Structure-based switch of regioselectivity in the flavin-dependent tryptophan 6-halogenase Thal

Ann-Christin Moritzer[‡1], Hannah Minges[§1], Tina Prior[‡], Marcel Frese[§], ● Norbert Sewald[§], and ● Hartmut H. Niemann[‡2]

From the ‡Structural Biochemistry and §Organic and Bioorganic Chemistry Research Groups, Department of Chemistry, Bielefeld University, Universitätsstrasse 25, D-33615 Bielefeld, Germany

Edited by Ruma Banerjee

Flavin-dependent halogenases increasingly attract attention as biocatalysts in organic synthesis, facilitating environmentally friendly halogenation strategies that require only FADH$_2$, oxygen, and halide salts. Different flavin-dependent tryptophan halogenases regioselectively chlorinate or brominate tryptophan's indole moiety at C5, C6, or C7. Here, we present the first substrate-bound structure of a tryptophan 6-halogenase, namely Thal, also known as ThdH, from the bacterium *Streptomyces albogriseolus* at 2.55 Å resolution. The structure revealed that the C6 of tryptophan is positioned next to the ε-amino group of a conserved lysine, confirming the hypothesis that proximity to the catalytic residue determines the site of electrophilic aromatic substitution. Although Thal is more similar in sequence and structure to the tryptophan 7-halogenase RebH than to the tryptophan 5-halogenase PyrH, the indole binding pose in the Thal active site more closely resembled that of PyrH than that of RebH. The difference in indole orientation between Thal and RebH appeared to be largely governed by residues positioning the Trp backbone atoms. The sequences of Thal and RebH lining the substrate binding site differ in only few residues. Therefore, we exchanged five amino acids in the Thal active site with the corresponding counterparts in RebH, generating the quintuple variant Thal-RebH5. Overall conversion of L-Trp by the Thal-RebH5 variant resembled that of WT Thal, but its regioselectivity of chlorination and bromination was almost completely switched from C6 to C7 as in RebH. We conclude that structure-based protein engineering with targeted substitution of a few residues is an efficient approach to tailoring flavin-dependent halogenases.

Whereas chemical halogenation of aromatic compounds requires toxic elemental halogens and Lewis acids as catalysts, flavin-dependent halogenases (FDHs)[3] can catalyze electrophilic aromatic substitutions with innoxious halide salts and O$_2$

as substrates and FADH$_2$ as co-factor (1, 2). FDHs are responsible for the regioselective chlorination or bromination of a vast variety of natural products (3) and have attracted a lot of attention due to their potential use as biocatalysts (2, 4, 5). Slow turnover and low stability still limit their large-scale application, but various strategies have been pursued to tackle these challenges (6–8). Bacterial tryptophan halogenases were the first FDHs to be described (9–11). In addition, FDHs were identified that halogenate substituted tryptophan (12, 13) or chlorinate tryptophan only when it is part of a substrate oligopeptide (14). FDHs acting on free substrates other than tryptophan were also described (15–18), including fungal enzymes that accept large substrates (19–21) or have rather broad substrate specificity (22, 23), raising the hope that diverse compounds can be enzymatically halogenated *in vitro*. Moreover, FDHs have been integrated into engineered biosynthetic pathways *in vivo* (24, 25). The first structure of an FDH showed that these enzymes consist of two subdomains, a conserved box-shaped subdomain harboring the FAD binding site and a variable subdomain that is pyramidal in tryptophan halogenases (26). Tryptophan is bound at the interface of the box and the pyramid.

In addition to using benign reactants under mild conditions, the regioselectivity of FDHs provides a further advantage of enzymatic halogenation in contrast to chemical approaches. Different FDHs are known that halogenate tryptophan selectively on C5, C6, or C7 of the indole moiety, whereas chemical halogenation preferably occurs at C2 (27). Thus, for the natural substrate L-Trp, the site of enzymatic halogenation by flavin-dependent tryptophan halogenases is not governed by electronic factors of the substrate but determined sterically by the orientation of the substrate in the active site. This model was suggested based on the structure of the tryptophan 7-halogenase PrnA (26) and confirmed by the structure of the tryptophan 5-halogenase PyrH (28), both in complex with their native substrate tryptophan. These structures showed that the carbon atom to be halogenated is positioned most closely to the ε-amino group of the catalytic lysine, which is conserved among FDHs. To achieve this positioning, tryptophan is bound with distinctly different poses by tryptophan 7-halogenases and by tryptophan 5-halogenases. The indole moiety of L-Trp bound to PyrH is flipped by about 180° relative to the orientation observed in the tryptophan 7-halogenases PrnA and RebH. Several FDHs have been described that halogenate tryptophan

on C6, namely Thal (29) (also known as ThdH (30)), SttH (31), and thermophilic Th-Hal (32), which are *bona fide* tryptophan 6-halogenases, and KtzR, which produces 6-chlorotryptophan from tryptophan but has higher catalytic efficiency for converting 7-Cl-L-Trp to 6,7-di-Cl-L-Trp (12). BorH is another putative tryptophan 6-halogenase; however, its activity has not yet been confirmed *in vitro* (33). Based on sequence identity, SttH, Th-Hal, and KtzR are closely related to each other and to the tryptophan 5-halogenase PyrH, whereas on the other hand, Thal and BorH are closely related to each other and to the tryptophan 7-halogenases RebH and PrnA (Table 1). Crystal structures of SttH (34) and Th-Hal (32) have been published, but only in the absence of substrate. No structural data are available for Thal, BorH, or KtzR. Modeling the substrate into the active site of the empty enzyme is hampered by protein conformational changes that occur upon substrate binding (28, 35, 36). One would expect that in tryptophan 6-halogenases C6 of the indole moiety occupies a position roughly equivalent to the position of C7 or C5 in tryptophan 7-halogenases or 5-halogenases, respectively. This positioning of C6 could be achieved by slightly rotating the indole moiety, starting either from the binding pose observed in the tryptophan 7-halogenases PrnA and RebH or from the flipped binding pose observed in the tryptophan 5-halogenase PyrH (37). Thus, it remains elusive how L-Trp binds to the active site of tryptophan 6-halogenases.

With their application as biocatalysts in mind, enzyme optimization has been performed on FDHs, aiming at increasing their lifetime and changing their substrate scope or regioselectivity (1, 2). A large amount of work has gone into both structure-based engineering and directed evolution (27). These efforts resulted in enzymes with higher stability (7, 8), higher activity (23, 38), different substrate scope (24, 37, 39, 40), or altered regioselectivity (34, 37, 40, 41). Despite the already achieved success, these tasks still pose formidable challenges. Directed evolution requires many rounds of high-throughput screening, and frequently it remains unclear how beneficial mutations exert their positive effect (37). Attempts to alter the regioselectivity with tryptophan as substrate have so far resulted in an incomplete change of selectivity so that only one-third of substrate was converted to a new regioisomer, whereas the majority was still halogenated on the initially preferred position (34, 41).

We embarked on a project to further elucidate the structural basis of regioselectivity of tryptophan halogenases. We aimed at crystallizing the tryptophan 6-halogenase Thal together with its native substrate L-Trp. Based on its structure, we designed a Thal variant with five amino acid substitutions in the active site with the goal to switch the regioselectivity from the 6- to the 7-position.
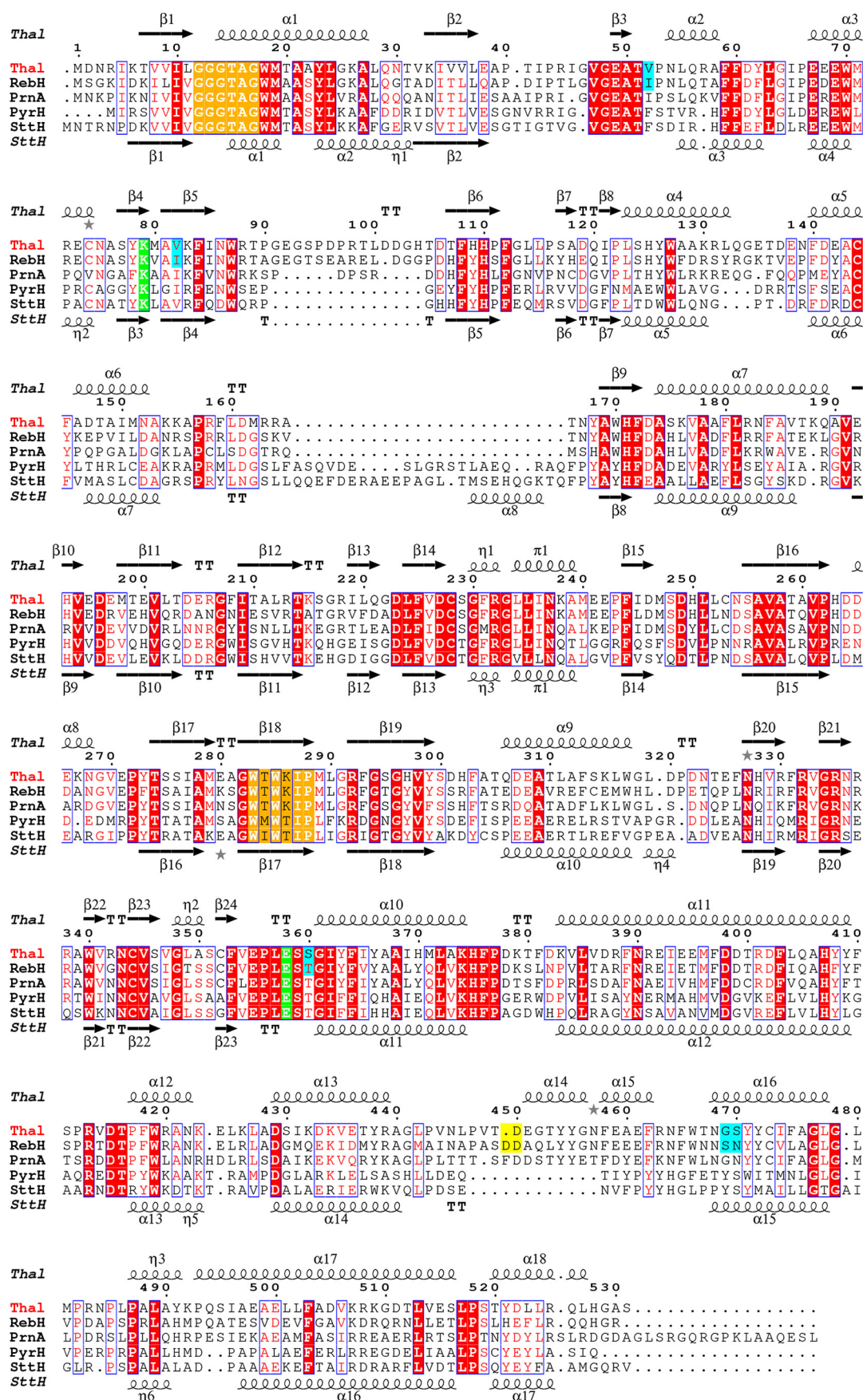
Results

Overall structure

The crystal structure of Thal without FAD, chloride, and tryptophan (apo-Thal) and a structure in complex with its substrate L-Trp (Trp-Thal) obtained by soaking contained two

copies of Thal per asymmetric unit (Table S1). Amino acids (aa) 2–529 could be modeled in all chains except for chain B of the apo-Thal structure, which lacked interpretable electron density for residues 450–456. In the apo-Thal structure, the active site of chain A contains positive difference density, which we could not faithfully model. This difference density suggests that the substrate binding site of apo-Thal chain A is not truly empty. Clear electron density was observed for the substrate L-Trp in both chains of Trp-Thal.

In solution, Thal behaved as a monomer according to the elution volume from size-exclusion chromatography and its hydrodynamic radius as determined by dynamic light scattering (data not shown). Nevertheless, Thal appears to form a dimer in the crystal (Fig. S1) like other tryptophan halogenases (26, 28, 32, 34–36). The arrangement of protomers in the dimer is very similar to that of other tryptophan halogenases with a root mean square deviation (r.m.s.d.) between the apo-Thal dimer and the apo-RebH (PDB code 2OAM) dimer of only 1.4426 Å for 1027 structurally corresponding residues.

The structural similarity is even higher for the protomers with an r.m.s.d. of 0.9675 and 0.9013 Å for 521 and 519 structurally corresponding residues, respectively. This is not surprising, given the sequence identity between Thal and RebH of 64% with almost no insertions or deletions, except for Asp-101 and the C-terminal Ser-531 that are inserted in Thal and Asp-449 inserted in RebH (Fig. 1). Each monomer consists of a single domain, which folds into a larger rectangular box containing the FAD binding site, and a smaller triangular pyramid (Fig. 2), as described for PrnA (26) and other tryptophan halogenases (28, 32, 34–36). The substrate binding site is located between the box and the pyramid and occupies the same position as in RebH (Fig. S2a). Thal residues Lys-79 and Glu-358 align with the catalytic key residues Lys-79 and Glu-357 of RebH. The difference from the catalytic residues Lys-83 and Glu-362 previously reported for Thal (29) is due to a difference in numbering with an offset of 4. At the level of the backbone, just a few structural differences between Thal and RebH become apparent, mainly at the sites of insertions (Fig. S2a).

Sequence conservation between Thal and the tryptophan 5-halogenase PyrH or the tryptophan 6-halogenases Th-Hal and SttH is lower than between Thal and RebH (Table 1). Accordingly, structural differences between Thal and PyrH, Th-Hal or SttH are bigger than between Thal and RebH. As a L-Trp–bound structure is available only for PyrH, we will focus on a structural comparison of Thal with PyrH rather than SttH or Th-Hal. Thal (PDB code 6H44) and PyrH (PDB code 2WEU) protomers align with only 466 structurally corresponding residues and an r.m.s.d. value of 1.22 Å (Fig. S2b). For the dimer, the r.m.s.d. between Thal and PyrH is even higher with 1.73 Å for 928 structurally corresponding residues instead of the total of 1056 possible residues. Structural conservation is highest in the box-shaped FAD binding module. In the pyramidal subdomain, some secondary structure elements are rearranged or connected differently, which is mostly due to insertions or deletions (Fig. 1).
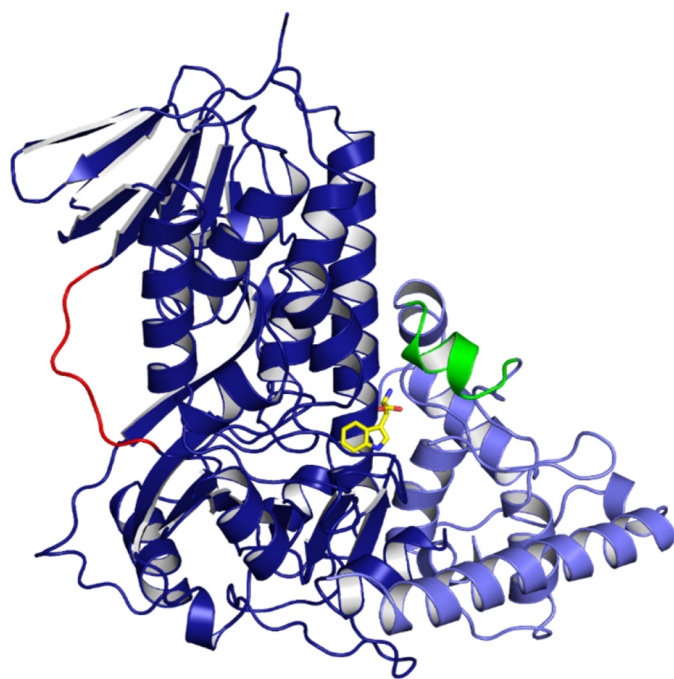
```
                    β1              α1                  β2            β3       α2          α3
Thal                →       oooooooooooo              →           →       ooooo       ooooo
        1         10              20            30          40       50        60          70
Thal    .MDNRIKTVVILGGGTAGWMTAAYLGKALQNTVKIVVLEAP.TIPRIGVGEATVPNLQRAFFDYLGIPEEEWM
RebH    .MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAP.DIPTLGVGEATIPNLQTAFFDFLGIPEDEWM
PrnA    .MNKPIKNIVIVGGGTAGWMAASYILVRALQQANITLIESAAIPRI.GVGEATIPSLQKVFFDFLGIPEREWM
PyrH    ....MIRSVVIVGGGTAGWMTASYLKAAFDDRIDVTLVESGNVRRIG.VGEATFSTVR.HFFDYLGLDEREWL
SttH    MNTRNPDKVVIVGGGTAGWMTASYLKKAFGERVSVTLVESGTIGTVG.VGEATFSDIR.HFFEFLDLREEEWM
SttH    →          ooooo        ooooo η1  β2                        oo ooooo    ooooo
        β1         α1           α2        β2                          α3      α4


                β4    β5               TT           β6      β7  β8   α4              α5
Thal  ooo  ★         →   →                          →       →   →   oooooo           ooo
        80              90          100         110      120      130            140
Thal    RECNASYKMAVKFINWRTPGEGSPDPRTLDDGHTDTFHHPFGLLPSADQIPLSHYWAAKRLQGETDENFDEAC
RebH    RECNASYKVAIKFINWRTAGEGTSEAREL.DGGPDHFYHSFGLLKYHEQIPLSHYWFDRSYRGKTVEPFDYAC
PrnA    PQVNGAFKAAIKFVNWRKSP.....DPSR....DDHEYHLFGNVPNCDGVPLTHYWLRKREQG.FQQPMEYAC
PyrH    PRCAGGYKLGIRFENWSEP........GEYEYHPFERLRVVDGFNMAEWWLAVG...DRRTSFSEAC
SttH    PACNATYKLAVRFQDWQRP.........GHHFYHPFEQMRSVDGFPLTDWWLQNG...PT.DRFDRDC
SttH    ooo  →          T..............T              →      TT oooooooo     ooooo
        η2   β3    β4                               β5       β6 β7    α5        α6


               α6               TT                                  β9     α7
Thal        ooooooo                                                →    ooooooooooo   -
        150              160                                      170        180        190
Thal    FADTAIMNAKKAPRFLDMRRA...............................TNYAWHFDASKVAAFLRNFAVTKQAVE
RebH    YKEPVILDANRSPRRLDGSKV...............................TNYAWHFDAHLVADFLRNFATEKLGVR
PrnA    YPQPGALDGKLAPCLSDGTRQ...............................MSHAWHFDAHLVADFLKRWAVE.RGVN
PyrH    YLTHRLCEAKRAPRMLDGSLFASQVDE.....SLGRSTLAEQ..RAQFPYAYHFDADEVARYLSEYAIA.RGVR
SttH    FVMASLCDAGRSPRYLNGSLLQQEFDERAEEPAGL.TMSEHQGKTQFPYAYHFEAALLAEFLSGYSKD.RGVK
SttH    oooooooo        TT                                  ooooooo      →  oooooooo   -
          α7                                                   α8          β8     α9


        β10       β11    TT      β12   TT   β13   β14  η1    π1       β15        β16
Thal    →         →               →        →     →   ooo  ooooo      →          →       o
        200          210         220        230         240      250         260
Thal    HVEDEMTEVLTDERGFITALRTKSGRILQGDLFVDCSGFRGLLINKAMEEPFIDMSDHLLCNSAVATAVPHDD
RebH    HVEDRVEHVQRDANGNIESVRTATGRVFDADLFVDCSGFRGLLINQALKEPFIDMSDYLLCDSAVASAVPNDD
PrnA    RVVDEVVDVRLNNRGYISNLLTKEGRTLEADLFIDCSGFMRGLLINQALEEPFIDMSDYLLCDSAVASAVPNDD
PyrH    HVVDDVQHVGQDERGWISGVHTKQHGEISGDLFVDCTGFRGLLINQTLGGRFQSFSDVLPNNRAVALRVPREN
SttH    HVVDEVLEVKLDDRGWISHVVTKEHGDIGGDLFVDCTGFRGVILLNQALGVPFVSYQDTLPNDSAVALQVPLDM
SttH    →         →      TT      →        →     →   ooo  ooo       →          →
        β9        β10            β11      β12   β13 η3   π1        β14        β15


        α8      β17   TT  β18       β19            α9              TT     β20    β21
Thal    oooo    →        →           →         oooooooooooo            ★          →
        270         280         290         300          310         320       330
Thal    EKNGVEPYTSSIAMEAGWTWKIPMLGRFGSCHVYSDHFATQDEATLAFSKLWGL.DPDNTEFNHVRFRVGRNR
RebH    DANGVEPFTSAIAMKSGWTWKIPMLGRFGTGYVYSDHFATEDEAVREFCEMWHL.DPETQPLNRIRFRVGRNM
PrnA    ARDGVEPYTSSIAMNSGWTWKIPMLGRFGSGYVYSSHFTSRDQATADFLKLWGL.S.DNQPLNQIKFRVGRNK
PyrH    D.EDMRPYTTATAMSAGWMWNTIPLFKRDGNGYVYSDEFISPEEAERELRSTVAPGR.DDLEANHIQMRIGRNE
SttH    EARGIPPYTRATAKEAGWIWTIPLIGRIGTGYVYAKDYCSPEEAERTLREFVGPEA.ADVEANHIRMRIGRSE
SttH    →         ★     →           →          oooooooooo   η4      →     β20
        β16            β17         β18           α10              β19


        β22 β23  η2 β24      TT          α10              TT        α11
Thal    →   → TT  →          →        oooooooooooo     TT      oooooooooooooooooo
        340       350         360         370         380         390         400         410
Thal    RAWVRNCVSVGLASCFVEPLESSGIYFIYAAIHMLAKHFPDKTFDKVLVDRFNREIEEMFDDTRDFLQAHYYF
RebH    RAWVGNCVSIGTSSCFVEPLESTGIYFIYAALYQLVKHFPDKSLNPVLTARFNREIETMFDDTRDFIQAHFYF
PrnA    RAWVNNCVSIGLSSCFLEPLESTGIYFIYAALYQLVKHFPDTSFDPRLSDAFNAEIVHMFDDCRDFVQAHYFT
PyrH    RTWINNCVAVGLSAAFVEPLESTGIFFIQHAIEQLVKHFPGERWDPVLISAYNERMAHMVDGVKEFLVLHYKG
SttH    QSWKNNCVAIGLSSGFVEPLESTGIFFIHHAIEQLVKHFPAGDWHPQLRAGYNSAVANVMDGVREFLVLHYLG
SttH    → TT  →     →       TT    oooooooooo                oooooooooooooooooo
        β21 β22    β23               α11                          α12


            α12           α13              α14   α15        α16
Thal      oooooo       oooooooooo         oooooo      ★     ooooooo
        420         430         440         450         460         470         480
Thal    SPRVDTPFWRANK.ELKLADSIKDKVETYRAGLPVNLPVT.DEGTYYGNFEAEFRNFWTNGSYYCIFAGLG.L
RebH    SPRTDTPFWRANK.ELRLADGMQEKIDMYRAGMAINAPASDDAQLYYGNFEEEFRNFWNNSNYYCVLAGLG.L
PrnA    TSRDDTPFWLANRHDLRLSDAIKEKVQRYKAGLPLTTT.SFDDSTYYETFDYEFKNFWLNGNYYCIFAGLG.M
PyrH    AQREDTPYWKAAKT.RAMPDGLARKLELSASHLLDEQ...........TIYPYYHGFETYSWITMNLGLG.I
SttH    AARNDTRYWKDTKT.RAVPDALAAERIERWKVQLPDSE...........NVFPYYHGLPPYSYMAILLGTCAI
SttH      oooooo  η5      oooooooooo         TT                          oooooooo
          α13            α14                                            α15


            η3      α17               α18
Thal      oooo   ooooooooooooooooo      oooooo.oo
        490         500         510         520         530
Thal    MPRNPLPALAYKPQSIAEAELLFADVKRKGDTLVESLPSTYDLLR.QLHGAS................
RebH    VPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLPSLHEFLR.QQHGR................
PrnA    LPDRSLPLLQHRPESIEKAEAMFASIRREAERLRTSLPTNYDYLRSLRDGDAGLSRGQRGPKLAAQESL
PyrH    VPERPRPALLHMD..PAPALAEFERLRREGDELIAALPSCYEYLA.SIQ..................
SttH    GLR.PSPALALAD..PAAAEKEFTAIRDRARFLVDTLPSQYEYFA.AMGQRV...............
SttH      oooo  ooooooooooooooo    oooooooo
          η6         α16               α17
```

ASBMB

Figure 2. Monomer of Trp-Thal. The box and pyramid subdomains are shown in *dark blue* and *light blue,* respectively. The substrate-binding loop that undergoes a conformational change upon L-Trp binding and closes the substrate binding site is *highlighted* in *green.* The FAD loop is shown in *red.* The substrate L-Trp is shown as a *stick model* with *yellow* carbon atoms.

Table 1

Pairwise sequence identity of various tryptophan halogenases

The tryptophan 6-halogenases Thal and BorH are closely related to each other and to the tryptophan 7-halogenase RebH. The tryptophan 6-halogenases SttH and Th-Hal are closely related to each other and to the tryptophan 5-halogenase PyrH. The sequence identity is lower between Thal or BorH on the one hand and SttH or Th-Hal on the other hand. The sequence identity was derived from pairwise structure-based sequence alignments with the DALI server (53) except for BorH (asterisks), for which no structure is available.

ID%	Thal	BorH*	RebH	PrnA	PyrH	SttH	Th-Hal
Thal	100	70*	64	57	42	40	41
BorH		100	63*	56*	40*	39*	39*
RebH			100	55	40	41	41
PrnA				100	42	40	41
PyrH					100	58	56
SttH						100	74
Th-Hal							100

Figure 3. Orientation of the substrate L-Trp in the Trp-Thal structure. *a,* $F_o - F_c$ electron density before building L-Trp contoured at $+3\ \sigma$ (*green*) and the final $2F_o - F_c$ electron density contoured at 1σ (*blue*). *b,* superposition of the bound L-Trp from the halogenases Thal (C6), RebH (C7), and PyrH (C5). The halogenated carbon atoms of the indole moiety superpose well and are located closest to the ϵ-amino group of the catalytic lysine.

Tryptophan binding site

The structure of apo-Thal, which we had determined first, revealed a substrate binding site very similar to that of RebH. All residues of the substrate binding site that differ between Thal and RebH are smaller in Thal, leaving more space for the substrate. As we were not able to determine the L-Trp binding pose or to deduce a reason for the different regioselectivity from the structure of apo-Thal, we determined the structure of Thal in complex with its native substrate L-Trp. As expected, L-Trp is oriented in such a way that its C6 position points toward the catalytic residues Lys-79 and Glu-358 (Fig. 3). The substrate binding site can be divided into three regions, namely the catalytic, the π-stacking, and the backbone-binding region (Fig. 4a). The catalytic region harbors the Lys-79 that is universally conserved among flavin-dependent halogenases and Glu-358 that is conserved among tryptophan halogenases (Fig. 4b). In addition, we assign several small amino acids, including Val-52, Val-82, and Ser-360 to this region. The π-stacking region fixes the indole moiety (Fig. 4c) and may shield the nonhalogenated aromatic carbon atoms of the indole from reactive HOCl. It consists of the aromatic residues His-110, Phe-112, Phe-465, and Trp-466. In addition, we assign Pro-53 and Pro-111 to this middle region (Fig. 4, *a* and *c*). The backbone-binding region, which is in part poorly ordered in the absence of substrate, contains residues that coordinate the α-amino and α-carboxyl group of L-Trp. It contains Gly-113 and Lys-114 from a loop that flips into position upon substrate binding; Tyr-454, Tyr-455, and Glu-461 from the substrate-binding loop; and other polar residues like Thr-467 and Ser-470 (Fig. 4, *a* and *d*).

Structural changes upon L-Trp binding

A comparison of apo-Thal and Trp-Thal indicates structural changes upon binding of L-Trp. The *B* factors of the loop covering the FAD binding site (FAD loop; aa 38–45) increase while the substrate-binding loop (aa 450–457) becomes more ordered. In chain B, Phe-112 undergoes a peptide flip, moving Gly-113 and Leu-114 closer to the substrate by 2.4 and 1.4 Å,
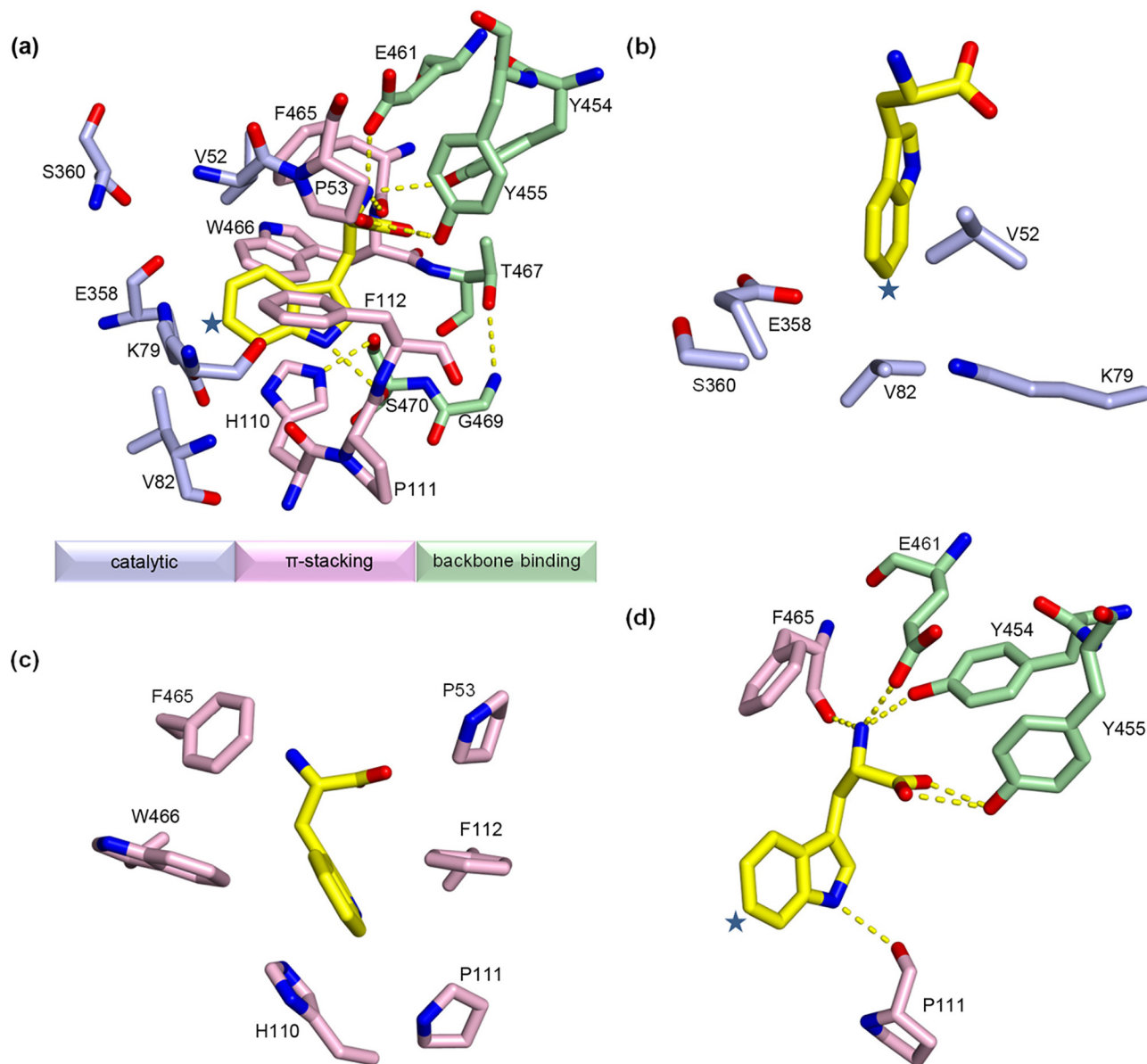
Figure 1. Structure-based sequence alignment. The tryptophan 6-halogenase Thal (PDB code 6H44), the two tryptophan 7-halogenases RebH (2E4G) and PrnA (2AQJ), the tryptophan 5-halogenase PyrH (2WEU) and the tryptophan 6-halogenase SttH (5HY5) are ordered by decreasing sequence identity with Thal. Secondary structure elements of Thal and SttH are shown *above* and *below* the sequence, respectively. The catalytic lysine and glutamate are *highlighted* in *green.* Fingerprint sequences conserved in FDHs (G*X*G*XX*G and W*X*W*X*IP) are *highlighted* in *orange.* The Asp-449 insertion in RebH relative to Thal is *highlighted* in *yellow.* The five residues mutated in Thal-RebH5 are *highlighted* in *cyan.* The figure was generated with ESPpript (54). Pairwise structure-based sequence alignments were performed with the DALI server (53).

Figure 4. Depiction of the tryptophan binding site. *a*, the binding site is divided into three regions: catalytic (*light blue*), π-stacking (*light pink*), and backbone-binding (*light green*). The substrate L-Trp is shown with *yellow* carbon atoms, and the halogenation position C6 is *highlighted* with a *star*. Direct hydrogen bonds or salt bridges between Thal and L-Trp are shown as *dotted yellow lines*. *b*, catalytic region. *c*, π-stacking region. *d*, backbone-binding region.

respectively. A similar peptide flip involving the equivalent residue Phe-111 is observed for RebH and Trp-RebH (35, 36). Phe-112 in chain A of apo-Thal adopts the same conformation as chain A and B of Trp-Thal, presumably because the substrate binding site is not truly empty, as indicated by positive difference density. Upon substrate binding, residues of the backbone-binding region that form polar contacts with L-Trp become ordered and block the way out through the substrate tunnel to the dimer interface (Fig. S3). The α-amino group of L-Trp forms a salt bridge with the carboxyl group of Glu-461 and two H-bonds, one to the hydroxyl group of Tyr-454 and the other to the backbone carbonyl of Tyr-465 (Fig. 4d). The carboxyl group of L-Trp forms H-bonds to the hydroxyl group of Tyr-455 and to two conserved water molecules (see below). Upon substrate binding, several residues move out of the active

site (Fig. 5). This makes room for L-Trp both in the catalytic region (Thr-51, 0.6 Å; Val-52, 0.7 Å; Glu-358, 1.0 Å) and in the π-stacking region (Pro-53, 0.7 Å; Phe-112, 1.0 Å; Trp-466, 0.6 Å).

Comparison of substrate binding in Thal, RebH, and PyrH

The indole moiety in Trp-Thal is almost parallel to that in Trp-RebH but flipped by 180°, so that the five-membered rings do not superpose at all (Fig. 3). In contrast, the halogenated carbon atoms (C6 in Trp-Thal and C7 in Trp-RebH) are only 0.5 Å apart. The same flip of ~180° relative to the orientation in RebH is present in PyrH. Hence, the binding pose of the substrate indole moiety in Thal is more like that in PyrH than that in RebH. Interestingly, L-Trp adopts different rotameric conformations in Thal and in RebH, whereas it is bound to PyrH in

Figure 5. Structural flexibility of conserved residues in the π-stacking region. Overlay of the π-stacking region of chain A of apo-Thal (*gray*), Trp-Thal (*dark blue*), and Trp-RebH (PDB code 2E4G; *violet*). The substrate tryptophan is shown in *yellow* for Thal and in *violet* for RebH. The numbering of residues corresponds to that of Thal. RebH contains Ser-110 instead of Pro-111. In apo-Thal, Pro-53, Phe-112, and Trp-466 have moved into the substrate binding site relative to Trp-Thal. In chain B, the movement toward the empty substrate binding site of apo-Thal is even larger (see Fig. S7).

a nonfavorable conformation. In Thal and PyrH, the halogenated carbon atoms (C6 in Trp-Thal and C5 in Trp-PyrH) again overlap closely with a distance of only 0.2 Å. The indole nitrogen forms an H-bond to structurally equivalent residues in Thal (Pro-111) and PyrH (Pro-93), but to a different residue in RebH (Glu-357).

The rather low sequence and structural conservation between Thal and PyrH makes a meaningful comparison of substrate binding to these proteins difficult. The catalytic and π-stacking regions of the Trp-PyrH (PDB code 2WEU) active site show only a few substitutions (V52F, V82I, S360T, and W466F; numbering according to Thal) and align fairly well with the structure of Trp-Thal. In contrast, the backbone-binding region is completely different in structure and sequence (Fig. 1 and Fig. S2b). The rearranged structure of the active site lid is the most obvious difference. In Thal, residues 445–465 form two minihelices and loops covering the tryptophan binding site. The corresponding PyrH loop aa 441–451 only partly covers the area above the active site. Therefore, in PyrH, the active site is further covered by an α-helix (aa 157–162), which structurally corresponds to residues 451–459 of Thal. In Thal, Tyr-454, Tyr-455, and Glu-461 are key residues of the substrate-binding loop and form direct or water-mediated H-bonds and a salt bridge with L-Trp. These residues are replaced by Gln-160, Gln-163, and Tyr-454 (which does not superpose with Thal Tyr-454) in PyrH. Accordingly, the distance of the backbone atoms of L-Trp are larger between Thal and PyrH than between Thal and RebH.

Due to the higher conservation, a comparison of Thal with RebH appears more rewarding. Considering the direct surrounding of L-Trp in Thal and RebH (*i.e.* residues with atoms within 4 Å of L-Trp), there are only four substitutions, namely

V52I, V82I, S470N, and Pro-111 → Ser-110 (Thal → RebH). In a slightly extended definition of the surroundings (*i.e.* residues with atoms within 5 Å of L-Trp or one of two conserved water molecules (see below)), there are another three substitutions, namely Ser-360 → Thr-359, T467N, and G469S (Fig. S4). The side chains of Ser-110 in RebH and Pro-111 in Thal point away from the substrate. Thus, the larger Pro side chain in Thal presumably has little influence on L-Trp binding, as the backbone atoms that line the substrate binding site superpose well. The other six substitutions result in smaller side chains in Thal than in RebH. Thus, the substrate binding site of Thal should be more spacious than that of RebH. Three of the substitutions (V52I, V82I, and Ser-360 → Thr-359) are in the catalytic region and, therefore, in close contact with the indole moiety. The remaining three substitutions are in the backbone-binding region. Two substitutions (T467N and G469S) will potentially impact the water structure involved in coordinating the L-Trp carboxyl group. Finally, the S470N substitution impacts both the coordination of a conserved water molecule (see below) and the indole binding site. Serine in position 470 provides the necessary space for the five-membered ring of L-Trp that is required upon the 180° flip of the indole moiety.

Water within the Trp binding site

Along with the substrate L-Trp, some water molecules get trapped in the active site. The position of these waters is conserved between the two crystallographically independent Thal monomers and sometimes even among Thal and RebH. There are two prominent water molecules forming H-bonds with the L-Trp carboxyl group. Their positions are conserved between Thal and RebH, and here we call them W1 and W2 for simplicity (Fig. 6). W1 corresponds to A710 and B721 in Trp-Thal and to water A1031 and B1125 in Trp-RebH (PDB code 2E4G). W1 forms an H-bond with the carboxyl group of the substrate and two H-bonds with the protein via the side chain carboxyl group of the Glu-461 and the backbone NH of Asn-54 (Fig. 6, *a* and *c*). The H-bond pattern between W1 and the protein is identical in Thal and RebH, and W1 is also present in chain A of apo-Thal (water A718) (Fig. 6b) and in apo-RebH (waters A777 and B646; PDB code 2OAM) (Fig. 6d). W2 corresponds to A752 and B712 in Trp-Thal and to waters A1080 and B1091 in Trp-RebH (PDB code 2E4G)). In contrast to W1, W2 is missing in the apo-structures of Thal and RebH (Fig. 6, *b* and *d*). Like W1, W2 forms an H-bond with the carboxyl group of the substrate L-Trp. In addition, W2 forms an H-bond with the carbonyl oxygen of Gly-113 (Gly-112 in RebH). This H-bond in combination with the absence of W2 in the apo-structures may be one reason for the peptide flip of Phe-112 upon substrate binding. Further H-bonds between W2 and the protein differ between Thal and RebH. W2 forms two direct H-bonds with RebH via the side-chain amides of Asn-467 and Asn-470 (Fig. 6c). In contrast, W2 in Thal forms an H-bond with a secondary water molecule (A717 and B713, W3 in Fig. 6a), which in turn is bound to Thal via two H-bonds to the hydroxyl groups of Thr-467 and Ser-470. W3 is conserved between Trp-Thal monomers, but not among Trp-Thal and Trp-RebH. This difference between Thal and RebH is presumably due to the amino acid substitutions in the backbone-binding region of the active site.
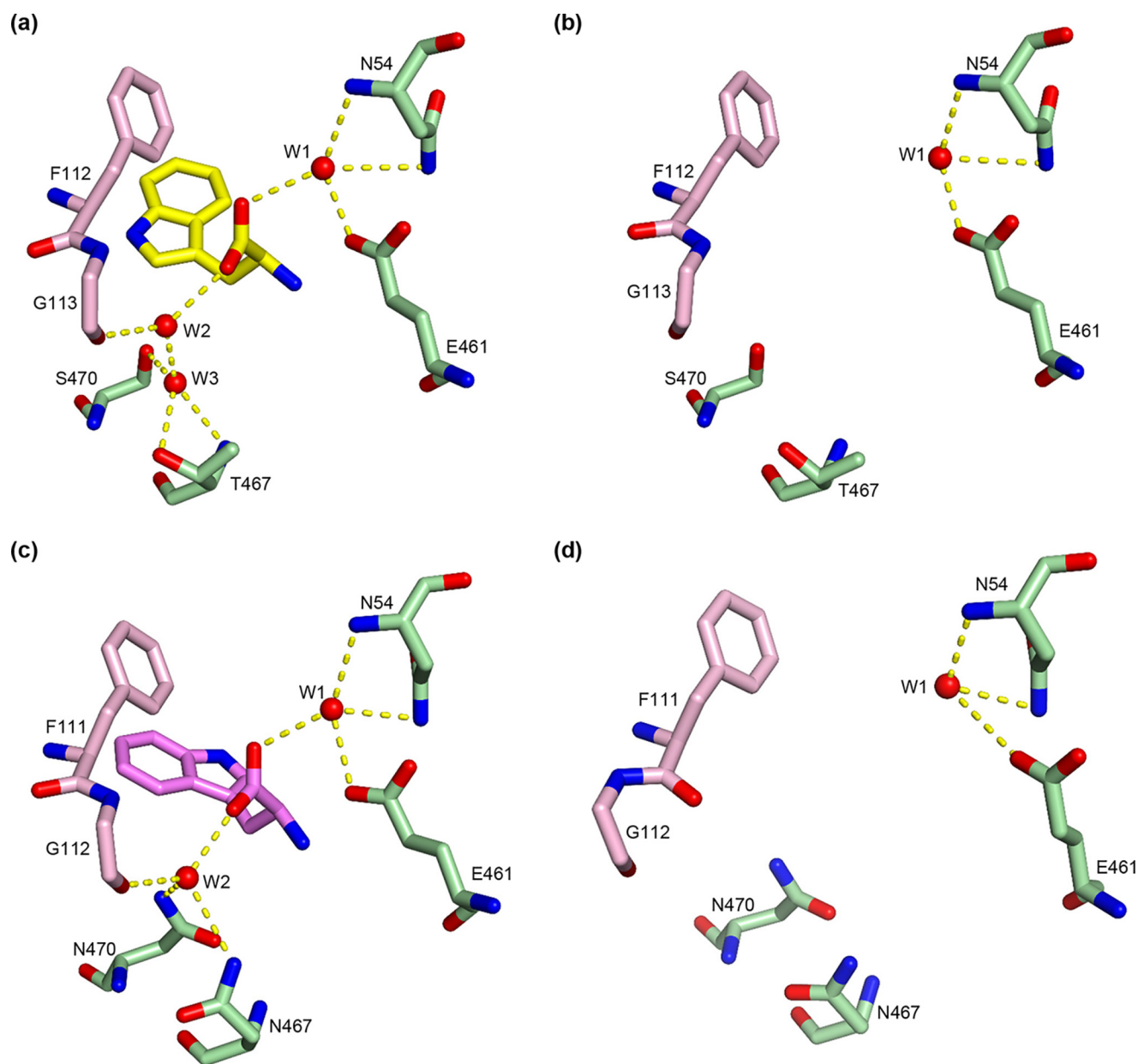
Figure 6. Conserved water molecules that form H-bonds to the L-Trp carboxylate in the substrate binding sites of Thal and RebH. Conserved water molecules W1 and W2 and their H-bond partners are shown. *a*, in Trp-Thal, W2 forms an H-bond to Gly-113. Further contacts to Thal are mediated by water W3. *b*, in chain A of apo-Thal, W1 is also present. *c*, in Trp-RebH (PDB code 2E4G), W2 forms three direct H-bonds with the protein. The contact with Gly-112 is conserved in Thal, but the contacts to Asn-467 and Asn-470 are not. *d*, in apo-RebH (PDB code 2OAM), W1 is present, and Phe-111 has undergone a peptide flip relative to Trp-RebH so that the carbonyl oxygen of Gly-112 could no longer form an H-bond to W2.

Switching the regioselectivity by structure-guided mutagenesis

To test whether a few targeted mutations are sufficient to switch the regioselectivity, we substituted five residues in the substrate binding site of Thal by the corresponding amino acids of RebH. The resulting variant Thal-RebH5 harbored the substitutions V52I, V82I, S360T, G469S, and S470N already described above (Fig. S4). To assess the overall conversion of tryptophan, assay mixtures using purified protein and concomitant cofactor regeneration were analyzed via HPLC-MS (Fig. 7). After a reaction time of 86 h, 70% conversion of 5 mM L-Trp to L-chlorotryptophan was reached. This proved that these substitutions in the active site do not affect the overall conversion.

In addition to the main signal observed in the chromatogram, a less intense signal indicated the formation of a second product (Fig. 7c). As both signals were identified as monochlorinated tryptophan with the same m/z pattern from their corresponding mass spectra, we assumed that another regioisomer might have formed. HPLC-MS analysis showed that both isomers were formed in a ratio of 95:5, whereas comparison with authentic standards revealed that 7-chlorotryptophan was formed as the main product. The minor signal corresponds to 6-chlorotryptophan.

In addition, chlorination of L-Trp was conducted on a preparative scale using *Escherichia coli* lysate containing overexpressed halogenase. After 26 h, a maximum conversion of 30%

Regioselectivity switch of tryptophan 6-halogenase Thal

Figure 7. Thal-RebH5-catalyzed halogenation yielded L-7-halotryptophan. *a,* halogenation scheme and final ratio of chlorotryptophan isomers resulting from Thal-RebH5-catalyzed halogenation assays indicate the dominant formation of C7 isomer. *b–d,* LC-MS analysis confirms the formation of two monochlorinated L-Trp isomers with the same *m/z* pattern. *m/z* $[M^{35Cl} + H]^+$ found 239.06, calc. 239.06, $[M^{37Cl} + H]^+$ found. 241.06, calc. 241.06. HPLC-MS-analysis of crude reaction mixture was compared with authentic standards. *b,* 7-Cl-Trp standard, $t_r = 6.6$ min. *c,* the trace of the reaction mixture revealed the formation of C7-chlorinated L-Trp as the major regioisomer (95.4%), whereas 6-Cl-Trp appeared as a minor side product (4.6%). Reaction conditions were as follows: 89 μM Thal-RebH5, 5 mM L-Trp, 1 mM NAD$^+$, 0.01 mM FAD, 1 unit·ml^{-1} *RR*-ADH, 2.5 units·ml^{-1} PrnF, 5% (v/v) 2-propanol, 10 mM Na$_2$HPO$_4$, pH 7.4, 30 mM NaX, X = Cl, Br. *d,* 6-Cl-Trp standard, $t_r = 6.9$ min.

into the expected monochlorinated products (95:5 ratio) was observed via HPLC-MS. Separation and purification of the main product by reversed-phase HPLC finally allowed the identification of L-7-chlorotryptophan by one- and two-dimensional NMR spectroscopy (Figs. S8–S13). This proves the successful and nearly quantitative change of the regioselectivity from position 6 to 7 by substituting five residues of Thal by the corresponding RebH residues.

Enzymatic bromination reactions were performed on a preparative scale analogously to the chlorination described above. After a reaction time of 48 h, HPLC-MS data revealed 15% conversion to monobrominated L-Trp. Like the chlorination, two product signals in a ratio of 95:5 were observed. As the overall conversion was lower, the minor isomer was hardly visible in the UV trace, but its quantity was sufficient for MS detection. After purification, C7-brominated L-Trp was isolated as the major bromination product of Thal-RebH5, which was also confirmed by NMR analysis (Figs. S14–S17). Corresponding to the previous observations made for Thal-RebH5–catalyzed chlorination, 6-bromotryptophan had been formed in negligible quantity, as identified via an authentic HPLC standard.

Structure of Thal-RebH5

Thal-RebH5 crystallized like the WT protein, and the 2.12 Å structure (Table S1) showed clear electron density for the five mutated residues. The rotamers of substituted residues in Thal-RebH5 correspond to those of RebH except for Ser-469. An overlay of apo-Thal, apo-Thal-RebH5, and apo-RebH showed that in Thal-RebH5, most active-site residues, including many conserved residues, adopt an intermediate conformation between Thal and RebH (Fig. S5). In some cases, like Thr-360 or Phe-465, Thal-RebH5 more closely resembles RebH than WT Thal. We also soaked L-Trp into Thal-RebH5 crystals and obtained a structure at 2.55 Å resolution. There was clear positive difference density in the active site of chain A (Fig. S6*a*) and some positive difference density in the active site of chain B. The difference density in chain A is akin to the shape of L-Trp but cannot be modeled faithfully. The best interpretation that we achieved is a superposition of L-Trp as it is bound to Thal (conformer A) and to RebH (conformer B) (Fig. S6*b*). Occupancy refinement converged to values of about 0.5 for conformer A and about 0.36 for conformer B. If only one conformer was modeled, the occupancy refined to values of about 0.8–0.9. Placing a single conformer, however, resulted in positive differ-
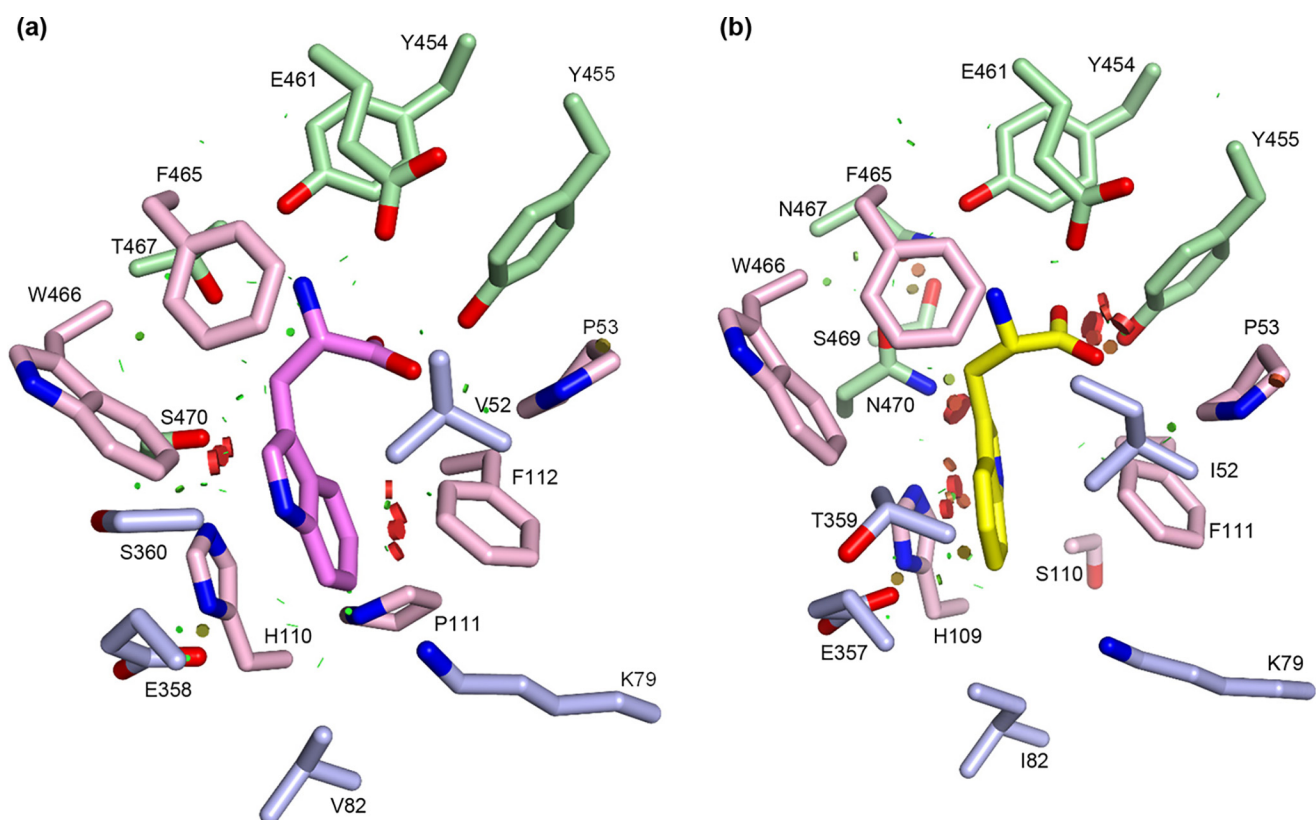
Figure 8. Steric conflicts upon placing L-Trp from Trp-RebH in Trp-Thal and vice versa. Minor clashes are indicated as *small green hexagons,* and more severe clashes are shown as *larger red hexagons. a,* L-Trp from Trp-RebH (PDB code 2E4G) was placed in Trp-Thal. Clashes mainly occur between the L-Trp indole and the side chains of conserved residues Phe-112 and Trp-466. *b,* L-Trp from Trp-Thal was placed in Trp-RebH (PDB code 2E4G). Clashes mainly occur between the L-Trp indole and the side chain of Asn-470 and between the L-Trp carboxylate and the side chain of Tyr-455.

ence density in the region occupied by the five-membered ring of the other conformer (Fig. S6, *c* and *d*), supporting the presence of the second conformer. While the plane of the indole moiety was well-defined, each conformer could be rotated or translated to a certain degree in this plane. Positional refinement of L-Trp did not converge to a uniquely defined solution, presumably due to the partial occupancy and the limited resolution. Therefore, we decided not to deposit this structure in the PDB but to provide coordinates and structure factors as supporting information. The Trp-Thal-RebH5 structure strongly suggests that Thal-RebH5 is permissive to binding L-Trp with its C7 positioned most closely to the ε-amino group of the catalytic Lys-79.

Discussion

Here, we present the first experimental structure of a tryptophan 6-halogenase in complex with its native substrate L-Trp. The structure confirms the model of regioselectivity put forward by the Naismith and van Pée groups (26, 28). As in the tryptophan 7-halogenases PrnA and RebH and the tryptophan 5-halogenase PyrH, the halogenated aromatic carbon atom is positioned most closely to the ε-amino group of the catalytic lysine in Thal. It clearly appears to be the binding pose of the substrate that determines the regioselectivity.

Our apo– and L-Trp–bound structures of Thal suggest that the empty active site of the enzyme may not be well-suited to reliably model the substrate binding pose for three reasons. First, in the absence of L-Trp, several conserved residues in the

catalytic region and the π-stacking region move toward the center of the substrate binding site. If one places L-Trp from the Trp-Thal structure into structurally aligned apo-Thal, there would be severe clashes (Fig. S7). Second, major conformational changes between the apo-Thal and Trp-Thal structure affect key residues for L-Trp binding in the backbone-binding region (Fig. 4d and Fig. S3). Finally, two water molecules that are conserved between Thal and RebH mediate interactions between the L-Trp carboxyl group and the protein, but only one of these waters is present in the substrate-free structures of Thal and RebH (Fig. 6).

Given the high degree of sequence and structural similarity between Thal and RebH, also in the active site, it may have come as some surprise to find in Thal an indole binding pose that more closely resembles that of PyrH than that of RebH. Altogether, this work underscores the importance of experimental structure determination, even in cases where structures of highly homologous proteins are available.

To gain more insight into which residues may be key to determine the regioselectivity in RebH and Thal, we performed a structural overlay of Trp-Thal and Trp-RebH (PDB code 2E4G) and copied the L-Trp from Trp-RebH to Trp-Thal and vice versa. We then checked for clashes indicating which residues might prevent binding of L-Trp to Thal in the same binding pose as in RebH and vice versa.

When placing L-Trp from RebH in the Trp-Thal structure (Fig. 8a), clashes mainly occur with conserved residues. This

may not be surprising, given the fact that all sequence differences in the active sites of Thal and RebH have smaller side chains in Thal. Several conserved residues of the catalytic and the π-stacking region move toward the substrate in Trp-Thal compared with Trp-RebH. Accordingly, there would be close contacts between L-Trp from RebH and Thal residues Pro-53, Phe-112, and Trp-466. Reasons why these residues move into the active site in Thal are not easily apparent, as even the shell of residues surrounding Pro-53, Phe-112, and Trp-466 is highly conserved between Thal and RebH. We suspect that the indole binding pose in Thal may not be determined primarily by steric hindrance from residues surrounding the indole moiety but by the different positioning of L-Trp Cα due to the different H-bond pattern between the L-Trp carboxylate and the backbone-binding region of the protein.

In contrast, moving L-Trp from Trp-Thal to Trp-RebH results in distinct clashes between RebH and L-Trp (Fig. 8b). Two RebH residues would form particularly short contacts with L-Trp. First, the side chain of Asn-470 would clash with C2 and C3 of the indole moiety. Second, the aromatic ring of Tyr-455 would clash with the L-Trp carboxyl group. Our conclusions are consistent with the previous notion of the van Pée and Naismith groups that in PrnA the equivalent residues Tyr-444 and Asn-459 prevent L-Trp from adopting the same orientation as that seen in PyrH (41). Intriguingly, even a random mutagenesis approach to select RebH variants that can halogenate larger substrates picked up the N470S substitution (39).

These observations suggest that the N470S substitution (RebH → Thal) may be of central importance to facilitate the indole binding pose observed in Trp-Thal. Interestingly, a serine is also present at this position in BorH (Ser-469), SttH (Ser-464), Th-Hal (Ser-460), and PyrH (Ser-455). In PyrH, however, mutating this Ser to Asn would not result in clashes with the L-Trp indole moiety. Still, the N470S substitution may be an enabling mutation in RebH that allows flipping the indole moiety relative to the orientation found in Trp-RebH. Due to the higher activity of the N470S variant on tryptamine, Andorfer *et al.* (37) used RebH N470S as starting point for their directed evolution of RebH toward C5 or C6 selectivity. Interestingly, RebH N470S retained regioselectivity for the 7-position. It produced >99% 7-chlorotryptamine and only minute amounts of 5- and/or 6-chlorotryptamine. However, all of the later variants that halogenated up to 95% of tryptamine on C5 or C6 had this founding N470S substitution.

The second RebH residue that would clash with L-Trp from Thal, namely Tyr-455, is conserved between RebH and Thal. Despite sequence conservation, its spatial arrangement is markedly different between RebH and Thal. In RebH, Tyr-455 is pushed toward the substrate relative to Thal. This displacement may be caused by the only insertion in RebH relative to Thal. A structure-based sequence alignment identifies Asp-449 as the inserted residue (Fig. 1). This different location of Tyr-455 decreases the space available to the L-Trp backbone atoms in RebH. Again, the extensive mutagenesis work on RebH by Andorfer *et al.* (37) provides circumstantial support for this hypothesis, as their RebH substitution variants with good halogenation activity on C5 or C6 of tryptamine had no activity toward the larger L-Trp.

Previous studies demonstrated that halogenase regioselectivity can be altered using different mutagenesis strategies. In a rational approach, a partial change of regioselectivity yielding a 2:1 ratio of 7- and 5-bromotryptophan was achieved for PrnA (41). In an approach more similar to ours, Micklefield and co-workers (34) tried to switch the regioselectivity of the tryptophan 6-halogenase SttH to position 5. Their work is based on structural differences between SttH and the closely related tryptophan 5-halogenase PyrH. A triple variant of SttH showed activity like WT SttH with relaxed regioselectivity, leading to 32% 5-Cl-L-Trp. A larger switch in regioselectivity was observed for the smaller substrate 3-indolepropionic acid, yielding 75% 5-chlorinated product compared with just 10% 5-chlorination for WT SttH. In a directed evolution approach, Lewis and co-workers (37) used random mutagenesis in combination with structure-guided mutation of active site residues to modify the regioselectivity of RebH. Their efforts to evolve the tryptophan 7-halogenase RebH into a C5- or C6-regioselective biocatalyst resulted in the generation of two new variants capable of chlorinating tryptamine with up to 95 and 90% regioselectivity for positions 5 and 6, respectively. For 2-methyltryptamine and *N*-methyltryptamine, the selectivity was even higher. Both the C5- and the C6-selective variant were inactive on the native substrate L-Trp.

In summary, previous studies achieved nearly complete selectivity switches for nonnative substrates like 3-indolepropionic acid and tryptamine both by rational, structure-based engineering (34) and by (random) mutagenesis coupled to high-throughput screening (37). For L-Trp, however, structure-based mutagenesis of both PrnA (41) and SttH (34) shifted regioselectivity to only about one-third of 5-Cl-L-Trp from 100% each of 7-Cl-L-Trp and 6-Cl-L-Trp, respectively. Our Thal-RebH5 variant exhibited a 95% change in regioselectivity toward C7 halogenation for both chlorination and bromination of L-Trp, constituting a nearly quantitative switch that has not been reported for any tryptophan halogenase until now. Both 3-indolepropionic acid and tryptamine are smaller than L-Trp and should have more flexibility to accommodate in the active site, offering a potential explanation why switching regioselectivity for smaller substrates was more successful. This may be particularly relevant for RebH, because, as pointed out before, the insertion of Asp-449 in RebH pushes the conserved Tyr-455 into the active site compared with Thal. Thus, Thal, which provides more space for the L-Trp backbone atoms, may represent a better starting point than RebH to evolve or engineer regioselectivity on the larger substrate L-Trp.

The success of our protein engineering is based on a careful structural analysis of substrate binding in closely related tryptophan halogenases with different regioselectivity. As such, our approach most closely resembles that of Shepherd *et al.* who also based their mutations on structures of two closely related tryptophan halogenases (34). Our results show that of the roughly 200 sequence differences between RebH and Thal, only a handful are required to switch regioselectivity. All of these substitutions affect active site residues. Substitution of residues distant from the substrate binding site that are frequently picked up in random mutagenesis but almost impossible to predict in rational structure-based approaches was not required.

Experimental procedures

Plasmids and molecular cloning

The pET-28a plasmid containing a codon-optimized gene for *E. coli* encoding Thal (Uniprot ID: A1E280) was reported earlier (43). The genes encoding Thal and Thal-RebH5 were separately subcloned into the vector pETM-11 (Gunter Stier, Heidelberg) using the NcoI and NotI restriction sites with primers listed in Table S2. The plasmid vector pET-21-*adh* encoding alcohol dehydrogenase from *Rhodococcus ruber* (*RR*-ADH) was kindly provided by Prof. Dr. Werner Hummel, Bielefeld University, and pCIBhis-*prnF* encoding the flavin reductase PrnF from *Pseudomonas fluorescens* was donated by Prof. Dr. Karl-Heinz van Pée, TU Dresden. Plasmid pGro7 was purchased from TaKaRa. *E. coli* DH5α and BL21(DE3) were purchased from Novagen. Generation of the Thal-RebH5 quintuple variant was performed according to the QuikChange II site-directed mutagenesis kit with primers listed in Table S2. Mutations were introduced consecutively starting with pET-28a-*thal* WT as DNA template. Mutations were confirmed by DNA sequencing (GATC Biotech).

Expression and purification of recombinant Thal for crystallization

N-terminal His$_6$-tagged Thal was overexpressed in *E. coli* BL21(DE3) using a pETM-11 vector. The culture was grown in lysogeny broth medium with kanamycin (30 mg·liter^{-1}) at 37 °C until an A_{600} of 0.6 was achieved. After that, the culture was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside and shaken for an additional 20 h at 25 °C. The cells were harvested by centrifugation (30 min, 4 °C, 5180 $\times$ *g*), resuspended in lysis buffer (50 mM sodium phosphate, pH 7.4, 300 mM NaCl) with additional DNase I and a cOmplete™ EDTA-free protease inhibitor tablet (Roche Applied Science), and disrupted using a SPCH high-pressure homogenizer (Stansted Fluid Power Ltd.). The lysate was clarified by centrifugation (4 °C, 1 h, 30,966 $\times$ *g*). Soluble Thal was purified by nickel-nitrilotriacetic acid (Ni-NTA) agarose chromatography (pure Cube Ni-NTA agarose, Cube Biotech; buffer: 50 mM sodium phosphate, pH 7.4, 300 mM NaCl, 1 mM β-ME, 10–300 mM imidazole), followed by ion exchange (Source 15Q, GE Healthcare; buffer: 50 mM Tris, pH 7.4, 50–500 mM NaCl, 1 mM DTT) and size-exclusion chromatography (HiLoad 16/60 Superdex 200 Prep grade, GE Healthcare; buffer: 10 mM Tris, pH 7.4, 50 mM NaCl, 1 mM DTT). After the Ni-NTA and during buffer exchange through dialysis, the N-terminal His$_6$ tag was cleaved via tobacco etch virus protease at 20 °C overnight using a mass ratio (w/w) of 1:50. Purified Thal was concentrated up to 25 mg·ml^{-1} using Vivaspin concentrators (Sartorius AG; molecular weight cutoff 10,000) and stored in crystallization buffer (10 mM Tris, pH 7.4, 50 mM NaCl, and 1 mM tris(2-carboxyethyl)phosphine) at −80 °C until further usage. The variant Thal-RebH5 was expressed and purified as described for Thal, except using a Co-NTA (pure Cube Co-NTA agarose, Cube Biotech) instead of Ni-NTA.

Protein expression and purification for enzymatic assays

Expression of PrnF and alcohol dehydrogenase was performed according to published procedures (5). For Thal and Thal-RebH5, 1.5 liters of lysogeny broth medium containing kanamycin (60 mg·liter^{-1}) and chloramphenicol (50 mg·liter^{-1}) were inoculated with 20 ml·liter^{-1} overnight culture of *E. coli* BL21 (DE3) pGro7 pET28a-*thal* WT or *-thal-rebh5*. The expression culture was incubated at 37 °C until an $A_{600} = 0.6$ was reached. Temperature was decreased to 25 °C for 30 min, and overexpression was subsequently induced by the addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside and 2 g·liter^{-1} L-arabinose. Cells were shaken at 150 rpm for 20 h, harvested by centrifugation (3220 $\times$ *g*, 30 min, 4 °C), washed with 40 ml of Na$_2$HPO$_4$ buffer (0.1 M, pH 7.4), and stored at −20 °C. His$_6$-tagged fusion protein was extracted from *E. coli* cells by resuspending the bacteria in 40 ml of 50 mM Na$_2$HPO$_4$, pH 7.4, 300 mM NaCl for subsequent lysis by French press (three times). Cell debris was removed by centrifugation (10,000 $\times$ *g*, 30 min, 4 °C), and the supernatant was filtered through a 0.2-μm Whatman filter. The cleared lysate was loaded on HisTALON agarose affinity resin (bed volume of 1.5 ml of resin, flow rate, 0.5 ml·min^{-1}), followed by washing steps (10 ml; flow rate, 1 ml·min^{-1}) with 50 mM NaH$_2$PO$_4$, pH 7.4, 300 mM NaCl and with 50 mM NaH$_2$PO$_4$, pH 7.4, 300 mM NaCl, 10 mM imidazole. The fusion protein was eluted by the addition of 300 mM NaCl, 300 mM imidazole, 50 mM Na$_2$HPO$_4$, pH 7.4 (flow rate, 1 ml·min^{-1}). Fractions of 0.5 ml were collected, and protein concentration was determined by Nano-Drop UV spectroscopy.

Crystallization and data collection

Purified Thal was crystallized using the sitting-drop vapor-diffusion method at 20 °C with a drop ratio of 2:1 of protein solution (~15 mg·ml^{-1}) and reservoir solution (0.1 M Tris, pH 8.5, 1.5 M K$_2$HPO$_4$/KH$_2$PO$_4$). To obtain an L-Trp-bound structure, Thal crystals were soaked in reservoir solution additionally containing 5 mM L-tryptophan. The Thal crystals for soaking experiments were obtained from a different crystallization condition (reservoir solution: 0.1 M Bicine, pH 9.0, 1.6 M K$_2$HPO$_4$/NaH$_2$PO$_4$; drop ratio, 1:1). Thal-RebH5 was crystallized as the WT but using another reservoir solution (0.1 M Bicine, pH 8.4, 1.3 M K$_2$HPO$_4$/KH$_2$PO$_4$). Trp-Thal-RebH5 crystals were obtained by soaking Thal-RebH5 crystals in reservoir solution saturated with L-tryptophan for 75 min. Crystals appeared as hexagonal prisms within 7 days. For cryoprotection, the crystals were transferred to reservoir solution supplemented with 25% glycerol, before being flash-cooled and stored in liquid nitrogen. All data were collected at a temperature of 100 K at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring at DESY, Hamburg, Germany (44).

Data processing, structure determination, and refinement

The data sets were processed with XDS and scaled with XSCALE (45). All crystals belonged to the space group P6$_4$ and showed no indication of twinning using the program Aimless (46) from the CCP4 package (47). A summary of the data processing is shown in Table S1. The structure of apo-Thal was determined by molecular replacement using the program

ASBMB

J. Biol. Chem. (2019) 294(7) 2529–2542 **2539**

Phaser (48). The search model consisted of an ensemble of structures of the tryptophan halogenases RebH (PDB code 2OAM), PrnA (PDB code 2AQJ), PyrH (PDB code 2WEU), and SttH (PDB code 5HY5). Phaser placed two molecules per asymmetric unit with no clashes and rotation function z-scores of 21.0 and 9.3 and translation function z-scores of 42.2 and 84.8. The Trp-Thal structure was determined using the apo-Thal structure as a search model and was used as search model for the apo-Thal-RebH5 and Trp-Thal-RebH5 structures.

The first layout of the apo-Thal structure was generated with AutoBuild (49) from the PHENIX suite (50) and improved by modeling in COOT (51) and restrained refinement in Refmac5 (52) using noncrystallographic symmetry restraints or, in the case of Thal-RebH5, improved by phenix.refine using TLS refinement and noncrystallographic symmetry restraints. All structures were optimized until R values converged (Table S1). The occupancy of L-Trp in Trp-Thal refined to a value of 1. Figures of the final structures were generated using PyMOL. All r.m.s.d. values were calculated with the CCP4 program TOPP with the setting "superpose using secondary structure matching."

Enzymatic chlorination of L-tryptophan using purified halogenase

Enzymatic chlorination was assayed in a total volume of 500 μl while shaking gently at 25 °C. Substrate L-tryptophan was added to a final concentration of 5 mM together with 1 mM NAD$^+$, 0.01 mM FAD, 1 unit·ml^{-1} RR-ADH, 2.5 units·ml^{-1} flavin reductase PrnF, 5% (v/v) isopropyl alcohol, 10 mM Na$_2$HPO$_4$, pH 7.4, and 30 mM NaCl. Purified halogenase was added to a final concentration of 89 μM. Reactions were performed as triplicates, and reaction progress was monitored by HPLC-MS or analytical HPLC. Aliquots of 30 μl were taken from the reaction mixture at different time points and quenched by adding an equal volume of methanol. The mixture was centrifuged (12,000 $\times$ g, 10 min), and the supernatant was analyzed via HPLC-MS or analytical HPLC.

Enzymatic bromination/chlorination of L-tryptophan using E. coli crude lysate

E. coli BL21(DE3) pGro7 cells overexpressing Thal-RebH5 from 1.5 liter of culture were resuspended in 20 ml of chlorination buffer (10 mM Na$_2$HPO$_4$, pH 7.4, 30 mM NaCl). The cells were lysed by French press (three times), and cell debris was removed by centrifugation (10,000 $\times$ g, 30 min, 4 °C). Subsequently, lysate was added to 5 mM L-tryptophan, 1 mM NAD$^+$, 0.01 mM FAD, 1 unit·ml^{-1} RR-ADH, 2.5 units·ml^{-1} PrnF, 5% (v/v) isopropyl alcohol, 10 mM Na$_2$HPO$_4$, pH 7.4, 30 mM NaCl, in a final volume of 30 ml. The reaction mixture was incubated at 25 °C in a shaking incubator (150 rpm), and the reaction progress was monitored by HPLC-MS as described above. Bromination assays were performed analogously, except that an aqueous buffer consisting of 10 mM K$_2$HPO$_4$, pH 7.4, and 30 mM NaBr was used for resuspending the cells and subsequent biocatalysis.

Analytical reversed-phase HPLC

A Shimadzu CBM-20A chromatography system was employed for analytical measurements equipped with a diode array detector SPD-M20A IVDD (Shimadzu), pump Nexera XR LC LC-20AD$_{XR}$ (Shimadzu), column oven CTO-20A (Shimadzu), and an autosampler, Nexera XR SIL-20A$_{xr}$ (Shimadzu). For separation, a Hypersil Gold C18 column (3 μm, 150 $\times$ 2.1 mm, from Thermo Scientific, with a flow rate of 0.5 ml·min^{-1}) was used. Absorbance was measured simultaneously at 220, 254, and 280 nm. Eluents were as follows: A, water/TFA (99.9:0.1); B, acetonitrile/TFA (99.9:0.1). Elution profile at 40 °C column temperature was as follows: A%: 0–5.5 min, linear gradient from 95 to 5%; 5.5–6 min, 5%; 6–6.1 min, gradient from 5 to 95%; 6.1–9.0 min, 95%.

Preparative reversed-phase HPLC

A Merck-Hitachi LaChrom HPLC was used for purification consisting of interface D-7000, pump L-7150, and detector L-7420 equipped with a Hypersil Gold C18 8-μm column (250 $\times$ 21.2 mm, Thermo Scientific): flow rate, 10 ml·min^{-1}; injection volume, 2.00 ml; detection at λ = 280 nm; solvents: A, water/acetonitrile/TFA (94.9:5:0.1); B, water/acetonitrile/TFA (5:94.9:0.1); gradient elution, A%: 0–5 min, 100%; 5–105 min, linear gradient from 100 to 0%; 105–110 min, gradient from 0 to 100%; 110–115 min, 100%.

HPLC MS (LC-MS)

Electrospray ionization (ESI) mass spectra were acquired employing an Agilent 6220 TOF mass spectrometer from Agilent Technologies (Santa Clara, CA) in extended dynamic range mode containing a Dual-ESI source, operating with a nitrogen generator, NGM 11. The mass axis was externally calibrated with ESI-L tuning mix (Agilent Technologies, Santa Clara, CA) as calibration standard. Samples were introduced from an HPLC outlet, and separation was performed with a 1200 HPLC system consisting of an autosampler, degasser, binary pump, column oven, and diode array detector (Agilent Technologies, Santa Clara, CA) using a C18 Hypersil Gold column (3 μm, 150 $\times$ 2.1 mm; Thermo Scientific). HPLC solvent A contained water/acetonitrile/formic acid (94.9:5:0.1); HPLC solvent B contained water/acetonitrile/formic acid (5:94.9:0.1). Gradient elution was as follows: A%: 0–10 min, gradient from 100 to 60%; 10–11 min, gradient from 60 to 2%; 11–11.5 min, gradient from 2 to 100%; 11.5–15 min, 100% with a flow rate of 0.3 ml·min^{-1} and column temperature of 40 °C.

NMR spectroscopy

NMR spectra were recorded on a Bruker Avance 500 (^{1}H: 500 MHz; ^{13}C: 126 MHz), Avance 500 HD (^{1}H: 500 MHz; ^{13}C: 126 MHz), or Avance 600 (^{1}H: 600.13 MHz; ^{13}C: 150.92 MHz) spectrometer. Chemical shifts (δ) were reported in ppm relative to tetramethylsilan (TMS) (δ_{TMS} = 0 ppm) and referenced to the residual nondeuterated solvent signal as internal standard (DMSO-d_6: ^{1}H: 2.50 ppm; ^{13}C: 39.5 ppm). Coupling constants (J) were reported in Hertz (Hz) with 0.05 Hz resolution. Multiplicities were described as singlet (s), doublet (d), triplet (t), quartet (q), or multiplet (m). The assignment of ^{13}C and ^{1}H

NMR signals was supported by two-dimensional NMR techniques (COSY, ROESY, and HSQC).

Accession numbers

The atomic coordinates and structure factors of apo-Thal, Trp-Thal, and apo-Thal-RebH5 have been deposited in the PDB as entries 6H43, 6H44, and 6IB5, respectively.

References

1. Schnepel, C., and Sewald, N. (2017) Enzymatic halogenation: a timely strategy for regioselective C-H activation. *Chemistry* **23,** 12064–12086 CrossRef Medline
2. Latham, J., Brandenburger, E., Shepherd, S. A., Menon, B. R. K., and Micklefield, J. (2018) Development of halogenase enzymes for use in synthesis. *Chem. Rev.* **118,** 232–269 CrossRef Medline
3. van Pée, K. H. (2001) Microbial biosynthesis of halometabolites. *Arch. Microbiol.* **175,** 250–258 CrossRef Medline
4. Payne, J. T., Andorfer, M. C., and Lewis, J. C. (2013) Regioselective arene halogenation using the FAD-dependent halogenase RebH. *Angew. Chem. Int. Ed. Engl.* **52,** 5271–5274 CrossRef Medline
5. Frese, M., Guzowska, P. H., Voss, H., and Sewald, N. (2014) Regioselective enzymatic halogenation of substituted tryptophan derivatives using the FAD-dependent halogenase RebH. *ChemCatChem* **6,** 1270–1276 CrossRef
6. Frese, M., and Sewald, N. (2015) Enzymatic Halogenation of Tryptophan on a Gram Scale. *Angew. Chem. Int. Ed. Engl.* **54,** 298–301 CrossRef Medline
7. Schnepel, C., Minges, H., Frese, M., and Sewald, N. (2016) A High-throughput fluorescence assay to determine the activity of tryptophan halogenases. *Angew. Chem. Int. Ed. Engl.* **55,** 14159–14163 CrossRef Medline
8. Poor, C. B., Andorfer, M. C., and Lewis, J. C. (2014) Improving the stability and catalyst lifetime of the halogenase RebH by directed evolution. *ChemBioChem* **15,** 1286–1289 CrossRef Medline
9. Hammer, P. E., Hill, D. S., Lam, S. T., van Pée, K. H., and Ligon, J. M. (1997) Four genes from *Pseudomonas fluorescens* that encode the biosynthesis of pyrrolnitrin. *Appl. Environ. Microbiol.* **63,** 2147–2154 Medline
10. Kirner, S., Hammer, P. E., Hill, D. S., Altmann, A., Fischer, I., Weislo, L. J., Lanahan, M., van Pée, K. H., and Ligon, J. M. (1998) Functions encoded by pyrrolnitrin biosynthetic genes from *Pseudomonas fluorescens*. **180,** 1939–1943
11. Keller, S., Wage, T., Hohaus, K., Hölzer, M., Eichhorn, E., and van Pée, K. H. (2000) Purification and partial characterization of tryptophan 7-halogenase (PrnA) from *Pseudomonas fluorescens. Angew. Chem. Int. Ed. Engl.* **39,** 2300–2302 CrossRef Medline
12. Heemstra, J. R., Jr., and Walsh, C. T. (2008) Tandem action of the O$_2$- and FADH$_2$-dependent halogenases KtzQ and KtzR produce 6,7-dichlorotryptophan for kutzneride assembly. *J. Am. Chem. Soc.* **130,** 14024–14025 CrossRef Medline
13. Smith, D. R. M., Uria, A. R., Helfrich, E. J. N., Milbredt, D., van Pée, K. H., Piel, J., and Goss, R. J. M. (2017) An unusual flavin-dependent halogenase from the metagenome of the marine sponge *Theonella swinhoei* WA. *ACS Chem. Biol.* **12,** 1281–1287 CrossRef Medline
14. Ortega, M. A., Cogan, D. P., Mukherjee, S., Garg, N., Li, B., Thibodeaux, G. N., Maffioli, S. I., Donadio, S., Sosio, M., Escano, J., Smith, L., Nair, S. K., and van der Donk, W. A. (2017) Two flavoenzymes catalyze the post-translational generation of 5-chlorotryptophan and 2-aminovinyl-cysteine during NAI-107 biosynthesis. *ACS Chem. Biol.* **12,** 548–557 CrossRef Medline
15. Agarwal, V., El Gamal, A. A., Yamanaka, K., Poth, D., Kersten, R. D., Schorn, M., Allen, E. E., and Moore, B. S. (2014) Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. *Nat. Chem. Biol.* **10,** 640–647 CrossRef Medline
16. Neubauer, P. R., Widmann, C., Wibberg, D., Schröder, L., Frese, M., Kottke, T., Kalinowski, J., Niemann, H. H., and Sewald, N. (2018) A flavin-dependent halogenase from metagenomic analysis prefers bromination over chlorination. *PLoS One* **13,** e0196797 CrossRef Medline
17. Zeng, J., and Zhan, J. (2010) A novel fungal flavin-dependent halogenase for natural product biosynthesis. *ChemBioChem* **11,** 2119–2123 CrossRef Medline
18. Zhu, T., Cheng, X., Liu, Y., Deng, Z., and You, D. (2013) Deciphering and engineering of the final step halogenase for improved chlortetracycline biosynthesis in industrial *Streptomyces aureofaciens. Metab. Eng.* **19,** 69–78 CrossRef Medline
19. Fraley, A. E., Garcia-Borràs, M., Tripathi, A., Khare, D., Mercado-Marin, E. V., Tran, H., Dan, Q., Webb, G. P., Watts, K. R., Crews, P., Sarpong, R., Williams, R. M., Smith, J. L., Houk, K. N., and Sherman, D. H. (2017) Function and structure of MalA/MalA′, iterative halogenases for late-stage C-H functionalization of indole alkaloids. *J. Am. Chem. Soc.* **139,** 12060–12068 CrossRef Medline
20. Chankhamjon, P., Tsunematsu, Y., Ishida-Ito, M., Sasa, Y., Meyer, F., Boettger-Schmidt, D., Urbansky, B., Menzel, K. D., Scherlach, K., Watanabe, K., and Hertweck, C. (2016) Regioselective dichlorination of a non-activated aliphatic carbon atom and phenolic bismethylation by a multifunctional fungal flavoenzyme. *Angew. Chem. Int. Ed. Engl.* **55,** 11955–11959 CrossRef Medline
21. Almeida, C., Pérez-Victoria, I., González-Menéndez, V., de Pedro, N., Martín, J., Crespo, G., Mackenzie, T., Cautain, B., Reyes, F., Vicente, F., and Genilloud, O. (2018) Non-geminal aliphatic dihalogenation pattern in dichlorinated diaporthins from *Hamigera fusca* NRRL 35721. *J. Nat. Prod.* **81,** 1488–1492 CrossRef Medline
22. Zeng, J., Lytle, A. K., Gage, D., Johnson, S. J., and Zhan, J. (2013) Specific chlorination of isoquinolines by a fungal flavin-dependent halogenase. *Bioorg. Med. Chem. Lett.* **23,** 1001–1003 CrossRef Medline
23. Menon, B. R. K., Brandenburger, E., Sharif, H. H., Klemstein, U., Shepherd, S. A., Greaney, M. F., and Micklefield, J. (2017) RadH: a versatile halogenase for integration into synthetic pathways. *Angew. Chem. Int. Ed. Engl.* **56,** 11841–11845 CrossRef Medline
24. Glenn, W. S., Nims, E., and O'Connor, S. E. (2011) Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. *J. Am. Chem. Soc.* **133,** 19346–19349 CrossRef Medline
25. Runguphan, W., Qu, X., and O'Connor, S. E. (2010) Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature* **468,** 461–464 CrossRef Medline
26. Dong, C., Flecks, S., Unversucht, S., Haupt, C., van Pée, K. H., and Naismith, J. H. (2005) Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science* **309,** 2216–2219 CrossRef Medline
27. Andorfer, M. C., and Lewis, J. C. (2018) Understanding and improving the activity of flavin-dependent halogenases via random and targeted mutagenesis. *Annu. Rev. Biochem.* **87,** 159–185 CrossRef Medline
28. Zhu, X., De Laurentis, W., Leang, K., Herrmann, J., Ihlefeld, K., van Pée, K. H., and Naismith, J. H. (2009) Structural insights into regioselectivity in the enzymatic chlorination of tryptophan. *J. Mol. Biol.* **391,** 74–85 CrossRef Medline
29. Seibold, C., Schnerr, H., Rumpf, J., Kunzendorf, A., Hatscher, C., Wage, T., Ernyei, A. J., Dong, C., Naismith, J. H., and van Pée, K. H. (2006) A flavin-

dependent tryptophan 6-halogenase and its use in modification of pyrrolnitrin biosynthesis. *Biocatal. Biotransform.* **24,** 401–408 CrossRef

30. Milbredt, D., Patallo, E. P., and van Pée, K. H. (2014) A tryptophan 6-halogenase and an amidotransferase are involved in thienodolin biosynthesis. *ChemBioChem* **15,** 1011–1020 CrossRef Medline

31. Zeng, J., and Zhan, J. (2011) Characterization of a tryptophan 6-halogenase from *Streptomyces toxytricini. Biotechnol. Lett.* **33,** 1607–1613 CrossRef Medline

32. Menon, B. R. K., Latham, J., Dunstan, M. S., Brandenburger, E., Klemstein, U., Leys, D., Karthikeyan, C., Greaney, M. F., Shepherd, S. A., and Micklefield, J. (2016) Structure and biocatalytic scope of thermophilic flavin-dependent halogenase and flavin reductase enzymes. *Org. Biomol. Chem.* **14,** 9354–9361 CrossRef Medline

33. Chang, F.-Y., and Brady, S. F. (2013) Discovery of indolotryptoline antiproliferative agents by homology-guided metagenomic screening. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 2478–2483 CrossRef Medline

34. Shepherd, S. A., Menon, B. R. K., Fisk, H., Struck, A. W., Levy, C., Leys, D., and Micklefield, J. (2016) A structure-guided switch in the regioselectivity of a tryptophan halogenase. *ChemBioChem* **17,** 821–824 CrossRef Medline

35. Yeh, E., Blasiak, L. C., Koglin, A., Drennan, C. L., and Walsh, C. T. (2007) Chlorination by a long-lived intermediate in the mechanism of flavin-dependent halogenases. *Biochemistry* **46,** 1284–1292 CrossRef Medline

36. Bitto, E., Huang, Y., Bingman, C. A., Singh, S., Thorson, J. S., and Phillips, G. N. (2008) The structure of flavin-dependent tryptophan 7-halogenase RebH. *Proteins* **70,** 289–293 Medline

37. Andorfer, M. C., Park, H. J., Vergara-Coll, J., and Lewis, J. C. (2016) Directed evolution of RebH for catalyst-controlled halogenation of indole C–H bonds. *Chem. Sci.* **7,** 3720–3729 CrossRef Medline

38. Belsare, K. D., Andorfer, M. C., Cardenas, F. S., Chael, J. R., Park, H. J., and Lewis, J. C. (2017) A simple combinatorial codon mutagenesis method for targeted protein engineering. *ACS Synth. Biol.* **6,** 416–420 CrossRef Medline

39. Payne, J. T., Poor, C. B., and Lewis, J. C. (2015) Directed evolution of Rebh for site-selective halogenation of large biologically active molecules. *Angew. Chem. Int. Ed. Engl.* **54,** 4226–4230 CrossRef Medline

40. Shepherd, S. A., Karthikeyan, C., Latham, J., Struck, A.-W., Thompson, M. L., Menon, B. R. K., Styles, M. Q., Levy, C., Leys, D., and Micklefield, J. (2015) Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. *Chem. Sci.* **6,** 3454–3460 CrossRef Medline

41. Lang, A., Polnick, S., Nicke, T., William, P., Patallo, E. P., Naismith, J. H., and van Pée, K. H. (2011) Changing the regioselectivity of the tryptophan 7-halogenase PrnA by site-directed mutagenesis. *Angew. Chem. Int. Ed. Engl.* **50,** 2951–2953 CrossRef Medline

42. Deleted in proof

43. Frese, M., Schnepel, C., Minges, H., Voss, H., Feiner, R., and Sewald, N. (2016) Modular combination of enzymatic halogenation of tryptophan with Suzuki-Miyaura cross-coupling reactions. *ChemCatChem* **8,** 1799–1803 CrossRef

44. Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I., Roessle, M., Cipriani, F., Fiedler, S., and Schneider, T. R. (2017) P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J. Synchrotron Radiat.* **24,** 323–332 CrossRef Medline

45. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 CrossRef Medline

46. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69,** 1204–1214 CrossRef Medline

47. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242 CrossRef Medline

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

49. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* **64,** 61–69 CrossRef Medline

50. Adams, P. D., Afonine, P. V., Bunkóczi, 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., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 CrossRef Medline

51. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501 CrossRef Medline

52. Kovalevskiy, O., Nicholls, R. A., Long, F., Carlon, A., and Murshudov, G. N. (2018) Overview of refinement procedures within *REFMAC* 5: utilizing data from different sources. *Acta Crystallogr. D Struct. Biol.* **74,** 215–227 CrossRef Medline

53. Holm, L., and Laakso, L. M. (2016) Dali server update. *Nucleic Acids Res.* **44,** W351–W355 CrossRef Medline

54. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31,** 3320–3323 CrossRef Medline

ASBMB

Structure-based switch of regioselectivity in the flavin-dependent tryptophan 6-halogenase Thal

Ann-Christin Moritzer, Hannah Minges, Tina Prior, Marcel Frese, Norbert Sewald and
Hartmut H. Niemann

Access the most updated version of this article at doi: 10.1074/jbc.RA118.005393

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 52 references, 3 of which can be accessed free at
http://www.jbc.org/content/294/7/2529.full.html#ref-list-1