheme 1. CHI-2s catalyze the cyclization of 6'-hydroxychalcones and 6'-deoxychalcones into (2S)-flavanones; CHI-1s are selective for 6'-hydroxychalcones.



active site waters to facilitate substrate reorganizations, enabling kinetically efficient ring closures of 6'-deoxychalcones into reduced flavanones such as (2S)-liquiritigenin.

All CHI enzymes evolved from a larger family of non-enzymatic, Fatty Acid Binding Proteins (FAPs), present throughout green plants, fungi, and protists.⁹ The emergence and evolution of enzyme activity in the CHI-fold was enabled, in part, by a founder mutation that expands the dynamism and conformational space of an active site arginine.¹⁰ In non-catalytic FAPs, the same arginine plays a structural role in fatty-acid binding.⁹ Although structures of CHIs have revealed an array of thermodynamically favored conformations for the active site arginine, evidence of its catalytic role has remained elusive.⁹⁻¹²

Protein X-ray crystal structures in conjunction with molecular dynamics simulations have further shaped our mechanistic understanding of the intramolecular oxa-Michael reactions in both CHI-1s and CHI-2s.^{11,13} From a protein crystal structure of product-bound CHI-1, we here delineate the key catalytic roles of the conserved active site arginine in both aligning a nucleophilic Michael donor (2'O-) and enhancing the electrophilicity of the enone Michael acceptor (C β) through both Brønsted and Lewis acid interactions. Consistent with these observed structured interactions, the data presented support a model in which the reaction terminates upon the asymmetric transfer of a proton from the guanidinium moiety of the arginine to the substrate carbanion tautomer at C α . The mechanistic details presented herein correlate features of this natural biocatalytic reaction with diverse, guanidine-based organocatalysts, developed for varied applications in asymmetric syntheses.^{6,14}

RESULTS AND DISCUSSION

X-ray crystal structure and molecular dynamics simulations of MtCHI-1 reveal the bifunctional catalytic roles of the active site arginine. To reveal the chemistry of the active site arginine and pinpoint the stereochemical differences that define substrate selectivity for CHI-1 versus CHI-2, we crystallized CHI-1 from *Medicago truncatula* (MtCHI-1), a close relative of the well-characterized CHI-2 from *Medicago sativa* (MsCHI-2).⁹ Protein crystals were transferred to a solution containing 4,2',4',6'-tetrahydroxychalcone, and, over time, the crystalline proteins catalyzed the conversion of the substrate to (2S)-naringenin (Figure S1). The product-bound co-crystal structure was solved with eight α/β -fold protein molecules per asymmetric unit, and electron density supporting (2S)-naringenin bound in five of the eight MtCHI-1 active sites (Figures 1a, S2, and Table S1).

Crystallographic details of the MtCHI-1 • (2S)-naringenin complex reveal multifunctional catalytic roles for the active site arginine (Arg37). Previous structures of MsCHI-2 • (2S)-naringenin show the catalytic arginine (Arg36) oriented away from the active site through a salt bridge interaction with Asp200 (Figure S3).⁹ However, mutation of Asp200 to Ala or Asn does not perturb enzyme activity, suggesting the crystallographic orientation of Arg36 in MsCHI-2 is non-productive (Table 1). Here, the MtCHI-1 • (2S)-naringenin complex reveals that the carbocation center (C ζ) of Arg37 resides within van der Waals radial distance of the product's pyran oxygen (10) supporting a role for Arg37 as a Lewis acid in stabilizing the nucleophilic 2'-oxoanion of the chalcone substrate (Figure 1b, Table S2). The guanidine of Arg37 forms a shared H η_2 -10 hydrogen bond with (2S)-naringenin, positing a mechanism in which Arg37 orients the substrate's 2'-oxyanion for ring closure. Furthermore, the guanidine of Arg37 is positioned to act asymmetrically as a Brønsted acid through a shared H ϵ bond with C3 of (2S)-naringenin (C α in 4,2',4',6'-tetrahydroxychalcone). This Brønsted acid interaction may increase the electrophilicity of the olefin at C β further favoring the approach to a catalytically productive substrate conformation.

Molecular dynamics of MtCHI-1 in complex with 4,2',4',6'-tetrahydroxychalcone indicate a clear preference for substrate conformations approaching cyclization. In these simulations, the nucleophilic Michael donor is stabilized by the guanidinium ion of Arg37 primarily through simultaneous interaction with the N η_2 and N ϵ as well as the carbocationic center of the ion (C ζ) (Figure 1c and Table S3). At the same time, the mutual orientation of Arg37 and the substrate in the simulated Michaelis complex, and the proximity of the N ϵ group to the C α atom supports the posited role of the guanidinium ion acting as a Brønsted acid in the enantioselective cyclization of 4,2',4',6'-tetrahydroxychalcone. Together with the structures, these results highlight how the electrostatics and hydrogen bond geometry of arginine's natural guanidine can simultaneously stabilize, organize and activate multiple moieties involved in conjugate addition reactions.

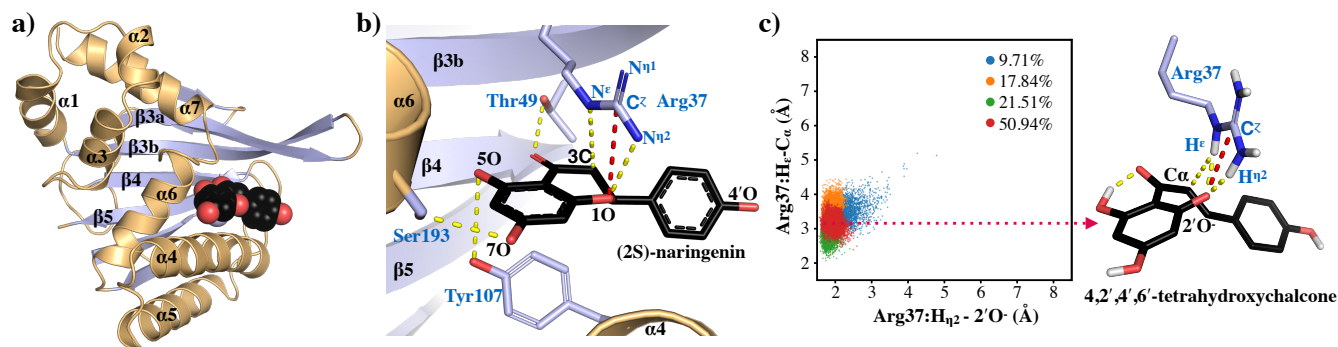


Figure 1. Structural role of the arginine in CHI catalysis. (a) X-ray crystal structure of *MtCHI-1* bound with (2*S*)-naringenin. (b) Enzyme-product interactions suggest how the guanidinium of Arg37 directs ring closure by positioning and activating the substrate through hydrogen bonds (yellow dotted lines) and Lewis acid interactions (red dotted line). (c) Conformational space defined by six Arg37-substrate contacts ($H_{\eta^2}-2'O$, $H_{\epsilon}-C_{\alpha}$, $N_{\eta^1}-2'O$, $N_{\eta^2}-2'O$, $H_{\epsilon}-2'O$, $C_{\zeta}-2'O$) and reactive $2'O-C_{\beta}$ distance observed in the simulations of *MtCHI-1* bound with **1** support the existence of enzyme-substrate interactions favorable for catalysis.

Intriguingly, these structural details correlate features of this biocatalytic reaction with asymmetric organocatalysts developed for the Strecker reaction,¹⁵ Michael reactions,¹⁶ and Claisen rearrangements,¹⁷ among others.¹⁴ For example, Lewis acid and Brønsted acid interactions, mediated through the guanidinium moiety of a chiral, triazabicyclo[4.4.0]dec-5-ene (TBD)-based catalyst, direct a tandem thiol Michael addition toward a particular stereochemical outcome.^{16a} Further, an X-ray crystal structure of the TBD catalyst clearly demonstrates the capacity of the guanidinium to simultaneously position and activate both a nucleophile and electrophile through Lewis acid and Brønsted acid interactions for conjugate addition.¹⁸ Together, the mechanistic similarity of these examples of guanidine based organocatalysts highlights the potential for chalcone isomerases to serve as proteinaceous catalysts in diverse enantioselective chemical applications, as direct, natural complements for synthetic catalyst design.

Recent studies have identified alternatives to guanidinium-based asymmetric catalysis of flavanones from chalcones. Chiral thioureas derived from quinine/quinidine scaffolds have been adapted as synthetic mimetics of CHI enantioselective functionality.¹⁹ The human gut bacterium, *Eubacterium ramulus*, which evolved CHI activity from an unrelated protein fold, reverses the reaction for flavonoid catabolism.²⁰ The bacterial CHI is similar to plant CHIs in its selectivity for (2*S*)-flavanones; however, for the bacterial CHI an active site histidine facilitates general acid/base catalysis.

The active site arginine is essential for efficient catalysis in plant CHIs. For *MsCHI-2*, mutations of Arg36 (R36A, R36H, R36M and R36Q) result in approximately 10^4 to 10^5 -fold reductions in catalytic turnover (k_{cat} min⁻¹) of 4,2',4',6'-tetrahydroxychalcone and 4,2',4'-trihydroxychalcone, respectively (Table 1). For both *MsCHI-2* and *MtCHI-1*, substitutions of active site arginines with lysines also dramatically reduce catalytic turnovers: approximately 10^3 -fold for 4,2',4',6'-tetrahydroxychalcone, and $>10^3$ -fold for 4,2',4'-trihydroxychalcone with *MsCHI-2* (Table 1). For *MsCHI-2* and *MtCHI-1*, these results support a role for the Arginine residue in facilitating CHI catalysis by acting as more than a simple point charge, but instead carrying out delocalized, multidentate functions.

Table 1. Michaelis-Menten kinetics of CHI mutations

	4,2',4',6'-tetrahydroxychalcone		
	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (s ⁻¹ M ⁻¹)
<i>MsCHI-2</i> - Wild-type	11,180 ± 1,380	112 ± 28	1,660,000
<i>MsCHI-2</i> - R36A ^a	2.1	<200	175
<i>MsCHI-2</i> - R36H	4.8 ± 0.2	135 ± 23	593
<i>MsCHI-2</i> - R36K	25.4 ± 1.1	98.3 ± 1.5	4,310
<i>MsCHI-2</i> - R36M ^a	2.5	<200	208
<i>MsCHI-2</i> - R36Q ^a	2.4	<200	200
<i>MsCHI-2</i> - G95S	530 ± 47	30.3 ± 8.5	292,000
<i>MsCHI-2</i> - G95T	528 ± 47	169 ± 55	52,000
<i>MsCHI-2</i> - G95V	512 ± 26	5.3 ± 2.6	1,720,000
<i>MsCHI-2</i> - K97A	2,450 ± 272	16.9 ± 3.7	2,400,000
<i>MsCHI-2</i> - K97E	1700 ± 139	31.2 ± 8.6	884,000
<i>MsCHI-2</i> - K97M	1,300 ± 154	14.5 ± 4.2	1,440,000
<i>MsCHI-2</i> - K97Q	375 ± 42	11.8 ± 2.8	514,000
<i>MsCHI-2</i> - K109M	9,610 ± 956	261 ± 34	614,000
<i>MsCHI-2</i> - D200A	12,100 ± 1,160	192 ± 41	1,050,000
<i>MsCHI-2</i> - D200N	12,300 ± 1,420	166 ± 30	1,240,000
<i>MtCHI-1</i> - Wild-type	1,683 ± 93	13 ± 2	2,050,000
<i>MtCHI-1</i> - R37K	1.16 ± 0.04	2.5 ± 0.3	7,810
<i>MtCHI-1</i> - T49A	1.72 ± 0.32	13.2 ± 5.4	2,170
<i>MtCHI-1</i> - Y107F	426 ± 120	10.8 ± 7	659,000
<i>MtCHI-1</i> - K110M	1,691 ± 409	224 ± 66.5	126,000
<i>MtCHI-1</i> - S195A	50.0 ± 7.3	18.9 ± 6.6	44,900
<i>AtCHI-1</i> - Wild-type	379 ± 35	62 ± 9	103,000
<i>AtCHI-1</i> - V106G	332 ± 34	19.0 ± 0.5	292,000
	4,2',4'-trihydroxychalcone		
	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (s ⁻¹ M ⁻¹)
<i>MsCHI-2</i> - Wild-type	2,250 ± 165	8.4 ± 2.0	4,460,000
<i>MsCHI-2</i> - R36A ^a	0.040	<50	13
<i>MsCHI-2</i> - R36H	0.16 ± 0.01	9.5 ± 1.5	280
<i>MsCHI-2</i> - R36K	0.61 ± 0.03	15.2 ± 2.1	670
<i>MsCHI-2</i> - R36M ^a	0.038	<50	9
<i>MsCHI-2</i> - R36Q ^a	0.024	<50	7
<i>MsCHI-2</i> - G95S	59.5 ± 5	6.5 ± 1.1	153,000
<i>MsCHI-2</i> - G95T	23.3 ± 1.7	117 ± 12	3310
<i>MsCHI-2</i> - G95V	2.16 ± 0.02	65 ± 10	556
<i>MsCHI-2</i> - K97A	210 ± 10	36.5 ± 3.7	100,000
<i>MsCHI-2</i> - K97E	98 ± 8	78.6 ± 10.4	25,000
<i>MsCHI-2</i> - K97M	658 ± 11	25.7 ± 2.1	427,000
<i>MsCHI-2</i> - K97Q	146 ± 12	65.1 ± 9.5	31,000
<i>MsCHI-2</i> - K109M	1,977 ± 183	16.3 ± 1.9	2,021,000
<i>MsCHI-2</i> - D200A	2,560 ± 176	14.2 ± 2.6	3,010,000
<i>MsCHI-2</i> - D200N	2,730 ± 161	12.6 ± 2.3	3,610,000
<i>MtCHI-1</i> - Wild-type	0.15 ± 0.13	44 ± 48	56
<i>AtCHI-1</i> - Wild-type	0.064 ± 0.005	16.8 ± 3.5	42
<i>AtCHI-1</i> - V106G	0.16 ± 0.02	31.9 ± 8.5	603

^a Where saturation was not observed, k_{cat}/K_M values were estimated using $k_{cat}/[S]$

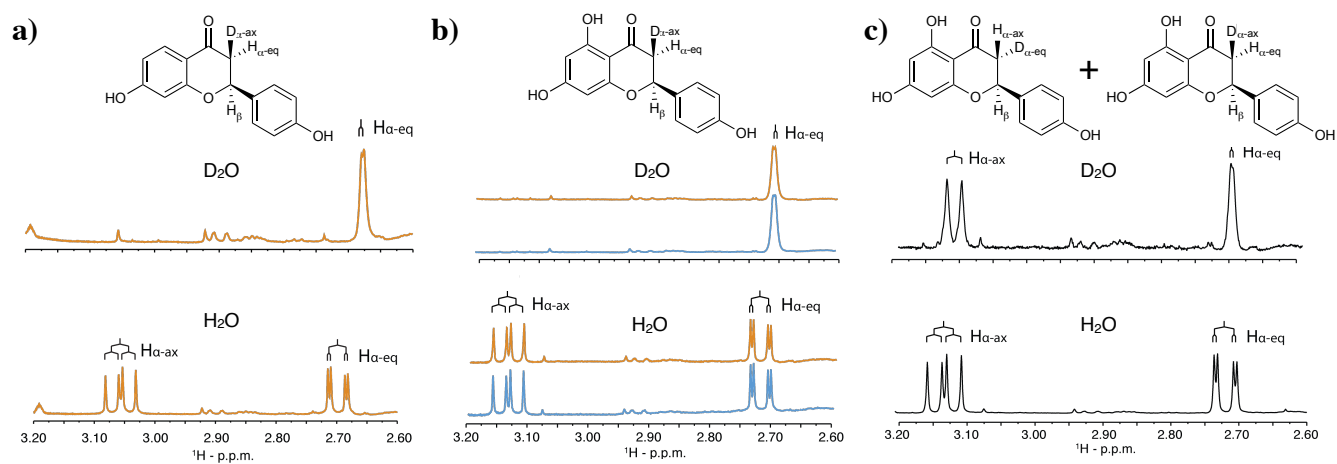


Figure 2. NMR analyses of flavanone protonation stereochemistry. (a) An expansion of a ¹H-NMR spectra of axial (H α -ax) and equatorial (H α -eq) protons at C3 of (2*S*)-liquiritigenin, generated from the *Ms*CHI-2-catalyzed cyclization of 4,2',4'-trihydroxychalcone in the presence of D₂O and H₂O. (b) Spectra of (2*S*)-naringenin formed by *Ms*CHI-2 (orange) and *Mt*CHI-1 (blue). (c) Cyclization of (2*R/2S*)-naringenin in D₂O or H₂O in the absence of CHI. When the reaction is run in D₂O, both the H α -ax and H α -eq protons are doublets of roughly equal intensity, revealing equivalent protonation to re and si faces

Stereochemistry of the proton transfer step. The orientation and proximity of the guanidinium N ϵ -H ϵ bond to the carbon 3C of (2*S*)-naringenin suggest that Arg37 delivers H ϵ to the carbanion tautomer of the substrate's enolate intermediate (Figure 1b). Using fractionated lysates from the Fabaceae, *Vigna radiata* (Mung Bean), an early investigation reported CHI-mediated protonation of C α consistent with the formation of the H-C α bond *syn* to the O-C β bond of 4,2',4'-trihydroxychalcone.²¹ However, ambiguities were noted in the interpretation of these NMR spectra due to unresolved heterogeneities. In light of the catalytic role and positioning of Arg37 in the structure of *Mt*CHI-1, we evaluated the proton transfer stereochemistry and expanded the scope of the study to include homogeneously purified *Mt*CHI-1, *Ms*CHI-2, and the ubiquitous substrate, 4,2',4',6'- tetrahydroxychalcone.

Flavanones formed by *Mt*CHI-1 and *Ms*CHI-2 in the presence of water exhibit two doublets of doublets by proton NMR, characteristic of equatorial and axial protons at the C α atoms in these products (Figures 2a-b, S4). When the enzyme reactions are run in D₂O, the flavanone proton peaks corresponding to H α -ax are absent and the H α -eq peaks occur as doublets with small *J*-coupling constants indicative of *gauche* orientations relative to H β -ax protons. These data unequivocally demonstrate that H α -ax derives from enzyme-mediated protonation of the *si* (*pro-S*) face of the carbanion tautomer intermediate for both CHI-1 and CHI-2. Non-enzymatic cyclization of 4,2',4',6'-tetrahydroxychalcone into naringenin in D₂O results in deuterium incorporation at both the H α -ax and H α -eq positions equivalently (Figure 2c). Together, these results are consistent with a structural model in which the H ϵ proton of the guanidinium is transferred to the carbanion tautomer of the enolate intermediate during enzymatic cyclization of (2*S*)-flavanones.

While arginine is often considered an unlikely candidate to function as a general acid/base in enzymatic reactions because of the apparent high pK_a of its guanidinium group (pK_a 13.8),²² exceptions have been described for citrate synthase,²³ serine recombinase,²⁴ IMP dehydrogenase,²⁵ endoribonuclease,²⁶ and fumarate reductase²⁷. Current hypotheses that posit the role of arginines as general acids/bases in enzyme catalysis depend upon: (1) pK_a values of substrates, (2) geometric constraints of the guanidinium cations in the postulated enzyme transition states,²⁸ (3) burial of side chain charges in hydrophobic environments,²⁹ and, in the case of fumarate reductase, (4) proton shuttling via a Grotthuss-like mechanism.^{27c,30} It is probable that CHI enzymes share mechanistic features with these studied systems. For example, similar to Grotthuss-like mechanisms, and mechanisms in which proximal charges affect pK_a values, a buried, active site Lys97-Zundel cation feature in CHI-2 may serve to reduce the effective pK_a of the catalytic arginine, priming it as a Brønsted base to deprotonate the substrate 2'OH.¹¹ Strong kinetic isotope effects for CHI-2, which are dependent upon the water-coordinating side chains of Thr48 and Tyr106, suggest proton transfer along the Lys97-linked water chain is rate limiting.^{11c}

Active site architectures support the stereochemistry of ring closure. To determine whether the catalytic arginine modulates the stereochemical outcome of CHI-mediated reactions, we resolved the flavanone enantiomers produced by *Ms*CHI-2 arginine mutants (R36A, R36H, R36M, R36Q and R36K). Using chiral HPLC analyses, we found that (2*S*)-flavanones are enzymatically produced exclusively across the panel of mutant enzymes lacking the catalytic Arg36, suggesting that additional factors associated with the shape and chemistry of the CHI active site direct the enantioselective reactions catalyzed by CHIs (Figure S5).

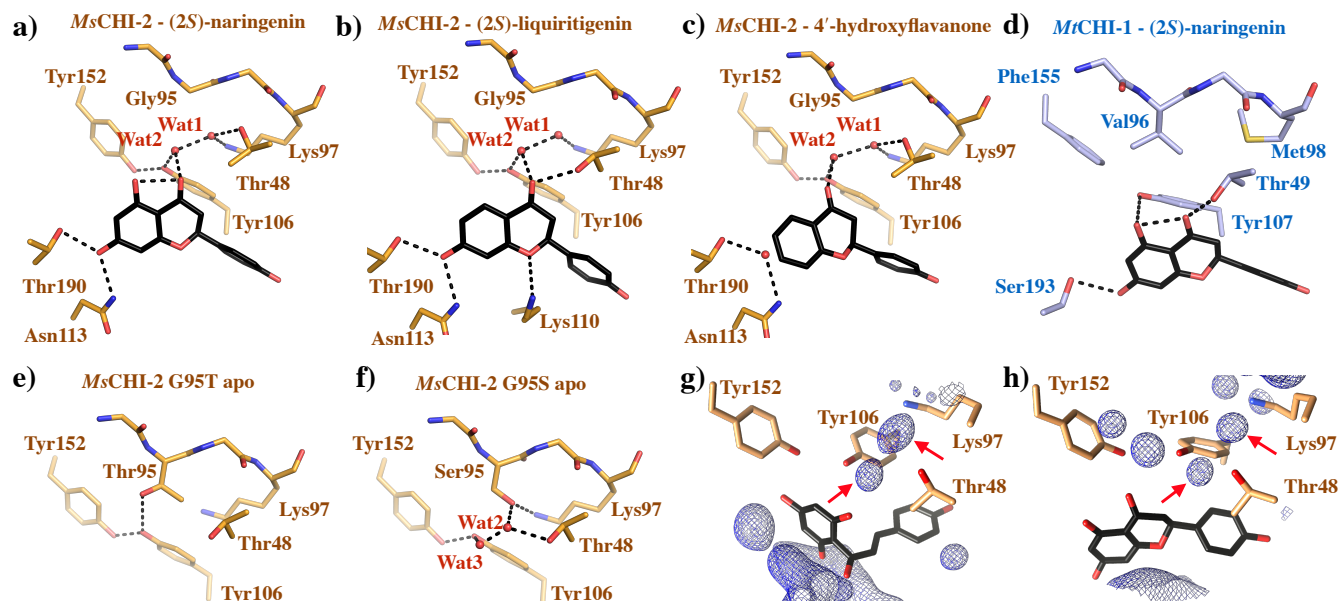


Figure 3. Structural rationale for CHI-1/CHI-2 functional divergence. (a-c) Water-mediated hydrogen bonds (dashed lines) between *MsCHI-2* Lys97, Tyr106 and Tyr152 and (2*S*)-naringenin (rendered from PDB:1EYQ, PDB:1FM7, PDB:1JEP). (d) Ordered water molecules are excluded from the active site of *MtCHI-1* • (2*S*)-naringenin due to steric occlusion by the side chain of Val96. (e) In the apo-crystal structure of the *MsCHI-2* mutant, G95T, Thr95 excludes ordered water and disrupts hydrogen bonds with the side chain of Lys97. (f) In the apo-crystal structure of the *MsCHI-2* G95S mutant, the hydroxyl of Ser95 maintains the hydrogen bonding network with Lys97 and water 2. Water 3 is positioned similar to the carbonyl oxygen of (2*S*)-flavanones. Water density seen in MD simulations of *MsCHI-2* with 4,2',4',6'-tetrahydroxychalcone (g) and (2*S*)-naringenin (h). The two indicated positions visited by water molecules during the simulations are similar to the positions of the water molecules observed in the crystal structures of *MsCHI-2* in complex with (2*S*)-flavanones products (red arrows).

Conserved waters in CHI-2s trigger the turnover of 6'-deoxychalcones. CHI-2s evolved to catalyze the cyclization of 6'-deoxychalcones, and are therefore enzymatically distinct from CHI-1s. The biochemical bases for these phenotypic differences have remained enigmatic.⁷ In the X-ray crystal structures of flavanone-bound *MsCHI-2* complexes, two ordered water molecules are present and well-coordinated near the back of the *MsCHI-2* active sites (Figure 3a-c).^{11a,b} Notably, in the *MtCHI-1* • (2*S*)-naringenin bound complexes, these two ordered water molecules are absent (Figure 3d). In *MsCHI-2*, Gly95 accommodates the hydrogen bonding environment necessary for the sequestration of the two water molecules. In contrast, in the active site of *MtCHI-1*, Val96 (equivalent to Gly95 in *MsCHI-2*) restricts binding of the ordered waters. The G95V mutation in *MsCHI-2* transforms *MsCHI-2* into a “type-1-like” CHI, evidenced by impairment of 4,2',4'-trihydroxychalcone turnover to (2*S*)-liquiritigenin (Table 1). The reciprocal mutation in *AtCHI-1*, V106G, results in a 10-fold gain in catalytic efficiency for the conversion of 4,2',4'-trihydroxychalcone to (2*S*)-liquiritigenin (Table 1). These results support a mechanistic role for CHI-2-specific active site water molecules in the enantioselective cyclization of 6'-deoxychalcones.

In *MsCHI-2*, the ordered active site water molecules are part of a hydrogen bonding network mediated by amino acid residues conserved in- and unique to-CHI-2s: Gly95, Lys97 and Tyr152 (Figure S6). This lattice of water-mediated hydrogen bonds links these residues to the carbonyl oxygen of the flavanone product, (2*S*)-naringenin (Figure 3a-c). The architecture of this extended set of hydrogen

bonds supports a role for the ordered waters in disruption of the intramolecular C=O--H-O hydrogen bonds, affording the facile A-ring flips of 6'-deoxychalcones necessary to reposition and activate the 2'-phenoxides for the ensuing enantioselective oxa-Michael reactions (Scheme 1).^{8b} Disruption of this hydrogen-bonding network is observed in the apo-crystal structure of the *MsCHI-2* G95T mutant (Table S1). Specifically, the methyl group of the Thr95 side chain supplants a water molecule, interrupting the hydrogen bonding interactions with Lys97 (Figure. 3e, S7). The *MsCHI-2* G95T mutant reduces k_{cat} ~10²-fold and increases K_M ~14-fold for 4,2',4'-trihydroxychalcone, yet for 4,2',4',6'-tetrahydroxychalcone, k_{cat} is reduced only 20-fold with no significant change to K_M (Table 1). By comparison, the crystal structure of apo-*MsCHI-2* G95S (a rare mutation present in at least one CHI-2, see Figure S6), reveals the hydroxyl of Ser95 positioned isosteric to one of the water molecules observed in wild-type *MsCHI-2* (Figure 3f, S7 and Table S1). As seen in the structure, this conserved chemical feature preserves the hydrogen bonding pattern of the natural enzyme. Accordingly, K_M for *MsCHI-2* G95S is similar to wild-type *MsCHI-2* for both 4,2',4'-trihydroxychalcone and 4,2',4',6'-tetrahydroxychalcone (decreased approximately 3-fold), and k_{cat} is similarly reduced for both substrates: 40-fold for 4,2',4'-trihydroxychalcone and 20-fold for 4,2',4',6'-tetrahydroxychalcone (Table 1).

Mutations of Lys97 in *MsCHI-2* (K97A, K97E, K97M, and K97Q), exhibit 3- to 10-fold decreases in K_M for 4,2',4',6'-tetrahydroxychalcone with 4- to 8-fold increases in K_M for

4,2',4'-trihydroxychalcone (Table 1). These results support the role of Lys97 in enhancing the binding specificity for the narrowly distributed 6'-deoxychalcones over the much more ubiquitous 6'-hydroxychalcone substrates in plants. Interestingly, k_{cat} values are similarly reduced for both substrates across Lys97 mutants. The positive effect of Lys97 on catalysis of the ubiquitous substrate, 4,2',4',6'-tetrahydroxychalcone, suggests that Lys97 and the associated active site water molecules distinguish the mechanisms of catalytic turnover of 6'-hydroxychalcones by CHI-1s versus CHI-2s.

Targeted mutations further suggest mechanistic distinctions between CHI-1 and CHI-2 enzymes. The losses of phenolic hydroxyl moieties by Y107F and Y106F mutations in *Mt*CHI-1 and *Ms*CHI-2 reduce k_{cat} values by 30-fold for both substrates in *Ms*CHI-2, yet only 4-fold for 4,2',4',6'-tetrahydroxychalcone in *Mt*CHI-1 (Table 1).^{11c} The T49A and T48A mutations also exhibit different effects in *Mt*CHI-1 versus *Ms*CHI-2, in which they reduce k_{cat} values by 10³-fold and 10²-fold, respectively (Table 1).^{11c} Finally, Lys110 plays a unique role in CHI-1s. Mutations of lysine to methionine leave k_{cat} values unchanged, yet increase K_M values 20-fold for K110M in *Mt*CHI-1, but only 2-fold for K109M in *Ms*CHI-2 (Table 1). These results support a role for a positive point charge at this position during the preferential binding and catalytic turnover of substrates in CHI-1 enzymes.

Molecular dynamics simulations have helped clarify the functional significance of the bridging water molecules and conserved active site residues in CHI-1 and CHI-2. In simulations of *Ms*CHI-2 and *Mt*CHI-1, the active site waters molecules are highly mobile, occurring at active site positions analogous to those observed in the crystal structure during up to 19% of the total simulation time in *Ms*CHI-2, and are not observed in the simulation of *Mt*CHI-1 (Figure 3g-h and Table S4). This result corroborates the trend observed from the static crystal structures and suggests that ordered waters are critically positioned in CHI-2 active sites but are poorly accommodated in CHI-1 active sites. In addition, simulations reveal that the intramolecular hydrogen bond between the 6'OH and the ketone oxygen of 4,2',4',6'-tetrahydroxychalcone is disrupted 63.3% of the time in the active site of *Ms*CHI-2 (Figure S8). In contrast, a simulation with *Mt*CHI-1 shows that this intramolecular hydrogen bond is disrupted only 1.2% of the time. Significantly, both tetrahydroxy- and trihydroxychalcone substrates display high conformational variability within the more spacious active site of *Ms*CHI-2 during the simulations, which may facilitate rotation of the A ring into a productive conformation for cyclization. Indeed, simulations of *Ms*CHI-2 with 4,2',4'-trihydroxychalcone capture the substrate in a productive conformation for cyclization that is not observed in simulations of *Mt*CHI-1 with the same substrate (Figures S9 and S11).

CONCLUSIONS

Together these results highlight the mechanistic importance of bifunctional substrate activation by a catalytic arginine in the oxa-Michael addition catalyzed by CHIs. The data presented herein support a model in which CHI-

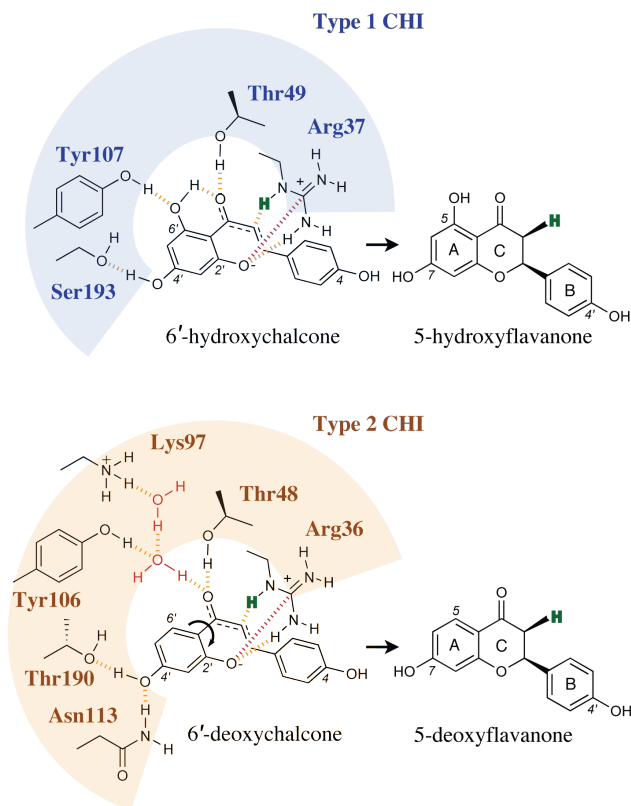


Figure 4. Mechanistic overview of CHI-2 versus CHI-1 catalysis. CHI-2s differ from CHI-1s due to the use of ordered water molecules to facilitate disruption of the intramolecular H-bond and rotation of the A ring into a productive conformation for cyclization of 6'-deoxychalcones. In both CHI-1s and CHI-2s, the catalytic arginine stabilizes the enone and positions the β carbon for nucleophilic addition. The geometry of the guanidine also enables precise positioning of the 2'-O-phenoxy relative to the β carbon for the ensuing enantioselective nucleophilic addition step. Both CHI-1 and CHI-2 enzyme mechanisms terminate with protonation at the alpha carbanion tautomer, delivered *syn* to the arginine residue, suggesting the proton is derived directly from the N ϵ -H ϵ bond of the guanidinium.

1 and CHI-2 bind and promote productive substrate conformations through distinct chemical pathways (Figure 4). CHI-2s utilize a Lys97-mediated, Zundel cation system to facilitate disruption of the C=O--H-O intramolecular hydrogen bond, enabling the flip of the 6'-deoxychalcone A ring into a catalytically productive conformation for ring closure. CHI-1s, which are ubiquitous across the green plant lineage, lack the water-mediated hydrogen bond network and possess low catalytic activity against 6'-deoxychalcones. For both enzyme types, cyclization of chalcones is enhanced through asymmetric guanidine-based catalysis, in which a conserved arginine provides transition state stabilizations through bifunctional Brønsted acid and Lewis acid interactions. At the termination of the reaction, asymmetric flavanone protonation at C α is consistent with the guanidinium of the arginine side chain serving as a Brønsted acid in this mechanism.

Here, we demonstrate the mechanism by which natural plant CHI enzymes utilize guanidinium to catalyze chalcone cyclizations with exquisite, enantioselective fidelity

and appreciable catalytic turnover. As such, these foundational discoveries provide informative starting points to design additional biological, chemoenzymatic, and asymmetric catalysts. In addition to the discovery of a bifunctional role of the delocalized guanidinium ion in CHI, the

open active site architecture, dynamism of the catalytic arginine, and ease of passage through directed evolution,¹⁰ suggest a potential for the CHI scaffold in applications to enzymatic, guanidine-based asymmetric catalysts.

ASSOCIATED CONTENT

Supporting Information

Detailed experimental protocols and computational protocols, crystallographic data, supporting data figures, NMR spectra, HPLC chromatograms, MD simulation data, and CHI-1/CHI-2 phylogenetic analyses are included.

AUTHOR INFORMATION

Corresponding Author

*noel@salk.edu

Present Addresses

Current affiliations: Manus Biosynthesis, Cambridge, MA, United States (R.N.P.) and Howard Hughes Medical Institute, Department of Biology, Washington University in St. Louis, St. Louis, MO, United States (J.M.J.)

Author Contributions

J.R.B., R.N.P., M.E.B., J.M.J., J.W. and J.P.N. performed mutagenesis, protein expression, and biochemical characterization of proteins. J.R.B., J.J.L., G.V.L., and J.P.N. performed and analyzed protein X-ray crystallography and NMR. A.P., G.A.C. and S.C.L.K. performed and analyzed MD simulations. K.B.W. and R.N.P. synthesized 4,2',4',6'-tetrahydrochalcone. M.K. and D.S.T. edited the manuscript. J.R.B., R.N.P., J.M.J., J.J.L., A.T.N., A.P., G.A.C., S.C.L.K. and J.P.N. designed the experiments. J.R.B., J.J.L., R.N.P., J.M.J., A.P., S.C.L.K. and J.P.N. wrote the manuscript. J.P.N. planned and directed the project.

Funding Sources

This work was funded by the Howard Hughes Medical Institute (J.P.N.), United States National Science Foundation grant EEC-0813570 (J.P.N.), the Knut and Alice Wallenberg Foundation (2013.0124), the Wenner-Gren Foundations (post-doctoral fellowship to A. P.), the European Research Council, ERC grant agreement 30647 (S.C.L.K.), Wallenberg Academy Fellowship to SCLK from the Knut and Alice Wallenberg Foundation (KAW 2013.0124). The Swedish National Infrastructure for Computing (SNIC) provided the computer time for the simulations conducted in this study.

ABBREVIATIONS

CHI, chalcone isomerase; CHI-1, type-1 chalcone isomerase; CHI-2, type-2 chalcone isomerase; TBD, triazabicyclo[4.4.0]dec-5-ene; MsCHI-2, *Medicago sativa* type-2 CHI, MtCHI-1, *Medicago truncatula* type-1 CHI, AtCHI-1, *Arabidopsis thaliana* type-1 CHI

REFERENCES

- (1) Bieza, K.; Lois, R. *Plant Physiol.* **2001**, *126*, 1105-1111.
- (2) Falcone Ferreyra, M. L.; Rius, S. P.; Casati, P. *Front. Plant Sci.* **2012**, *3*, 222.

- (3) (a) Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. *Chem. Int. Ed. Engl.* **2009**, *48*, 5001-5004. (b) Lindner, S.; Geu-Flores, F.; Bräse, S.; Sherden, N. H.; O'Connor, S. E. *Chem. Biol.* **2014**, *21*, 1452-1456.

- (4) van der Meer, J. Y.; Poddar, H.; Baas, B. J.; Miao, Y.; Rahimi, M.; Kundendorf, A.; van Merkerk, R.; Tepper, P. G.; Geertsema, E.M.; Thunnissen, A.M.; Quax, W.J.; Poelarends, G.J. *Nat. Commun.* **2016**, *7*, 10911.

- (5) (a) Ley, S. V.; Brown, D. S.; Clase, A. J.; Fairbanks, A. J.; Lennon, I. C.; Osborn, H. M. I.; Stokese Owen, E. S. E.; Wadsworth, D. J. *J. Chem. Soc. Perkin Trans. 1*, **1998**, *15*, 2259-2276. (b) Nicolaou, K. C.; Lim, Y. H.; Papageorgiou, C. D.; Piper, J. L. *Angew. Chem. Int. Ed. Engl.* **2005**, *117*, 8131-8135.

- (6) Taylor, M. S.; Jacobsen, E. N. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 1520-1543.

- (7) (a) Bednar, R. A.; Hadcock, J. R. *J. Biol. Chem.* **1998**, *263*, 9582-9588. (b) Cheng, A. X.; Zhang, X.; Han, X. J.; Zhang, Y. Y.; Gao, S.; Liu, C. J.; Lou, H. X. *New Phytol.* **2018**, *217*, 909-924. (c) Dastmalchi, M.; Dhaubhadel, S. *Planta* **2015**, *241*, 507-523. (d) Ralston, L.; Subramanian, S.; Matsuno, M.; Yu, O. *Plant Physiol.* **2005**, *137*, 1375-1388. (e) Shimada, N.; Aoki, T.; Sato, S.; Nakamura, Y.; Tabata, S. *Plant Physiol.* **2003**, *131*, 941-951.

- (8) (a) Furlong, J. J. P.; Nudelman, N. S. *J. Chem. Soc., Perkin Trans. 2*, **1985**, *5*, 633-639. (b) Miles, C. O.; Main, L. J. *Chem. Soc., Perkin Trans. 2*, **1985**, *10*, 1639-1642.

- (9) Ngaki, M. N.; Louie, G. V.; Philippe, R. N.; Manning, G.; Pojer, F.; Bowman, M. E.; Li, L.; Larsen, E.; Wurtele, E. S.; Noel, J. P. *Nature* **2012**, *485*, 530-533.

- (10) Kaltenbach, M.; Burke, J. R.; Dindo, M.; Pabis, A.; Munsberg, F. S.; Rabin, A.; Kamerlin, S. C. L.; Noel, J. P.; Tawfik, D. S. *Nat. Chem. Biol.* **2018**, *14*, 548-555.

- (11) (a) Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nat. Struct. Biol.* **2000**, *7*, 786-791. (b) Jez, J. M.; Noel, J. P. *J. Biol. Chem.* **2002**, *277*, 1361-1369. (c) Jez, J. M.; Bowman, M. E.; Noel, J. P. *Biochemistry* **2002**, *41*, 5168-5176.

- (12) Park S. H.; Lee, C. W.; Cho, S. M.; Lee, H.; Park, H.; Lee, J.; Lee, J. H. *PLoS One* **2018**, *13*, e0192415.

- (13) (a) Hur, S.; Newby, Z. E. R.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2730-2735. (b) Ruiz-Pernía, J. J.; Silla, E.; Tunon, I. *J. Phys. Chem. B* **2006**, *110*, 20686-20692. (c) Ruiz-Pernía, J. J.; Silla, E.; Tunon, I. *J. Am. Chem. Soc.* **2007**, *129*, 9117-9124. (d) Ruiz-Pernía, J. J.; Tunon, I.; Moliner, V.; Hynes, J. T.; Roca, M. J. *Am. Chem. Soc.* **2008**, *130*, 7477-7488.

- (14) Leow, D.; Tan, C. H. *Chem. Asian J.* **2009**, *4*, 488-507.

- (15) Corey, E. J.; Grogan, M. J. *Organic Letters* **1999**, *1*, 157-160.

- (16) (a) Cho, B.; Tan, C.; Wong, M. W. *J. Org. Chem.* **2012**, *77*, 6553-6562. (b) Xue, H.; Jiang, D.; Jiang, H.; Kee, C. W.; Hirao, H.; Nishimura, T.; Wong, M.W.; Tan, C. H. *J. Org. Chem.* **2015**, *80*, 5745-5752. (c) Yu, Z.; Liu, X.; Zhou, L.; Lin, L.; Feng, X. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 5195-5198.

- (17) (a) Uyeda, C.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2008**, *29*, 9228-9229. (b) Uyeda, C.; Jacobsen, E. J. *J. Am. Chem. Soc.* **2011**, *133*, 5062-5075.

- (18) Lee, R.; Lim, X.; Chen, T.; Tan, G. K.; Tan, C.; Huang, K. *Tetrahedron Lett.* **2009**, *50*, 1560-1562.

- (19) (a) Nibbs A.E.; Scheidt, K.A. *Eur. J. Org. Chem.* **2012**, *3*, 449-462. (b) Biddle, M.M.; Lin, M.; Scheidt, K.A. *J. Am. Chem. Soc.* **2007**, *129*, 3830-3831.

- (20) (a) Gall, M.; Thomsen, M.; Peters, C.; Pavlidis, I. V.; Jonczyk, P.; Grünert, P. P.; Beutel, S.; Scheper, T.; Gross, E.; Backes, M.; Geisler, T.; Ley, J. P.; Hilmer, J. M.; Krammer, G.; Palm, G. J.; Hinrichs, W.; Bornscheuer, U. T. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*,

1439-1442. (b) Thomsen, M.; Tuukkanen, A.; Dickerhoff, J.; Palm, G. J.; Kratzat, H.; Svergun, D. I.; Weisz, K.; Bornscheuer, U. T.; Hinrichs, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 907-917.

(21) (a) Hahlbrock, K.; Wong, E.; Schill, L.; Grisebach, H. *Phytochemistry* **1970**, *9*, 949-958. (b) Hahlbrock, K.; Zilg, H.; Grisebach, H. *Eur. J. Biochem.* **1970**, *15*, 13-18.

(22) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno, B. E.; McIntosh, L.P. *Protein Sci.* **2015**, *24*, 752-761.

(23) van der Kamp, M. W.; Perruccio, F.; Mulholland, A.J. *Chem. Commun.* **2008**, *28*, 1874-1876.

(24) Keenholtz, R. A.; Mouw, K. W.; Boocock, M. R.; Li, N. S.; Piccirilli, J. A.; Rice, P. A. *J. Biol. Chem.* **2013**, *288*, 29206-29214.

(25) Guillen Schlippe, Y. V.; Hedstrom, L. *Biochemistry* **2005**, *44*, 11700-11707.

(26) (a) Dunican, B. F.; Hiller, D. A.; Strobel, S.A. *Biochemistry* **2015**, *54*, 7048-7057. (b) Griffin, M. A.; Davis, J. H.; Strobel, S.A.

Biochemistry **2013**, *52*, 8633-8642. (c) Neubauer, C.; Gao, Y. G.; Andersen, K. R.; Dunham, C. M.; Kelley, A. C.; Hentschel, J.; Gerdes, K.; Ramakrishnan, V.; Brodersen, D.E. *Cell* **2009**, *139*, 1084-1095.

(27) (a) Doherty, M. K.; Pealing, S. L.; Miles, C. S.; Moyley, R.; Taylor, P.; Walkinshaw, M. D.; Reid, G.A.; Chapman, S. K. *Biochemistry* **2000**, *39*, 10695-701. (b) Mowat, C. G.; Moyley, R.; Miles, C. S.; Leys, D.; Doherty, M. K.; Taylor, P.; Walkinshaw, M. D.; Reid, G.A.; Chapman, S.K. *Biochemistry* **2001**, *40*, 12292-12298. (c) Pankhurst, K. L.; Mowat, C. G.; Rothery, E. L.; Hudson, J. M.; Jones, A. K.; Miles, C. S.; Walkinshaw, M. D.; Armstrong, F. A.; Reid, G. A.; Chapman S. K. *J. Biol. Chem.* **2006**, *281*, 20589-20597.

(28) (a) Guillen Schlippe, Y. V., Hedstrom, L. *Arch. of Biochem. Biophys.* **2005**, *433*, 266-278. (b) Raczynska, E. D.; Cyrański, M. K.; Gutowski, M.; Rak, J.; Gal, J.-F.; Maria P.-C.; Darowska, M.; Duczmal, K. *J. Phys. Org. Chem.* **2003**, *16*, 91-106.

(29) Highbarger L. A.; Gerlt, J.A. *Biochemistry* **1996**, *35*, 41-46.

(30) Cukierman, S. *Biochim. Biophys. Acta* **2006**, *8*, 876-885.