# Residues Controlling Facial Selectivity in an Alkene Reductase and Semirational Alterations to Create Stereocomplementary Variants

Adam Z. Walton, Bradford Sullivan, Athéna C. Patterson-Orazem, and Jon D. Stewart*

Department of Chemistry, 126 Sisler Hall, University of Florida, Gainesville, Florida 32611 United States

ABSTRACT: A systematic saturation mutagenesis campaign was carried out on an alkene reductase from *Pichia stipitis* (OYE 2.6) to develop variants with reversed stereoselectivities. Wild-type OYE 2.6 reduces three representative Baylis–Hillman adducts to the corresponding *S* products with almost complete stereoselectivities and good catalytic efficiencies. We created and screened 13 first-generation, site-saturation mutagenesis libraries, targeting residues found near the bound substrate. One variant (Tyr78Trp) showed high *R* selectivity toward one of the three substrates, but no change (cyclohexenone derivative) and no catalytic activity (acrylate derivative) for the other two. Subsequent rounds of mutagenesis retained the Tyr78Trp mutation and explored other residues that impacted stereoselectivity when altered in a wild-type background. These efforts yielded double and triple mutants that possessed inverted stereoselectivities for two of the three substrates (conversions >99% and at least 91% ee (*R*)). To understand the reasons underlying the stereochemical changes, we solved crystal structures of two key mutants: Tyr78Trp and Tyr78Trp/Ile113Cys, the latter with substrate partially occupying the active site. By combining these experimental data with modeling studies, we have proposed a rationale that explains the impacts of the most useful mutations.

KEYWORDS: alkene reductase, old yellow enzyme, protein engineering, X-ray crystallography, mutagenesis, directed evolution
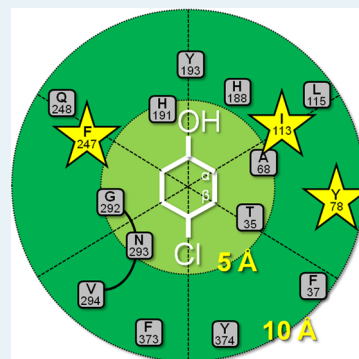
INTRODUCTION

Alkene reductase enzymes have become increasingly popular biocatalysts for converting prochiral substrates into optically pure building blocks.[1] Flavoproteins of the Old Yellow Enzyme superfamily are particularly useful for this purpose, and a variety of homologues have been studied since the seminal work by Massey on *Saccharomyces pastorianus* OYE 1.[2,3] Although *S. pastorianus* OYE 1 and its homologues often show very high stereoselectivities toward many substrates, natural enzymes are practically available only in a single enantiomeric form. This means that only one product enantiomer is readily available. Bioprospecting for stereocomplementary OYE 1 homologues has been successful for a few specific substrate/enzyme pairs,[4] but no generally applicable solution has been identified. Protein engineering is an attractive alternative, and previous success by our group[5] and others[1k,6] has suggested that this strategy might prove fruitful. One major concern with such projects, however, is that the evolutionary pathway may be highly substrate-specific, and it can be difficult to follow a path that encounters a dead end at an intermediate point.[7]

We chose *Pichia stipitis* OYE 2.6 as the starting point for these studies because the wild-type enzyme shows higher stability under process conditions than *S. pastorianus* OYE 1.[8] Moreover, the native enzyme shows very high stereoselectivities toward a series of Baylis–Hillman adducts whose reductions yield useful synthetic building blocks.[9] The major drawback of OYE 2.6 is that it provides only one of the two enantiomers of these products. In addition to addressing this practical problem, our larger goal in this study was to discover key structure–

function relationships in OYE 2.6 in particular and within the OYE superfamily in general.

In principle, there are two strategies that can be used to alter the enantioselectivity of an alkene reductase: move the substrate to the opposite side of the FMN cofactor (and provide a new general acid for α-carbon protonation) or flip the orientation of substrate binding within the existing active site. Because only one face of the FMN is accessible within the protein environment, moving the substrate binding site to the opposite side of the cofactor appeared impractical. We therefore concentrated on changing the substrate binding orientation by altering residues that directly or indirectly contacted the alkene. Our strategy was inspired by the iterative saturation mutagenesis approach of the Reetz group (for recent examples, see ref 10 and references therein), in which the "best" single mutants from first-round site-saturation libraries serve as the basis for subsequent site-saturation mutagenesis libraries at other key positions (also identified within the same collection of first-generation libraries). The key problem in alkene reductase engineering is to identify those residues that control substrate binding orientation with as little effort as possible. This is one issue that we hoped to address for this class of enzymes in the present study.

To avoid becoming trapped within local sequence space minima during directed evolution and to provide more generally useful structure–function correlations, we carried
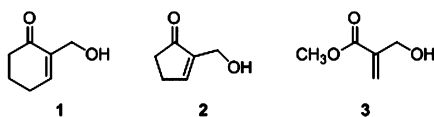
out three parallel studies. Each focused on a different alkene substrate (Figure 1). Wild-type OYE 2.6 reduces Baylis–

**Figure 1.** Structures of Baylis–Hillman substrates employed in this study.

Hillman adducts **1** and **3** to the corresponding *S* products with very high enantioselectivities and good rates. It also converts cyclopentenone **2** to the *S* product in 76% ee.[11] Unfortunately, no synthetically useful, complementary *R*-selective alkene reductase with good catalytic efficiency has been identified for these substrates. The only exception was wild-type *S. pastorianus* OYE 1, which reduces Roche ester precursor **3** to the *R* product with >98% ee; however, the conversion rate is too low to be practically useful.[9] Our goal was therefore to evolve *R*-selective, enantiocomplementary catalysts for **1–3**, starting from wild-type OYE 2.6.

■ RESULTS AND DISCUSSION

Residue Selection. In the high-resolution OYE 2.6/*p*-chlorophenol complex (PDB code 4DF2),[11] the phenolic oxygen lies between the side chains of His188 and His191. On the basis of our earlier experience with *S. pastorianus* OYE 1,[9] this is also the likely position of the substrate carbonyl oxygen in electron-deficient, conjugated alkenes (Figure 2). The α- and β-carbons of the conjugated system can likewise be inferred from the positions of carbon atoms in the *p*-chlorophenol ring. On the basis of this presumed substrate binding mode, we sought residues that might control the stereoselectivity of OYE 2.6 by direct interactions with the substrate. Our selections were guided by three criteria. First, we only considered residues lying above the *si* face of the flavin since the cofactor's *re* face is shielded by the protein interior. In addition, we also concentrated on residues whose side chains were oriented toward the center of the active site. Finally, we selected only residues whose distances from the likely locations of bound substrates made direct interactions plausible.

Once the initial list of candidates that fulfilled all three criteria had been generated, we measured the distance from the residue's C_β to the nearest atom of the bound inhibitor; the results are shown schematically in Figure 2. Those residues with C_β/inhibitor distances of ≤5 Å were predicted to have high probabilities of interacting directly with bound substrates (and therefore impacting stereoselectivity). Those with distances between 5 and 10 Å were judged somewhat less likely to impact substrate binding. These efforts yielded a group of 14 residues considered as candidates for site-saturation mutagenesis.[12] Of these, two were subsequently eliminated because their side chains were partially masked by adjacent active site residues (Leu115 and Gln248).

One of the most significant differences between the structures of OYE 2.6 and *S. pastorianus* OYE 1 occurs in a loop region centered on Pro295 (OYE 1 numbering). This loop is shorter by several residues in OYE 2.6, which lacks a direct counterpart to Pro295. This allows the side chain at position 292 to project into the active site, and this residue was therefore added to our group of candidates.[13] Finally, in OYE 2.6, the side chain of Tyr78 shifts position upon *p*-chlorophenol
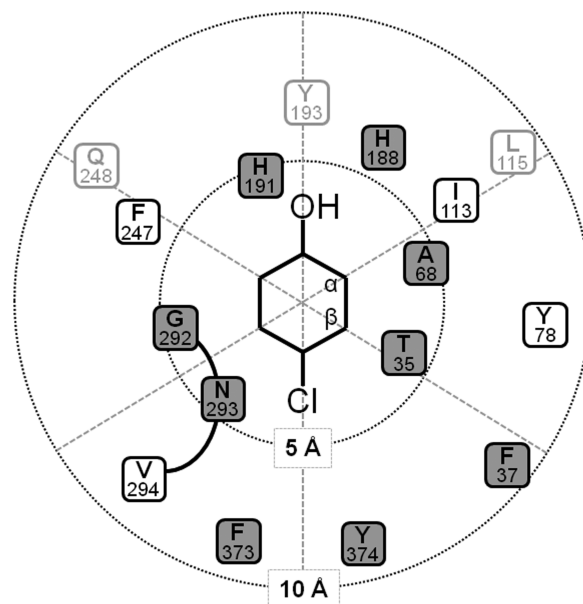
Figure 2. Schematic diagram of the OYE 2.6 active site and residues targeted for site-saturation mutagenesis. In this view, the FMN lies beneath the substrate analogue (*p*-chlorophenol). Distances between the C_β atom of selected side chains and the center of the aromatic ring of the inhibitor were determined from the X-ray crystal structure of OYE 2.6 and the centers of the boxes depicting the residues correspond to this distance. For reference, circles of 5 and 10 Å are shown in dotted lines. The heavy line linking Gly292, Asn293, and Val294 indicates the loop connecting these residues. Residues shown in gray type were not included in the site-saturation libraries (Leu115, Tyr193, and Gln248). Residues whose boxes are gray yielded no improved variants in site-saturation libraries when tested against the three substrates examined in this study.

binding. This alternate conformation would place its C_β significantly closer to a bound substrate, and we therefore included it in our group of mutagenesis candidates.

First-Generation Libraries. This project involved three parallel directed evolution studies, and each sample required individual chiral-phase GC analysis (~15 min/sample). Compact and complete site-saturation libraries devoid of artifacts were therefore essential to avoid unnecessary additional screening efforts. We used a rolling circle methodology to prepare site-saturation libraries at each of the 13 selected positions using an NNK degeneracy scheme.[14] This approach afforded a high probability of sampling all 20 amino acids in a library contained on a single 96-well plate. Library creation was optimized in a previous study to maximize diversity at the targeted position, minimize wild-type carryover, and eliminate primer concatamers.[15,16] All pooled libraries were analyzed quantitatively after the initial transformation step by fluorescence sequencing to assess diversity at the randomized positions (Table 1). The Q_{pool} value indicates the base composition at all targeted positions within a given library (0.0 = no diversity; 1.0 = perfect randomization).[15] In our experience, libraries with Q_{pool} values ≥0.70 likely contained at least 17 of the 20 amino acids within a 95-member collection. In some cases, individual library members were sequenced prior to screening to verify the relationship between Q_{pool} values and amino acid content. This also allowed a single example of each codon to be preselected prior to screening, which decreased the number of mutants to 32 (or fewer if a given codon was absent from a library).

Table 1. Summary of OYE 2.6 libraries[a]

starting protein	randomized position	Q_{pool} value
wild-type	Thr35[b]	0.76
	Phe37	0.86
	Ala68	0.73
	Tyr78	0.82
	Ile113[b,c]	0.78
	His188[b]	0.80
	His191[b]	0.89
	Phe247	0.83
	Gly292	0.80
	Asn293	0.83
	Val294	0.81
	Phe373	0.74
	Tyr374	0.73
Tyr78Trp	Ile113	0.78
	Phe247	0.74
	Val294	0.84
Ile113Asp	Tyr78	0.91
	Val294	0.80
Ile113Trp	Tyr78	0.80
	Val294	0.83
Val294Pro	Tyr78	0.93
	Ile113	0.84
Tyr78Trp, Phe247Ala	Ile113	0.82
Tyr78Trp, Phe247His	Ile113	0.82

[a]Q_{pool} values indicate nucleotide diversity across all three bases of the targeted codon (0.0 = single base present, 1.0 = all relevant nucleotides present at equal levels). Values were calculated from fluorescence sequencing data as described in ref 15. Codons were randomized by an NNK doping scheme that encodes all 20 amino acids. Libraries with $Q_{pool} \geq 0.75$ likely contain at least 15 of the 20 possible amino acids at the indicated position and were considered acceptable for screening experiments. [b]From ref 11. [c]Codons absent from this library were subsequently added individually to achieve complete sequence space coverage.

Individual clones were grown in 96-deepwell plates under autoinducing conditions, then cell pellets were harvested and used for whole-cell alkene reductions under nongrowing conditions. For libraries sequenced prior to screening, one example of each codon was assayed;[17] when individual sequences were unknown, we assessed 95 randomly chosen mutants along with a wild-type control. Reaction mixtures were extracted, and the extents of conversion and product optical purities were determined by chiral-phase GC. Clones showing improved properties were sequenced (if necessary) to determine the amino acid replacement responsible. In all cases, mutations were confined to the targeted positions, and no spurious changes were detected.

Our previous studies had shown that altering Thr35, His188, or His191 did not significantly change the facial selectivity of alkene binding within the OYE 2.6 active site.[11] This was also the case for Phe37, Ala68, Gly292, Asn293, Phe373, and Tyr374 (some or all of the mutants at these positions had lower percent conversion values, but all those retaining activity showed essentially unchanged S stereoselectivity for 1−3) (Figure 2). Clearly, despite their proximity to the bound substrate, these residues—when altered singly—do not strongly impact substrate binding.

A handful of single-site changes *did* impact the orientation of substrate binding in OYE 2.6. Changing Val294 to Pro somewhat diminished the enzyme's S selectivity toward

cyclohexenone **1** (from 95% to 61% ee, S), although this was accompanied by relatively poor conversion (∼10%). No other variant at position 294 showed any change in stereoselectivity toward **1**. Moreover, the Val294Pro mutant (along with all other Val294 replacements) reduced alkenes **2** and **3** with the same stereoselectivity as wild-type OYE 2.6. These results provided the first evidence that changes to the OYE 2.6 active site did not affect binding of the three homologous substrates equally. They also showed that proximity to the bound substrate was not the sole determinant of importance with respect to stereoselectivity

Substituting either Tyr or Trp for Phe247 decreased S selectivity for cyclopentenone **2** by a small amount without harming conversion. Other amino acid changes at this position were ineffective for **2**, and no alterations were identified that altered the stereochemistry of reduction for substrates **1** and **3**.

By contrast, changes at Ile113 altered the stereochemistry of alkene reductions more significantly. Ile113 in OYE 2.6 corresponds to Trp116 in *S. pastorianus* OYE 1, and our previous experience with Trp116 in OYE 1 underscored the dramatic effect of changes at this active site location.[9,18] In the case of OYE 2.6, the impacts were more subtle: the Asp mutant showed a significant decrease in stereoselectivity toward cyclohexenone **1** (but not toward alkenes **2** or **3**), whereas substituting Ile113 with Trp gave slight R selectivity for **2** (but unchanged stereoselectivity for **1** and **3**). Unfortunately, the enhanced R selectivity came at the cost of poor conversion efficiency (Figure 3A).

Replacing Tyr78 with Trp caused the largest alteration in OYE 2.6's stereoselectivity by a single amino acid change. This substitution reversed the enzyme's stereoselectivity toward cyclopentenone **2** (from >99% ee (S) to 63% ee (R)) while still retaining good catalytic efficiency. Interestingly, this phenomenon was restricted to Trp, and no other amino acid replacement showed similar effects (Figure 3B).[19] Moreover, although the Tyr78Trp mutant also reduced cyclohexenone **1** with high efficiency (>99% conversion after 24 h), its stereoselectivity toward this substrate was unchanged (>98% ee (S)). Finally, it was important to note that this mutant was ineffective at reducing Roche ester precursor **3**.

Three major themes emerged from our first-round site-saturation libraries. First, side chain proximity to the bound substrate, although relevant, was not a perfect predictor of control over substrate binding orientation. Of the 13 targeted positions, only four—Tyr78, Ile113, Phe247, and Val294—actually impacted stereoselectivities upon mutation. A second, unexpected lesson was that the changes in stereoselectivity provoked by mutagenesis were not consistent, even among structurally similar substrates. Indeed, the most interesting mutant, Tyr78Trp, showed significant R stereoselectivity and high activity toward cyclopentenone **2**, but no change for **1** and no catalytic activity whatsoever for **3**. *It is also important to note that none of the three substrates, if tested alone, would have uncovered all four improved OYE 2.6 variants.*

Crystal Structure of the Tyr78Trp Mutant. We solved the Tyr78Trp OYE 2.6 crystal structure to understand why this mutation almost completely reversed the enzyme's enantioselectivity toward alkene **2** (Table 2). Trp substitution did not appreciably change the overall protein structure, and its indole ring lay in approximately the same plane as the phenol moiety of the wild type Tyr at this position (Figure 4). Nearby neighbors included the side chains of Leu115, Ala68, Phe70, and Ile113. A Trp at position 78 adds to the existing
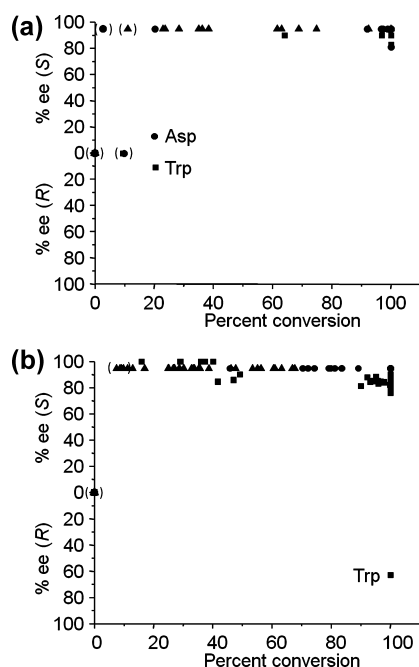
Figure 3. Results from first-generation site-saturation libraries. (A) Ile113 library. A collection of wild-type and all 19 amino acid replacements at position 113 of OYE 2.6 was tested against Baylis–Hillman substrates **1** (●), **2** (■), and **3** (▲). Optical purity data for points enclosed by parentheses could not be determined accurately because of the low conversion levels. Data points for two key variants (Asp and Trp) are highlighted. (B) Trp 78 library. A collection of 95 clones containing NNK-randomized bases at position 78 of OYE 2.6 was tested against Baylis–Hillman substrates **1** (●), **2** (■), and **3** (▲). Optical purity data for points enclosed by parentheses could not be determined accurately because of the low conversion levels. The data point for one key variant (Trp) is highlighted.

hydrophobic character of this pocket, and the decreased polarity of this active site region may be one factor that influences the orientation of substrate binding by discouraging the polar hydroxymethyl side chain of alkene **2** from binding in this region.

Beyond polarity, we also identified another, more direct impact of Trp substitution on the facial selectivity of substrate binding: an indole side chain at position 78 will sterically interfere with the hydroxyl side chain of alkene **2** and thereby prevent it from occupying this location in the "normal" binding mode. In wild-type OYE 2.6, an ordered solvent molecule (WAT 611) is hydrogen-bonded to the side chain of Tyr78[11] (Figure 4). Modeling suggested that this water was replaced by the substrate hydroxyl group upon binding of Baylis–Hillman substrates. No ordered solvent molecule corresponding to WAT 611 was observed in the Tyr78Trp variant, and it would be impossible to place a water at the same position in the mutant because the closest approach distance would be only 1.8 Å. This implies that alkene **2** would be unable to place its side chain hydroxyl group at an analogous position. The most logical alternative would be a "flipped" binding mode that places the substrate side chain on the other side of the active site. This arrangement would yield the opposite stereoisomer upon alkene reduction, as was observed experimentally.

Second- and Third-Generation Libraries. In a typical ISM strategy, one uses the single "best" mutant from the previous round as an "anchor" for libraries focused on other positions that had positive (albeit lower) impacts in first-round

Table 2. Protein Crystal Data Collection and Refinement Statistics

protein	Y78W OYE 2.6	Y78W/I113C OYE 2.6
accession code	4QAI	4M5P
observed active site ligand	None	malonate, alkene **3**
X-ray source	NSLS beamline X25	NSLS beamline X6A
space group	$R32:H$	$P6_322$
unit cell dimensions		
$a = b$, c (Å)	226.8, 360.9 (90°, 120°)	126.5, 123.3 (90°, 120°)
resolution (Å)	48.65–2.75	35.01–1.50
unique reflections	92 137 (9067)[a]	91 708 (8830)
completeness (%)	99.47 (98.63)	99.4 (97.0)
I/σ (I)	12.47 (2.4)	15.7 (2.1)
R_{work}[b], R_{free}[c]	0.224, 0.280	0.141, 0.170
Ramachandran statistics[d]		
favored (%)	91	97
allowed (%)	7.5	2.8
outliers (%)	1.5	0.2
no. of protein, solvent, ligand atoms	19 218, 186, 22	3243, 45, 498
average B factors (Å²)		
protein	55.1	16.2
solvent	30.0	25.6

[a]Values in parentheses denote data for the highest resolution bin. [b]$R_{work} = \Sigma |F_0 (hkl)| - |F_c (hkl)|/\Sigma |F_0 (hkl)|$. [c]$R_{free}$ is calculated in the same manner as R_{work} using 10% of the reflection data not included during the refinement. [d]Statistics generated using MOLPROBITY.[41]
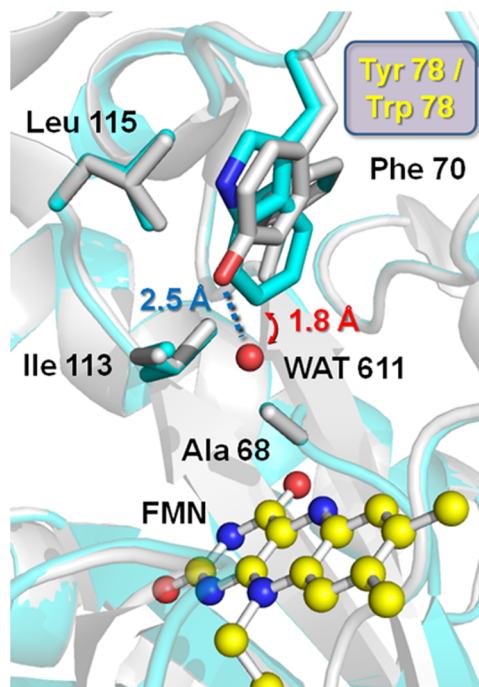

Figure 4. Active sites of wild-type and Tyr78Trp OYE 2.6. A global best fit between the structures of wild-type (PDB code 3TJL, gray) and Tyr78Trp OYE 2.6 (PDB code 4QAI, cyan) was calculated by PyMOL.[40] Key active side chains are shown along with an ordered solvent molecule (WAT 611) from the wild-type structure. The observed distances between WAT 611 and Tyr78 as well as that calculated between Trp78 in the overlaid structures are indicated.

libraries.[10a] This was not appropriate in our case, however, because the three different substrates yielded either different

"best" mutants in round 1 (alkenes **1** and **2**) or no significant improvement (alkene **3**). Moreover, because our goal was to understand structure−function relationships in substrate binding by OYE 2.6, we pursued parallel strategies for all three substrates that started from different "best" round 1 mutants, regardless of whether they were optimal in the context of single amino acid changes. By doing this, we hoped to answer three key questions. First, would the multiple, parallel evolutionary pathways converge to the same solution, or would we obtain different sets of mutations for a given substrate, depending on the pathway? Second, would the same mutations be optimal for all three of the structurally related substrates, or would the solutions be idiosyncratic? Finally, could we uncover useful multiply mutated proteins from singly mutated precursors that lacked detectable catalytic activity?

We created second-generation libraries from four "anchor" mutants that showed significant changes in substrate binding orientation preferences (as reflected by enhanced *R* stereoselectivity)—Tyr78Trp, Ile113Asp, Ile113Trp, and Val294-Pro—using the same methods as described above (Table 1, Figure 5). All of these libraries were screened "blind", that is,
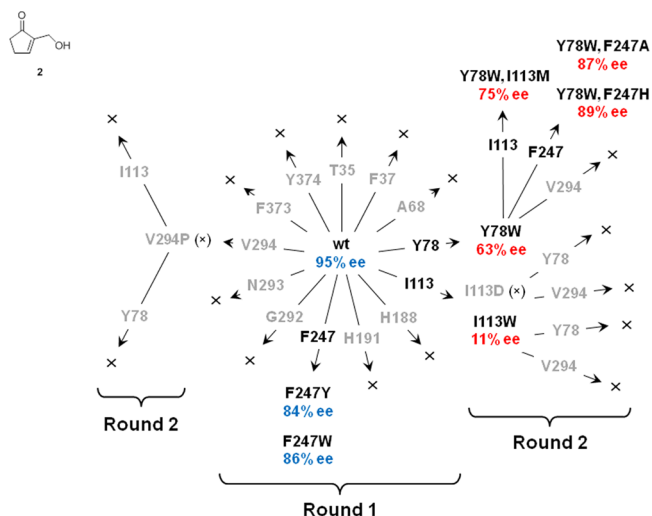
Figure 5. Summary of mutagenesis focused on cyclopentenone **2**. Starting from wild-type OYE 2.6, 13 different amino acid positions were subjected to site-saturation mutagenesis, and the resulting first-generation libraries were screened against alkene **2** (round 1). Enantioselectivities of key proteins are indicated (red = *R*; blue = *S*). Libraries that failed to yield mutants with enhanced *R* selectivity are shown in gray type and marked by "×". Four useful variants were identified in round 1: Y78W, I113W, F247Y, and F247W. These four, as well as two others identified in a screen against substrate **1**, were used as starting proteins for site-saturation libraries anchored by the indicated mutations (round 2). Starting mutants with no change in stereoselectivity or no catalytic activity toward substrate 1 are indicated by gray type and "(×)".

not sequenced prior to screening, and only those clones showing improved properties were subsequently examined by fluorescence sequencing. Because the changes were most significant for cyclopentenone **2**, this substrate is discussed first, and the evolutionary path for this substrate is summarized in Figure 5.

Cyclopentenone 2. *Val294 Pro Anchored Libraries.* We created two site-saturation libraries starting from the Val294 Pro variant. In a wild-type background, this change decreased *S* stereoselectivity toward **1**, although catalytic efficiency was also

significantly reduced. It was therefore not surprising that variations at both Tyr78 and Ile113 in a Val294 Pro background showed no improvement with respect to alkene **2**, and this anchor position was dropped from further consideration.[20]

Tyr78Trp Anchored Libraries. The side chain of Tyr78 is located on the "eastern" side of the OYE 2.6 active site (as viewed from above). Since replacement by Trp reversed the stereoselectivity toward alkene **2**, we used this mutant as a starting point for three second-generation libraries focused on amino acids found important in other contexts (Ile113, Phe247, and Val294) (Figure 5). When Val294 was randomized in a Tyr78Trp background, no additional change in protein properties toward **2** was observed (the substrate was reduced to the *R* product in ∼80% ee).

When Ile113 was varied in a wild-type background, changes at this position had very little impact on stereoselectivity toward **2** (Figure 3A). By contrast, when the same residue was examined in the context of a Tyr78Trp anchor, most Ile113 replacements showed worse conversion or stereoselectivity as compared with the parent, although the Tyr78Trp/Ile113Met double mutant showed slight improvement (75% ee (*S*)) (Figure 6). Given the close proximity of the side chains of
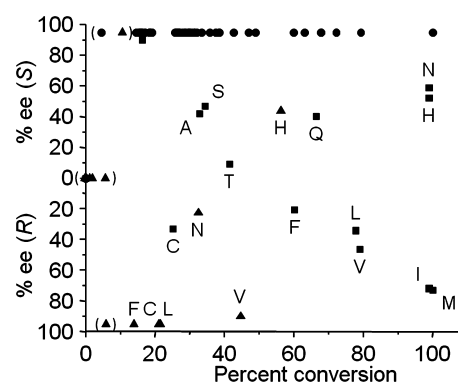
Figure 6. Influence of Trp at position 78 on the impact of amino acid replacements for Ile113. A site-saturation library was created at position 113 (wt = Ile) in the Y78W variant of OYE 2.6. A collection 95 clones containing NNK-randomized bases at position 113 of OYE 2.6 was tested against Baylis−Hillman substrates **1** (●), **2** (■), and **3** (▲). Optical purity data for points enclosed by parentheses could not be determined accurately as a result of the low conversion levels. Data points for several key variants are highlighted.

amino acids at positions 78 and 113, it is perhaps not surprising that the effects of mutations at these locations would be nonadditive. Although covarying these two residues simultaneously might be logical strategy, the long analysis times required for complete library coverage dissuaded us from following this avenue.

We also prepared the converse second-generation libraries (randomized at Tyr78 or Val294) in an Ile113 mutant background. The "best" Ile113 changes were originally identified for alkene **1** (Ile113Asp) or **2** (Ile113Trp). Unfortunately, no improved variants were detected for cyclopentenone **2** in the second-generation libraries (Figure 5). This further underscores the cooperativity between side chains at positions 78 and 113.

Finally, we also examined the effects of replacing Phe247 in the context of pre-existing a Tyr78Trp mutation. Changes here in a wild-type background had very little impact on stereo-

selectivity in round 1 (a slight decrease in *S* selectivity toward **2**, but no improvement for **1** or **3**). Nonetheless, we selected this position for a second-generation library because the side chain of Phe247 is located on the "western" side of the active site. We reasoned that a bulky indole side chain at position 78 might have altered the steric environment on the "eastern" side, which might now make amino acid replacements on the "western" side more important. Although contrary to a strict ISM strategy, this turned out to be the case: both the Ala and His replacements for Phe247 showed significantly higher stereoselectivity for cyclopentenone **2** in a Tyr78Trp background, as compared with the parent single mutant (Figure 5). In fact, these double mutants showed almost completely reversed stereoselectivities toward substrate **2** with good conversion, thereby providing a pair of practically useful, enantiocomplementary counterparts to wild-type OYE 2.6.

Cyclohexenone 1. After successfully inverting the stereoselectivity of OYE 2.6 toward cyclopentenone **2** by altering residues Tyr78, Ile113, and Phe247, we examined the effects of these and other positions on the binding of cyclohexenone **1**. All nine second-generation libraries were individually screened for stereoselectivity in the reduction of **1** (Figure 7). Two of the
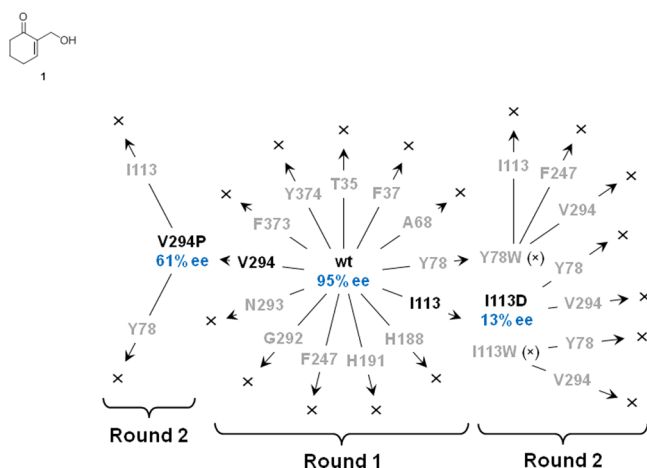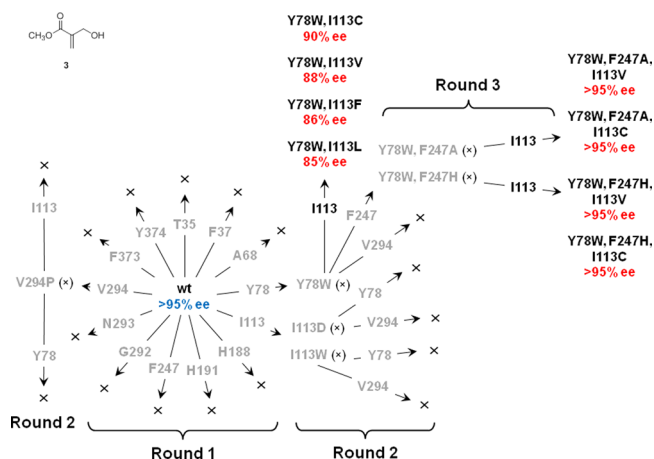

Figure 7. Summary of mutagenesis focused on cyclohexenone **1**. Starting from wild-type OYE 2.6, 13 different amino acid positions were subjected to site-saturation mutagenesis, and the resulting libraries were screened for reduction of **1** (round 1). Enantioselectivities of key proteins are indicated (red = *R*; blue = *S*). Libraries that failed to yield mutants with greater *R* selectivity are shown in gray type and marked by "×". Two useful variants were identified in round 1: I113D and V294P. These two, as well as four others identified in a screen against substrate **2**, were used as starting proteins for site-saturation libraries anchored by the indicated mutations (round 2). Starting mutants with no change in stereoselectivity or no catalytic activity toward substrate 1 are indicated by gray type and "(×)".

anchor positions (Val294Pro and Ile113Asp) were optimal first-generation substitutions for **1** with diminished *S* stereoselectivity in a wild-type background. Unfortunately, variations at additional positions (even those critical for controlling the binding of **2**) yielded no additional improvement toward **1** in any of the second-generation libraries. This was particularly surprising because the same libraries had yielded significant differences in substrate binding orientation toward the homologous Baylis–Hillman adduct **2**. Despite their structural similarities, subtle differences in binding interactions between the way that OYE 2.6 binds alkenes **1** and **2** and must be

present and account for the difference in facial selectivity (vide infra).

Roche Ester Precursor 3. The major challenge in altering OYE 2.6's facial selectivity toward binding alkene **3** was that no first-generation library yielded any variants that moved stereoselectivity in the desired *R* direction. *In a traditional ISM approach, this means that the program must be abandoned because no pathway is available for further improvement.* Despite these apparently bleak prospects, we screened all nine second-generation libraries for the ability to reduce alkene **3**. Eight of the nine libraries failed to provide a mutant with improved *R* selectivity (Figure 8).


Figure 8. Summary of mutagenesis focused on Roche ester precursor **3**. Starting from wild-type OYE 2.6, 13 different amino acid positions were subjected to site-saturation mutagenesis, and the resulting libraries were screened for reduction of **1** (round 1). Enantioselectivities of key proteins are indicated (red = *R*; blue = *S*). Libraries that failed to yield mutants with greater *R* selectivity are shown in gray and marked by "×". Note that no improved variants were identified in any first-generation library. Nine second-round libraries were examined, each with an anchor mutation that improved *S* selectivity toward substrate **1** or **2**. One second-round library (Ile113 in a Tyr78Trp background) yielded four variants with good *S* stereoselectivity toward **3**. Randomizing the same position in a double mutant background (Tyr78Trp/Phe247Ala or Tyr78Trp/Phe247His) also provided four mutants, each with >95% ee (*S*).

The pivotal observation was that varying Ile113 in a preexisting Tyr78Trp background yielded several variants with very high *R* stereoselectivities and good conversions (Figure 8). Replacing Ile113 with Cys, Val, Phe, or Leu afforded proteins with ≥85% ee (*R*). *This behavior is remarkable because the Tyr78Trp single mutant had no detectable catalytic activity toward* **3**. We also carried out a third-generation library study anchored by the Tyr78Trp/Phe247Ala or Tyr78Trp/Phe247His double alterations and varying Ile113, despite the fact that both of these starting proteins showed no catalytic activity toward alkene **3**. Again, we recovered activity when Ile113 was substituted by smaller residues. As before, Cys and Val were the most effective in conferring *R* selectivity, in this case ≥95% ee. That all the Ile113 libraries converged to the same two "best" choices (Cys and Val) suggests that these replacements are optimal in a Tyr78Trp background; the role of changing Phe247 is to fine-tune the *R* stereoselectivity. From a practical point of view, these two triple mutants have solved the original synthetic problem. We can prepare both the *S* and *R* Roche

ester antipodes in >95% ee using wild-type OYE 2.6 or one of the triple mutants, respectively.

Crystal Structure of the Tyr78Trp/Ile113Cys Double Mutant. To help understand why these two simultaneous mutations so dramatically altered the catalytic activity and stereoselectivity of OYE 2.6 with respect to alkene **3**, we solved the crystal structure of this double mutant. As expected, the overall protein architecture was essentially identical to that of the wild type. After molecular replacement and a few rounds of refinement, however, it became apparent that electron density in the vicinity of Cys 113 could not be fit by a single side chain conformation. Even allowing for multiple Cys side chain conformations failed to account for all of the observed electron density. We therefore hypothesized that sulfur oxidation had taken place before or during X-ray data collection.[21] We attempted to fit the observed electron density data by substituting the most commonly observed cysteine side chain oxidation products at position 113 and carried out one round of refinement.[22] When no single residue type, even with multiple conformations, could account for all of the observed electron density, judicious combinations of an unmodified Cys and a Cys oxidation product were tried. The best fit to the observed electron density was obtained from a native cysteine residue along with a 3-sulfinoalanine with occupancies of 0.35 and 0.65, respectively (Figure 9).
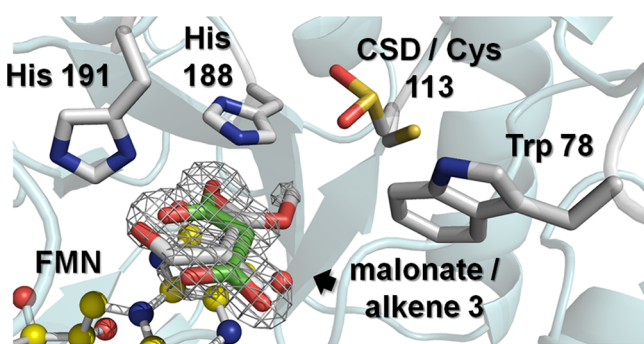
Figure 9. X-ray structure of the Tyr78Trp/Ile113Cys mutant complexed with alkene **3**. Electron density in the vicinity of position 113 was best fit by a combination of Cys and 3-sulfinoalanine (PDB residue code CSD) with occupancies of 0.35 and 0.65, respectively. Electron density in a plane above the FMN could be explained by a combination of malonate and alkene **3**, with occupancies of 0.7 and 0.3, respectively. The locations of key active site residues are indicated.

By postcrystallization soaking with alkene **3**, we hoped to observe a pseudo-Michaelis complex in which the substrate was bound in the "flipped" binding orientation. We indeed found positive electron density in a plane above the FMN that could not be explained by either protein residues or the cofactor. Malonate (present in the crystallization buffer at 1.2−2.4 M) is the "default" ligand in this location for wild-type OYE 2.6,[11] and malonate accounted for some of the observed electron density in the double mutant; however, closer inspection revealed that malonate bound by the double mutant was rotated by ~23° relative to its location in the wild-type protein.[11] This placed one of its carboxylate moieties farther from the density associated with the side chain at position 113, which may be due to the negative charge of 3-sulfinoalanine. Malonate alone, however, could not explain all of the electron density above the FMN cofactor.[23] We therefore modeled alkene **3** into the active site (along with malonate). The

carbonyl oxygen of **3** was fixed within hydrogen bonding distance of both His188 and His191 because these interactions are known to be crucial for catalysis by OYE 2.6.[11,24] The ligand was then rotated about this point until the π-electron system of the alkene coincided with the additional positive electron density above the FMN and the best fit was obtained after refinement. This located the alkene terminus near the side chains of Thr35 and Tyr374 (Figure 9). Occupancies of 0.7 and 0.3 for malonate and alkene **3**, respectively, gave the best fit to the observed electron density data.

Two features argue that the modeled complex with **3** in Figure 9 closely mimics the actual Michaelis complex that precedes alkene reduction. First, the β-carbon of alkene **3** (C_4) was 3.34 Å away from N_5 of the FMN and the $C_4-N_5-N_{10}$ angle was 100.6°. Both values are within the ranges observed for efficient hydride transfer from flavins.[25] Second, this binding mode predicts that **3** would be reduced to the R product, consistent with behavior observed experimentally.
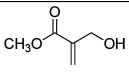
The crystallographically deduced binding arrangement for alkene **3** placed the hydroxymethyl side chain of alkene **3** near Gly292 and Asp293, which are part of a flexible active site loop (loop 6) that is rather hydrophilic. The best fit to the observed electron density suggested an intramolecular hydrogen bond between the hydroxyl group of alkene **3** and the carbonyl oxygen.[26] This "flipped" binding mode located the methoxy group of alkene **3** in a pocket partially formed by the two mutated resides (Trp78 and Cys113).

On the basis of these data and structures of wild-type OYE 2.6 and the Tyr78Trp single mutant, we can provide a rationale for the experimental results observed during library screening. Using the criteria described above, we modeled reasonable complexes of alkene **3** in both the "normal" and "flipped" orientations with wild-type OYE 2.6 as well as the Tyr78Trp and Tyr78Trp/Ile113Cys mutants. When alkene **3** is bound to the wild type protein in the "normal" orientation, the substrate hydroxymethyl group is located sufficiently close to the side chain of Tyr78 to form a critical hydrogen bond with this residue (mimicking the interaction of this residue with WAT 611 in Figure 4). The alternative, "flipped" orientation is highly disfavored in wild-type OYE 2.6 because the methoxy moiety is too large to fit in the pocket partially defined by Tyr78 and Ile113. The net result is that wild-type OYE 2.6 can bind **3** productively only in the "normal" orientation, which yields the S product.

When Tyr78 is changed to Trp, there is no longer sufficient room for the hydroxymethyl moiety of alkene **3** in the "normal" orientation (Figure 4). As noted above, the methoxy group is also too large for this active site volume in the alternative, "flipped" orientation. Because the substrate cannot be accommodated productively in *either* binding mode, the Tyr78Trp mutant loses the ability to reduce **3**. On the other hand, when the hydroxymethyl side chain of alkene **3** cannot hydrogen bond with the Tyr at position 78 (because it was mutated Trp), *but* additional room is created in this region by changing Ile to Cys to accommodate the methoxy group, the "flipped" binding mode is not only enabled, but also favored. This is also consistent with the observation that all of the best R-selective double mutants (anchored with Tyr78Trp) for alkene **3** possessed smaller residues at position 113 as compared with the wild-type Ile.

Overivew of Library Results. One key goal in this project was to determine which residue(s) controlled substrate binding orientation in OYE 2.6. On the basis of our results, Tyr78 and

Table 3. Catalytic Properties of Purified OYE 2.6 and Key Variants

Protein		1	2	3
Wt	Reaction time (hr)	<1	<1	<1
	% Conversion	>99%	>99%	>99%
	% ee	>98% (S)	95% (S)	>98% (S)
Tyr 78 Trp	Reaction time (hr)	4	6	24
	% Conversion	98	99%	10%
	% ee	81% (S)	75% (R)	---[a]
Tyr 78 Trp / Ile 113 Met	Reaction time (hr)	24	8	24
	% Conversion	99%	99%	10%
	% ee	79% (S)	75% (R)	---
Tyr 78 Trp / Ile 113 Cys	Reaction time (hr)	24	24	24
	% Conversion	98%	65%	57%
	% ee	60% (S)	16% (R)	90% (R)
Tyr 78 Trp / Ile 113 Phe	Reaction time (hr)	4	24	24
	% Conversion	99%	99%	81%
	%ee	79% (S)	19% (R)	91% (R)
Tyr 78 Trp / Ile 113 Leu	Reaction time (hr)	4	24	24
	% Conversion	99%	99%	96%
	% ee	81% (S)	27% (R)	87% (R)
Tyr 78 Trp / Ile 113 Val	Reaction time (hr)	24	24	24
	% Conversion	99%	99%	78%
	% ee	81% (S)	40% (R)	90% (R)
Tyr 78 Trp / Phe 247 Ala	Reaction time (hr)	24	4	24
	% Conversion	43%	99%	26%
	% ee	37% (S)	91% (R)	>98% (R)
Tyr 78 Trp / Phe 247 His	Reaction time (hr)	24	2	24
	% Conversion	84%	99%	30%
	% ee	73% (S)	91% (R)	94% (R)
Tyr 78 Trp / Ile 113 Val / Phe 247 Ala	Reaction time (hr)	24	24	24
	% Conversion	31%	99%	91%
	% ee	55% (S)	69% (R)	>98% (R)
Tyr 78 Trp / Ile 113 Cys / Phe 247 Ala	Reaction time (hr)	24	24	24
	% Conversion	10%	79%	72%
	% ee	36% (S)	47% (R)	>98% (R)
Tyr 78 Trp / Ile 113 Val / Phe 247 His	Reaction time (hr)	24	24	24
	% Conversion	75%	99%	97%
	% ee	77% (S)	67% (R)	>98% (R)
Tyr 78 Trp / Ile 113 Cys / Phe 247 His	Reaction time (hr)	24	24	24
	% Conversion	39%	85%	86%
	% ee	56% (S)	52% (R)	>98% (R)

[a]Enantiomeric purity could not be accurately assessed because of low conversion.

Ile113 are most critical, with Phe247 playing a subordinate role. The first two residues make close contacts within the active site, and the synergistic impacts of mutations at these positions are not unexpected. Together, they form the "eastern" wall of the active site. The side chain of Phe247 is located on the opposite site, and by substituting this amino acid by with smaller homologues, additional space is created to accommodate a "flipped" substrate binding orientation. In this way, changes at Phe247 complement those at the Tyr78/Ile113 couple. It is worth noting that the three key residues controlling stereo-selectivity lie close to the substrate, but others with less influence lie even closer (or alternatively, those lying closer cannot be changed without harming catalytic activity).

Key OYE 2.6 variants selected during the course of this study were purified as glutathione (S)-transferase fusion proteins and assayed for reductions of substrates 1 − 3 (Table 3). This allowed for consistent protein quantities that avoided artifacts due to differences in expression levels during whole-cell-catalyzed alkene reductions. In the case of cyclopentenone 2, there was clear interplay between the side chains at positions 78 and 113, and the two most R-selective mutants paired Trp with either Ile or Met. Substituting even similar residues at position

113 (e.g., Phe, Leu, or Val) significantly diminished the R selectivity. These observations parallel our earlier results from S. pastorianus OYE 1, in which very similar amino acid replacements at position 116 (which corresponds to position 113 of P. stipitis OYE 2.6) provoked highly divergent stereoselectivities.[18]

When all three substrates were considered together, the best overall performer was the Tyr78Trp/Phe247Ala variant. This protein gave essentially complete R selectivity for alkene 3, 91% ee favoring the R product from cyclopentenone 2, and the lowest observed S selectivity for cyclohexenone 1 (37% ee). That the two targeted positions lie on opposite sides of the active site and create additional site volume provides one explanation for their effects. From a preparative standpoint, the only deficiency of the Tyr78Trp/Phe247 Ala double mutant was its poor catalytic efficiency against alkene 3; however, this was remedied by the additional replacement of Ile113 by Cys or Val.

Interestingly, it was possible to create OYE 2.6 variants with altered stereoselectivity toward substrates 2 and 3, but not for 1. Given the apparent structural similarity between cyclic homologues 1 and 2, this was unexpected. One possible reason

emerged from a comparison of potential enzyme−substrate complexes for the three Baylis−Hillman adducts. Low-energy conformations for alkenes **1**−**3** (generated by MM2 molecular mechanics) were overlaid using the enone (C═O, C$_\alpha$ and C$_\beta$) and side chain methylene moieties, then the aggregate was modeled into the active site of wild-type OYE 2.6 in the "flipped" orientation by reference to the experimentally determined position of *p*-chlorophenol (Figure 1; PDB 4DF2). The side chain hydroxymethyl group was rotated to match its position in a crystal structure of **2** complexed to *S. pastorianus* OYE 1 (PDB 3RND). For comparison, the experimentally determined structure of the Tyr78Trp OYE 2.6 variant (PDB 4QAI) was also overlaid (Figure 10).
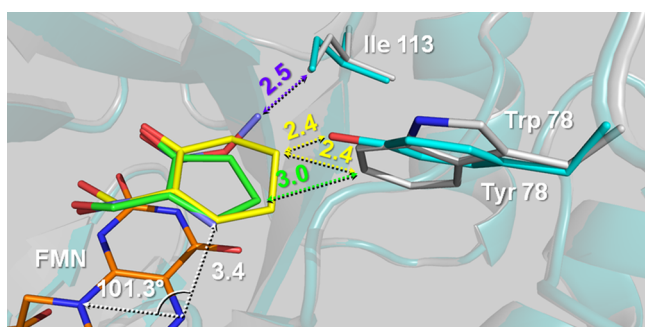
Figure 10. "Flipped" substrate binding in wild-type and Tyr78Trp OYE 2.6. Alkenes **1**, **2**, and **3** were overlaid as described in the text and modeled into the active site of wild-type OYE 2.6 (PDB code 4DF2) by aligning the substrate C$_\beta$ and carbonyl oxygens with the corresponding atoms of *p*-chlorophenol (substrate carbons are colored by substrate: **1**, yellow; **2**, green; **3**, blue). Experimentally determined coordinates for wild-type and Tyr78Trp OYE 2.6 (PDB codes 4DF2 and 4QAI; carbons shown in light blue and gray, respectively) were overlaid using the bound FMN moieties, and side chains of residues at positions 78 and 113 are shown in stick form. Key interatomic distances (in Å) are indicated along with the angle for hydride transfer from N$_5$ of reduced FMN (101.3°).

One key difference between **1** and **2** is that the former occupies a half-chair conformation and the latter is nearly planar (Figure 10). This also means that a CH$_2$ moiety in **1** projects above or below the plane of the ring, bringing it very close to the side chain at position 78.[27] Although the precise distances depicted in Figure 10 are subject to uncertainties, it is clear that the out-of-plane CH$_2$ group of **1** makes a significantly closer approach to the side chain at position 78 than does alkene **2**. Moreover, although the overall size of the Trp side chain is larger than that of Tyr, the latter more closely approaches the cyclopentenyl ring of **2** (Figure 10). This might explain why the "flipped" substrate binding orientation is not observed when **2** is reduced by wild-type OYE 2.6 but can occur when Trp replaces Tyr78. That both a Tyr and Trp side chain at position 78 appear to be too large to allow binding of **1** in a "flipped" orientation may explain why an alternative binding mode is not observed for this alkene.

The major surprise in evolving a catalyst for *S*-selective reduction of Roche ester precursor **3** was that all useful variants were forced to pass through an intermediate stage devoid of catalytic activity (Tyr78Trp).

Conclusions. Several lessons emerged from this study. First, when high-quality randomized libraries are available, even protein engineering projects requiring relatively lengthy analyses such as chiral-phase GC are feasible. In addition,

although side chain proximity to the substrate can provide a rough guide to importance in controlling binding interactions, one must experimentally interrogate each position to establish its role. Like its counterpart in *S. pastorianus* OYE 1 (Trp116),[9] Ile113 in OYE 2.6 plays a crucial role in controlling substrate binding; however, its influence is not felt fully until its neighbor (Tyr78) is altered. This behavior stands in contrast to OYE 1, where substitutions for Trp116 have significant impacts even in a wild-type background. The interplay between substitutions at positions 78 and 113 highlights the nonadditivity of mutational effects that can occur during protein engineering studies. In such cases, relying on successive rounds of site-saturation mutagenesis at single positions may not always yield the most improved multiply mutated variant(s). On the other hand, the potential benefits of simultaneously substituting several amino acids within single libraries must also be balanced against the much larger screening efforts required to interrogate the resulting libraries. [10]

With regard to the questions posed at the start of the project, the "best" evolutionary pathways for two of the substrates were similar, and common changes were appropriate for both (Figures 5 and 8). For example, in a Tyr78Trp background, substitutions of Phe247 yielded Ala and His as the best replacements for both alkenes **2** and **3**. Since both screening efforts were carried out independently, we believe that this indicates that sequence space has been completely sampled in this region. As a second example, random replacements of Ile113 with pre-existing mutations at Tyr78Trp, Tyr78Trp/Phe247Ala, or Tyr78Trp/Phe247His all yielded Ile113Cys and Ile113Val as the optimal replacements for alkene **3**. That all three starting proteins yielded the same Ile113 replacements strongly suggests that the system has converged and that Cys and Val are the "best" residue at this position (at least in the context of the substrates studied).

The general utility of the final OYE 2.6 variants remains a somewhat open question. Although a subset provided good conversion efficiency and reversed stereoselectivity for alkenes **2** and **3** (Table 3), this was not the case for cyclohexenone **1**. We are currently studying the reductions of other alkenes with the variants listed in Table 3 as a means of establishing their general utility.

The final lesson from these studies is that one must sometimes cross a "desert" of sequence space before re-establishing catalytic activity and seeking better variants. This was clearly the case for Roche ester precursor **3**, and it was only by pursuing three parallel protein engineering studies that a solution was found for this synthetically valuable substrate.

■ EXPERIMENTAL SECTION

General. Restriction endonucleases, Phusion Hot Start II High-Fidelity DNA Polymerase, restriction enzyme *Dpn*I and T4 DNA ligase were purchased from New England Biolabs. Primers were obtained from Integrated DNA Technologies. *Escherichia coli* ElectroTen-Blue and BL21-Gold (DE3) strains were purchased from Agilent Technologies, USA. Plasmids were purified on small scales by EconoSpin columns (Epoch Life Sciences) and on large scales using CsCl density gradient ultracentrifugation.[28] Plasmids were transformed into *E. coli* cells by a Bio-Rad Gene Pulser apparatus using the manufacturer's protocol. DNA sequencing was carried out by the University of Florida ICBR using capillary fluorescence methods and employing either standard protocols (single samples or pooled plasmid samples) or rolling circle

amplification (96 well bacterial culture plates). Plasmid pBS2, which encodes wild-type *P. stipitis* OYE 2.6, has been described elsewhere.[11] LB medium contained 5 g/L Bacto-Yeast extract, 10 g/L Bacto-Tryptone, and 10 g/L NaCl. When required, ampicillin was included at 200 μg/mL. ZYP-5052 autoinducing media contained 10 g/L tryptone, 5 g/L yeast extract, 1 mM MgSO$_4$, 25 mM (NH$_4$)$_2$SO$_4$, 50 mM KH$_2$PO$_4$, 50 mM Na$_2$HPO$_4$, 5 g/L glycerol, 0.5 g/L anhydrous glucose, and 2 g/L α-lactose monohydrate.[29] All protein isolation steps were carried out at 4 °C unless otherwise indicated. Glucose dehydrogenase (GDH102) was a generous gift from Bio-Catalytics.

Alkenes **1**−**3** and authentic standards for their reduction products were prepared as described previously.[9] All other reagents were obtained from commercial suppliers and were used as received unless otherwise indicated. Chiral-phase GC analyses were carried out with a 30 m × 0.25 mm β-Dex 225 column (Supelco) with He as the carrier gas and FID. For analysis of **1** and **2** and their reduction products, the temperature program involved 140 °C (10 min) followed by a 20 °C/min increase to 180 °C (5 min). Under these conditions, the *S* and *R* reduction products from **1** eluted at 10.2 and 10.8 min, respectively, and **1** was observed at 13.1 min. The *S* and *R* reduction products from **2** eluted at 11.3 and 10.2 min, respectively, and **2** was observed at 13.4 min.[30] For analysis of **3** and its reduction products, the temperature program involved 100 °C (12 min), followed by a 20 °C/min increase to 180 °C (5 min). Under these conditions, the *S* and *R* reduction products from **3** eluted at 10.7 and 11.3 min, respectively, and substrate **3** was observed at 11.8 min.

Library Construction. Details of the library construction methodology have been published elsewhere.[15] A list of primers used to create first-generation libraries beyond those reported in ref 15 is shown in Table 4. Second- and third-generation

Table 4. Sequences of Primers Used in This Study Not Reported Previously[15]

library	wt codon		primer sequence[a]
A68X	GCT	for	TT ATC ACT GAA **NNK** ACT TTT GTC TCT CCT CAA GCC TCT GG
		rev	GA GAC AAA AGT **MNN** TTC AGT GAT AAG CAA AGT ACC TGG GA
Y78X	TAT	for	AA GCC TCT GGT **NNK** GAA GGT GCT GCT CCA GGT ATT TGG AC
		rev	GC AGC ACC TTC **MNN** ACC AGA GGC TTG AGG AGA GAC AAA AG
G292	GGT	for	CT CGT GTA AGT **NNK** AAC GTC GAC GTC TCT GAA GAA GAC CA
		rev	AC GTC GAC GTT **MNN** ACT TAC ACG AGG CTC AAC GAC AGA AA

[a]N = A, C, G, T (equimolar amounts); K = G, T (equimolar amounts); M = A, C (equimolar amounts).

libraries were prepared by the same methods, with the appropriate mutant OYE 2.6 gene replacing the wild-type as the template in PCR amplifications. The pooled degenerate plasmid sample (1 ng) was used to transform 40 μL of electrocompetent BL21 Gold (DE3) cells. After recovery for 1 h at 37 °C in 600 μL of SOC medium, cells were spread on LB agar plates containing 200 μg/mL ampicillin and incubated overnight at 37 °C. Ninety five individual transformants (plus a control, wild-type for first-generation libraries or the starting mutant for second- and third-generation libraries) were used to

seed a 2 mL deepwell plate containing LB medium supplemented with 200 μg/mL ampicillin. For some libraries, all 96 plasmids were prepared by rolling circle amplification and sequenced. On the basis of the sequencing data, clones representing each of the 32 possible codons (when available) were used to seed 600 μL of LB media containing 200 μg/mL ampicillin arrayed on one-third of a deepwell plate (2 mL total volume per well) to yield library master plates. The initially prepared first-generation library of Ile113 random replacements was missing three members (Asp, Lys, and Phe); these were prepared individually and added to yield a complete set of amino acid replacements. For those libraries not sequenced prior to screening, catalytic properties of all 96 clones were evaluated, and only members that showed interesting properties were analyzed by DNA sequencing ("blind screening").

Library Screening. Library screening plates were prepared from library master plates, either freshly grown in LB medium supplemented with 200 μg/mL ampicillin or from frozen stocks. In the former case, the library master plate was shaken at 200 rpm and 37 °C. After 6 h, the cultures were visibly turbid, and a 20 μL aliquot from each well was used to inoculate wells in a duplicate deepwell plate containing 600 μL of ZYP-5052 autoinducing medium supplemented with 200 μg/mL ampicillin. Alternatively, the autoinduction plates were inoculated directly from frozen library master plates. After seeding, library screening plates were then mounted in a growth apparatus designed to facilitate maximal oxygen transfer rates (modified from the design of Duetz and Witholt[31,32]) and shaken at 300 rpm and 37 °C. After 16−18 h, cells were harvested by centrifuging at 3000 rpm for 30 min, and the supernatant was removed by aspiration. Cell pellets were resuspended in 300 μL of 50 mM KP$_i$, 100 mM glucose, 10 mM substrate, pH 7.0. The plate was shaken at 200 rpm at room temp. After 6 h, wells were individually extracted with 500 μL of EtOAc prior to chiral-phase GC analysis.

Isolation of GST-Fusion Proteins. Overexpression plasmids for each alkene reductase (produced as GST-fusion proteins) were separately used to transform *E. coli* BL21 Gold (DE3). Overnight cultures grown at 37 °C in LB medium supplemented with ampicillin were diluted 1:100 into 500 mL of the same medium and shaken at 37 °C until the O.D.$_{600}$ value was between 0.5 and 1.0. Alkene reductase overexpression was induced by adding IPTG (to a final concentration of 100 μM) and glucose (at a final concentration of 4 g/L). The cultures were shaken at room temperature for 4 h prior to collecting the cells by centrifugation. Cell pellets were resuspended in 10 mL of cold buffer (50 mM Tris-Cl, 1 mM PMSF, pH 7.5) and lysed using a French pressure cell (10 000−20 000 psi). Insoluble debris was removed by centrifuging at 15 000g for 60 min at 4 °C. The resulting supernatant was passed through 10 mL of glutathione agarose (Clontech) using an FPLC system (Pharmacia) with 1× PBS buffer as the mobile phase. Once the A_{280} returned to a baseline reading, the desired protein was eluted by adding 10 mL of reduced glutathione (10 mM, freshly prepared) in Tris−HCl (50 mM, pH 8.0), and then the appropriate fractions were concentrated to ~20−40 mg/mL by ultrafiltration (Amicon Ultra-4 membrane, 10 000 NMWL). An equal volume of glycerol was added prior to storage at −20 °C.

Isolation of Nontagged Alkene Reductases for X-ray Crystallography. Overexpression plasmids encoding a mutant alkene reductase (Tyr78Trp OYE 2.6 and Tyr78Trp/Ile113Cys OYE 2.6) were separately used to transform *E. coli* BL21 Gold

(DE3). Cells were grown and induced using the procedure described in the previous section. Cell pellets were resuspended in 100 mM Tris−HCl, 100 mM $(NH_4)_2SO_4$, 1 mM PMSF, pH 8.0, and lysed using a French pressure cell (10 000−20 000 psi). Insoluble debris was removed by centrifuging at 15 000g for 60 min at 4 °C. DNA was precipitated from the supernatant by the addition of 0.1 volume of protamine sulfate (10 mg/mL stock solution) followed by stirring for 10 min and centrifuging at 15 000g for 15 min at 4 °C. Solid $(NH_4)_2SO_4$ was added to the supernatant until 78% saturation, then the pellet was collected by centrifuging at 15 000g for 30 min at 4 °C. The resulting pellet was resuspended in a minimal volume (7−10 mL) of buffer (100 mM Tris−HCl, 100 mM $(NH_4)_2SO_4$, pH 8.0), then dialyzed against two changes of the same buffer over a total of 8 h. A final dialysis was performed against the same buffer supplemented with 10 mM $Na_2S_2O_4$ over 4 h.[33] The final dialysis step (8 h) was performed with the omission of the $Na_2S_2O_4$. The resulting bright yellow dialysate was loaded onto a 3 mL N-(4-hydroxybenzoyl)aminohexyl agarose affinity resin[34] previously equilibrated with starting buffer (100 mM Tris−HCl, 100 mM $(NH_4)_2SO_4$, pH 8.0) using an FPLC system (Pharmacia). The resin turned dark brown, indicating successful OYE 2.6 binding. The column was washed with starting buffer until the A_{280} value returned to its baseline reading, then the desired protein was eluted by adding 3 resin volumes of deoxygenated starting buffer supplemented with 10 mM $Na_2S_2O_4$ (degassed under vacuum for 20 min and flushed with N_2 prior to adding the reducing agent).[35] The appropriate fractions were concentrated to ~20−40 mg/mL by ultra-filtration (Amicon Ultra-4 membrane, 10 000 NMWL). Final purification was achieved by gel filtration chromatography using a Superdex 200 (Pharmacia) column equilibrated with 30 mM Tris−HCl, 30 mM NaCl, pH7.5. The eluted protein was concentrated to ~30−40 mg/mL prior to use in crystallization studies.

Crystallogenesis. Crystals were grown using the sitting drop vapor diffusion method from 2.4 M sodium malonate pH 7.0. After growth, the crystals were mounted into appropriately sized loops and transferred to a harvesting buffer (3.0 M sodium malonate, 10% glycerol v/v, pH 7.0) supplemented with a saturated amount of substrate prior to flash cooling in liquid N_2.

Data Collection and Structure Solution. Diffraction data were collected using synchrotron radiation under cryogenic conditions. Reflection data for the Tyr78Trp OYE 2.6 were processed by an in-house program to yield a merged set suitable for refinement; those for the Tyr78Trp/Ile113Cys OYE 2.6 were processed using the HKL3000 program suite.[36] Phases were obtained by molecular replacement using the AutoMR utility of PHENIX[37] using the wild-type *P. stipitis* OYE 2.6 (PDB code 3TJL) as the search model (after removing FMN, water molecules, and nonidentical side chains). The initially calculated $2F_o−F_c$ and $F_o−F_c$ maps showed easily identifiable electron density patterns for the FMN, which served as a validation for the molecular replacement solution. Iterative cycles of refinement and model building, with PHENIX and COOT,[38] respectively, were performed until the error statistics reached acceptable levels. At this stage of refinement, the active site contained areas of unaccounted-for electron density. Potential ligand structures were built using proDrg[39] and then incorporated into the model for further refinement. All protein structure figures were created using PyMOL (Schrödinger, LLC).

Alkene Reductions by Purified Proteins. Reaction mixtures (0.30 mL total volume) contained wild-type or a mutant OYE 2.6 (~100 μg), glucose dehydrogenase (75 μg), $NADP^+$ (0.20 mM), glucose (220 mM), and alkene substrate **1**, **2**, or **3** (10 mM) in KP_i, pH 7.0. Reactions were incubated at room temp, then extracted with an equal quantity of EtOAc and analyzed by chiral-phase GC.

■ AUTHOR INFORMATION
Corresponding Author
*Phone, Fax: 352.846.0743. E-mail: jds2@chem.ufl.edu.
Notes
The authors declare no competing financial interest.

■ REFERENCES
(1) (a) Swiderska, M. A.; Stewart, J. D. *J. Mol. Catal. B: Enzym.* **2006**, *42*, 52−54. (b) Chaparro-Riggers, J. F.; Rogers, T. A.; Vazquez-Figueroa, E.; Polizzi, K. M.; Bommarius, A. S. *Adv. Synth. Catal.* **2007**, *349*, 1521−1531. (c) Hall, M.; Stueckler, C.; Kroutil, W.; Macheroux, P.; Faber, K. *Angew. Chem., Int. Ed.* **2007**, *46*, 3934−3937. (d) Stueckler, C.; Hall, M.; Ehammer, H.; Pointner, E.; Kroutil, W.; Macheroux, P.; Faber, K. *Org. Lett.* **2007**, *9*, 5409−5411. (e) Stuermer, R.; Hauer, B.; Hall, M.; Faber, K. *Curr. Opin. Chem. Biol.* **2007**, *11*, 203−213. (f) Hall, M.; Stueckler, C.; Ehammer, H.; Pointner, E.; Oberdorfer, G.; Gruber, K.; Hauer, B.; Stuermer, R.; Kroutil, W.; Macheroux, P.; Faber, K. *Adv. Synth. Catal.* **2008**, *350*, 411−418. (g) Hall, M.; Stueckler, C.; Hauer, B.; Stuermer, R.; Friedrich, T.; Breuer, M.; Kroutil, W.; Faber, K. *Eur. J. Org. Chem.* **2008**, 1511−1516. (h) Kosjek, B.; Fleitz, F. J.; Dormer, P. G.; Kuethe, J. T.; Devine, P. N. *Tetrahedron: Asymmetry* **2008**, *19*, 1403−1406. (i) Schweiger, P.; Gross, H.; Wesener, S.; Deppenmeier, U. *Appl. Microbiol. Biotechnol.* **2008**, *80*, 995−1006. (j) Adalbjörnsson, B. V.; Toogood, H.; Fryszkowska, A.; Pudney, C. R.; Jowitt, T. A.; Leys, D.; Scrutton, N. S. *ChemBioChem.* **2010**, *11*, 197−207. (k) Bougioukou, D. J.; Kille, S.; Taglieber, A.; Reetz, M. T. *Adv. Synth. Catal.* **2009**, *351*, 3287−3305. (l) Burda, E.; Kraußer, M.; Fischer, G.; Hummel, W.; Müller-Uri, F.; Kreis, W.; Gröger, H. *Adv. Synth. Catal.* **2009**, *351*, 2787−2790. (m) Hirata, T.; Matsushima, A.; Sato, Y.; Iwasaki, T.; Nomura, H.; Watanabe, T.; Toyoda, S.; Izumi, S. *J. Mol. Catal. B: Enzym.* **2009**, *59*, 158−162. (n) Winkler, C. K.; Tasnádi, G.; Clay, D.; Hall, M.; Faber, K. *J. Biotechnol.* **2012**, *162*, 381−389.
(2) Formerly known as *Saccharomyces carlsbergensis*.
(3) Vaz, A. D. N.; Chakraborty, S.; Massey, V. *Biochemistry* **1995**, *34*, 4246−4256.
(4) (a) Bougioukou, D. J. Ph.D. Thesis, University of Florida, Gainesville, FL, 2006. (b) Bougioukou, D. J.; Stewart, J. D. *J. Am. Chem. Soc.* **2008**, *130*, 7655−7658. (c) Oberdorfer, G.; Gruber, K.; Faber, K.; Hall, M. *Synlett* **2012**, *23*, 1857−1864.
(5) (a) Padhi, S. K.; Bougioukou, D. J.; Stewart, J. D. *J. Am. Chem. Soc.* **2009**, *131*, 3271−3280. (b) Walton, A. Z. Ph.D. Thesis, University of Florida, Gainesville, FL, 2012.

(6) (a) Reich, S.; Hoeffken, H. W.; Rosche, B.; Nestl, B. M.; Hauer, B. *ChemBioChem.* **2012**, *13*, 2400−2407. (b) Daugherty, A. B.; Govindarajan, S.; Lutz, S. *J. Am. Chem. Soc.* **2013**, *135*, 14425−14432.

(7) Gumulya, Y.; Sanchis, J.; Reetz, M. T. *ChemBioChem.* **2012**, *13*, 1060−1066.

(8) Bougioukou, D. J.; Walton, A. Z.; Stewart, J. D. *Chem. Commun.* **2010**, *46*, 8558−8560.

(9) Walton, A. Z.; Conerly, W. C.; Pompeu, Y.; Sullivan, B.; Stewart, J. D. *ACS Catal.* **2011**, *1*, 989−993.

(10) (a) Reetz, M. T. *Angew. Chem., Int. Ed.* **2011**, *50*, 138−174. (b) Reetz, M. T.; Krebs, G. P. L. *C. R. Chim.* **2011**, *14*, 811−818.

(11) Pompeu, Y. A.; Sullivan, B.; Walton, A. Z.; Stewart, J. D. *Adv. Synth. Catal.* **2012**, *354*, 1949−1960.

(12) Thr35, Phe37, Ala68, Tyr78, Ile113, Leu115, His188, His191, Tyr193, Phe247, Gln248, Asn293, Val294, Phe373, and Tyr374.

(13) No C_β/inhibitor distance could be measured for the residue at position 292 because it is occupied by Gly in wild-type OYE 2.6.

(14) N = any base; K = G or T.

(15) Sullivan, B.; Walton, A. Z.; Stewart, J. D. *Enzyme Microb. Technol.* **2013**, *53*, 70−77.

(16) Byproducts in which one or more copies of the PCR primers become incorporated. Because this inserts additional amino acids within the protein, such mutants are highly unlikely to retain function.

(17) Because 32 codons are possible in the NNK scheme, this allowed three different randomized positions to be examined on a single 96 well "Superplate".

(18) Pompeu, Y. A.; Sullivan, B.; Stewart, J. D. *ACS Catal.* **2013**, *3*, 2376−2390.

(19) A complete set of 32 NNK codons was screened to ensure complete coverage at position 78.

(20) In all cases, positive controls showed very high conversions, which ruled out experimental artifacts as the explanation for the lack of catalytic activities in these libraries.

(21) To retain consistency with earlier studies, no disulfide reducing agents were included in the purification or crystallization buffers for the Tyr78Trp/Ile113Cys double mutant.

(22) (*S*)-Oxycysteine, PDB residue type CSX; 3-sulfinoalanine, PDB residue type CSD; cysteine-(*S*)-dioxide, PDB residue type CSW; cysteinesulfonic acid, PDB residue type OCS.

(23) Attempts to explain the additional electron density by a second malonate in a different location were unsuccessful.

(24) The carbonyl oxygen of 3 was initially placed at the same location as the phenolic oxygen of *p*-chlorophenol bound in the active site of OYE 2.6 (a structure solved in an earlier study; PDB code 4DF2).

(25) Fraaije, M. W.; Mattevi, A. *Trends Biochem. Sci.* **2000**, *25*, 126−132.

(26) Although the side chain was modeled as a single conformation, the somewhat weak electron density in this location suggests that other rotamers may coexist within the crystals.

(27) The alternative half-chair conformation of **1** cannot be accommodated because the out-of-plane CH_2 group would collide with the FMN cofactor (the distance would be ~1.4 A).

(28) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning. A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.

(29) Studier, F. W. *Protein Express. Purif.* **2005**, *41*, 207−234.

(30) Note that the order of elution is opposite for reduction products derived from substrates **1** and **2**.

(31) Duetz, W. A.; Witholt, B. *Biochem. Eng. J.* **2001**, *7*, 113−115.

(32) Details of the growth apparatus construction can be found in the Ph.D. thesis of A.Z.W., University of Florida, 2012.

(33) This step, referred to as "degreening", removes low-molecular-weight ligands from the enzyme active site that would otherwise interfere with binding to the affinity column.

(34) Abramovitz, A. S.; Massey, V. *J. Biol. Chem.* **1976**, *251*, 5321−5326.

(35) The elution buffer was prepared immediately before use.

(36) Otwinowski, Z.; Minor, W. *Meth. Enzymol.* **1997**, *276*, 307−326.

(37) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr., Sect. D* **2010**, *66*, 213−221.

(38) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D* **2010**, *66*, 486−501.

(39) Schuttelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallogr., Sect. D* **2004**, *60*, 1355−1363.

(40) *The PyMOL Molecular Graphics System, Version 1.5.0.4*; Schrödinger, LLC.

(41) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D* **2010**, *66*, 12−21.