

Struktur und Funktion der ersten bakteriellen Chalconisomerase  
und einer (*R*)-selektiven Amin-Transaminase

Inauguraldissertation

zur  
Erlangung des akademischen Grades eines  
Doktors der Naturwissenschaften (Dr. rer. nat.)

der  
Mathematisch-Naturwissenschaftlichen Fakultät  
der

Ernst-Moritz-Arndt-Universität Greifswald

vorgelegt von

Maren Thomsen

geboren am 23.01.1988

in Itzehoe

Greifswald, Dezember 2014

Dekan: Prof. Dr. Klaus Fesser

1. Gutachter: Prof. Dr. Winfried Hinrichs

2. Gutachter: Prof. Dr. Norbert Sträter

Tag der Promotion: 25. Februar 2015

## Inhaltsverzeichnis

	Seite
<b>1. Einleitung</b>	1
<b>2. Einleitung zu Chalconisomerasen</b>	2
2.1 Flavonoide	2
2.2 Pflanzliche Chalconisomerasen	3
2.3 Reaktionsmechanismus der pflanzlichen Chalconisomerasen	5
2.4 Degradation der Flavonoide durch die intestinale Mikroflora	6
<b>3. Ergebnisse zur Chalconisomerase aus <i>Eubacterium ramulus</i></b>	8
3.1 <i>De novo</i> Strukturlösung der bakteriellen CHI mittels SIRAS	8
3.2 Strukturanalyse der bakteriellen CHI aus <i>E. ramulus</i>	8
3.3 Aktives Zentrum und Reaktionsmechanismus	9
3.4 Bedeutung des Lids für die bakterielle CHI	10
3.5 Suche nach Faltungshomologen zur bakteriellen CHI	11
<b>4. Einleitung zu (<i>R</i>)-selektiven Amin-Transaminasen</b>	13
<b>5. Ergebnisse zur (<i>R</i>)-selektiven Amin-Transaminase aus <i>Aspergillus fumigatus</i></b>	18
5.1 Kristallisation und Röntgendiffraktion der ( <i>R</i> )-selektiven Amin-Transaminase	18
5.2 Strukturanalyse der ( <i>R</i> )-selektiven Amin-Transaminase	18
5.3 Untersuchungen zur dualen Substraterkennung der ( <i>R</i> )-selektiven Amin-Transaminase	20
<b>6. Literatur</b>	22
<b>7. Publikationsliste</b>	28

## 1. Einleitung

Proteine können eine Vielzahl an unterschiedlichen Funktionen ausüben. Darunter fallen Aufgaben wie die Übernahme von Stützfunktionen zum Aufbau des Zytoskeletts, andere übernehmen Transportprozesse und vermitteln den gerichteten Stofftransport in der Zelle. Wieder andere Proteine haben eine Signalübermittlungsfunktion und beeinflussen so Transkriptions-regulierende Proteine. Eine der interessantesten Gruppe unter den Proteinen aber sind die Enzyme.

Enzyme sind in der Lage biochemische Reaktionen durch eine Herabsetzung der Aktivierungsenergie zu katalysieren und stellen damit den Grundpfeiler unseres Stoffwechsels dar. Die Biotechnologie bedient sich dieser Enzyme zur Herstellung oder Verarbeitung vieler Produkte, da die Verwendung von Enzymen vielfältige Vorteile gegenüber chemischen Prozessen hat. So kann in Gegensatz zu vielen Metall-basierenden Katalysatoren durch die Verwendung von Enzyme bei Raumtemperatur, Atmosphärendruck und ohne organische Lösungsmittel gearbeitet werden. Dies ist nicht nur aufgrund der Energieeinsparungen weniger kostenintensiv, sondern auch weitaus schonender für die Umwelt.

Ein limitierender Faktor für die generelle Anwendung von Enzymen ist ihr meist begrenztes Substratspektrum. Um diesen Nachteil zu überwinden, können Methoden des *protein engineering* zum rationalen Design des aktiven Zentrums der Enzyme angewendet werden. Eine grundlegende Voraussetzung für die Anwendung dieser Methoden ist jedoch die Kenntnis der Proteinstruktur.

In dieser Arbeit konnten mittels Röntgenkristallographie die Strukturen und Funktionsbeziehungen von zwei biotechnologisch wichtigen Enzymen aufgeklärt werden. Der erste Teil der Arbeit beschreibt die Aufklärung der Proteinstruktur wie auch des Reaktionsmechanismusses der ersten bakteriellen Chalconisomerase. Dieses Enzym ist ein wichtiger Bestandteil der Flavonoiddegradation und kann biotechnologisch zur Umsetzung und Modifizierung von Flavonoiden benutzt.

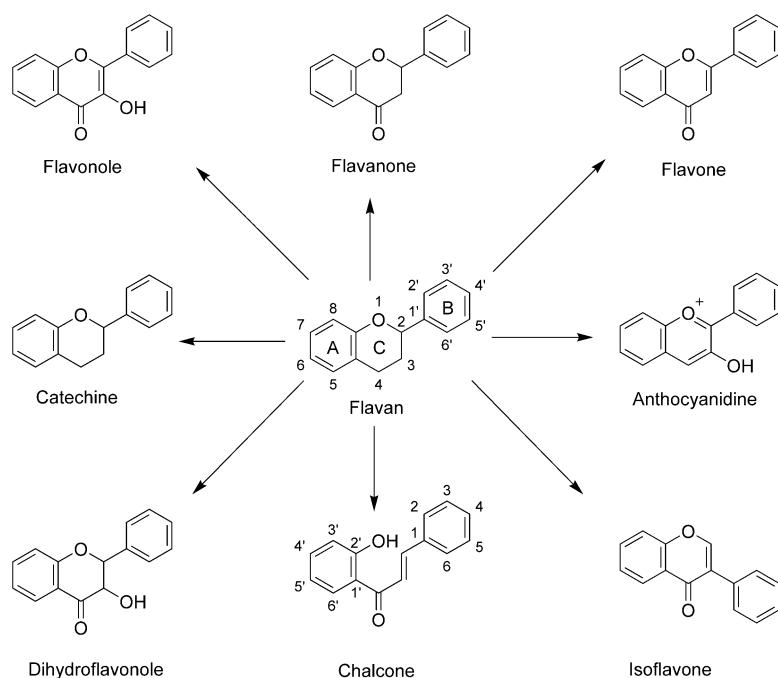
Die Aufklärung der Struktur-Funktionsbeziehung einer (*R*)-selektiven Amin-Transaminase wird im zweiten Teil dieser Arbeit behandelt. Es konnten grundlegende Kenntnisse über die Architektur des aktiven Zentrums und zum Mechanismus der dualen Substraterkennung gewonnen werden. Diese Erkenntnisse bilden die Grundlage für weiterführende Untersuchungen zur Erweiterung des Substratspektrums und deren Anwendung in der Biotechnologie.

## 2. Einleitung zu Chalconisomerasen

### 2.1 Flavonoide

Schon im Altertum wussten die Menschen um die heilbringenden Eigenschaften bestimmter Pflanzenextrakte bzw. -bestandteile<sup>1</sup>. Diese beruhen auf einer Vielzahl von pflanzlichen Sekundärmetaboliten mit gesundheitsprotektiven Eigenschaften<sup>2</sup>. Eine wichtige Klasse dieser Sekundärmetabolite stellen die Flavonoide dar. Sie wurden bereits in den 1930er Jahren vom Nobelpreisträger Albert von Szent-Györgyi Nagyrápolt entdeckt und zuerst als Vitamin P bezeichnet<sup>3</sup>.

Die Struktur der Flavonoide beruht auf dem Grundgerüst Flavan<sup>4</sup>, das aus zwei aromatischen Ringen A und B und einem heterozyklischen C-Ring aufgebaut ist (Abbildung 2.1). Dieses Grundgerüst kann durch eine Reihe von möglichen Modifikationen, wie z. B. Hydroxylierung, Prenylierung und Glykosylierung, variiert werden. Mittlerweile sind etwa 10.000 Flavonoide isoliert und beschrieben worden<sup>5</sup>. In der Pflanze liegen die Flavonoide zur Verbesserung ihrer Löslichkeit meist als Glycoside vor. Das Hauptvorkommen der Aglycone und methylierter Flavonoide ist in der Rinde und dem Wachs auf Blättern<sup>6</sup> zum Schutz vor Fraßfeinden. Aufgrund der reinen Größe dieser Stoffklasse ist die Art und Anzahl ihrer Eigenschaften und potentiellen Anwendungsbereiche sehr vielfältig.



**Abbildung 2.1:** Das Flavan-Grundgerüst und einige daraus entstehende Subklassen der Flavonoide<sup>4</sup>.

In der Lebensmittelindustrie bereits verwendete Flavonoide sind unter anderem Anthocyane (3-Monoglykoside der Anthocyanidine). Diese weisen eine intensive gelbe Farbe auf und werden daher als Ersatz für synthetisch hergestellte Farbstoffe verwendet<sup>7,8</sup>.

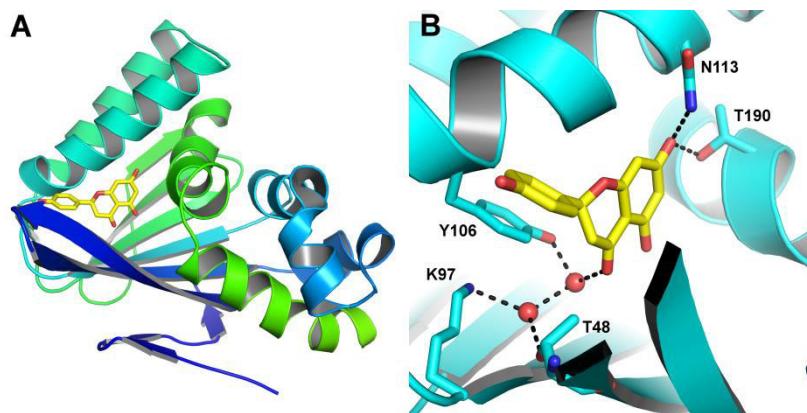
Neben ihren farblichen Eigenschaften werden Flavonoide, wie z.B. Quercetin, auch als Nahrungsergänzungsmittel aufgrund ihrer antioxidativen Wirkung vermarktet und es gibt zahlreiche Studien zur Untersuchung ihrer gesundheitsprotektiven Eigenschaften. Bisher beschrieben wurden Eigenschaften wie antioxidative<sup>9</sup>, antithrombische<sup>10</sup>, antilipoperoxidierende<sup>11</sup>, antiinflammatorische<sup>12, 13</sup>, anticancerogene<sup>14, 15</sup>, antifungizide<sup>16</sup>, antidiabetische<sup>17</sup>, antibakterielle<sup>18</sup>, antivirale<sup>19</sup>, Östrogenmangel-ausgleichende Effekte<sup>20</sup> sowie protektive Eigenschaften gegenüber kardiovaskulären Erkrankungen<sup>10</sup>. Diese sehr breit gefächerten physiologisch und pharmazeutisch relevanten Eigenschaften beruhen auf drei grundlegenden Wirkungsmechanismen der Flavonoide: der Beseitigung freier Radikale insbesondere reaktiver Sauerstoffspezies, der Chelatierung und folglich Maskierung von Übergangsmetallen und der Inhibierung spezifischer Enzyme<sup>21</sup>. Unter letzteren befinden sich vor allem Cytochrom P450-Monooxygenasen der Leber, die für die Degradation von Xenobiotika verantwortlich sind<sup>22, 23</sup>. So ist Hesperitin ein Inhibitor für die humane Cytochrom P450-Monooxygenase CYP1B1<sup>24</sup> und kann aufgrund dessen die Wirkungsweise und den Katabolismus einiger Medikamente entscheidend beeinflussen.

## 2.2 Pflanzliche Chalconisomerasen

Die Biosynthese der Flavonoide in Pflanzen ist bereits umfassend untersucht worden. Das erste Enzym im Biosynthese-Weg, die Chalcon synthase<sup>25, 26</sup> (CHS; EC 2.3.1.74) aus der Familie der Polyketidsynthasen<sup>27</sup>, katalysiert über mehrere aufeinander folgende Kondensationsreaktionen die Bildung des Naringeninchalcons aus drei Molekülen Malonyl-CoA und einem Coumaroyl-CoA. Das gebildete Naringeninchalcon wird anschließend über eine intramolekulare oxa-Michael-Addition selektiv zum (2S)-Naringenin durch eine Chalconisomerase (CHI; EC 5.5.1.6) zyklisiert. Die große Diversität der Flavonoide wird im weiteren Verlauf der Biosynthese durch die Einführung unterschiedlichster Modifikationen basierend auf dem (2S)-Naringenin erreicht. Diese Modifikationen variieren von einfachen Glykosylierungen, Methylierungen, Hydroxylierungen und Prenylierungen bis zur komplexen Phenylringverschiebung zur Bildung der Isoflavonoide<sup>28</sup>.

Bereits 1967 konnte die erste pflanzliche Chalconisomerase aus *Soja hispida* isoliert werden<sup>29</sup>. Artübergreifend weisen pflanzliche Chalconisomerasen mit Sequenzidentitäten von 50 – 82 % zwar eine hohe Homologie untereinander auf, doch ist diese Enzymklasse aufgrund ihrer großen Anzahl an Vertretern sehr unterschiedlich im Hinblick auf kinetische Parameter und Substratspezifität<sup>30</sup>. Die meisten Pflanzen besitzen außerdem mehrere Isoenzyme der CHI, deren Expressionsmuster in Abhängigkeit ihrer Lokalisation in der Pflanze, z.B. Wurzel oder Staubbeutel, stark variieren können<sup>31</sup>.

Die Struktur der pflanzlichen Chalconisomerase von *Medicago sativa* ist seit dem Jahr 2000 bekannt<sup>32</sup>. Pflanzliche CHIs liegen vorwiegend als Monomere vor und bestehen aus ca. 220 Aminosäuren. Die Faltung des Enzyms beruht auf einem *open-faced β-sandwich fold* (Abbildung 2.2A). Dies wird aus einem großen β-Faltblatt mit sechs antiparallelen β-Strängen und einem auf ihm liegenden Bündel von sieben α-Helices gebildet. Ein weiteres kleines β-Faltblatt, aus drei kleinen β-Strängen gebildet, formiert sich auf der anderen Seite. Diese Art der Proteinfaltung war lange Zeit nur für pflanzliche Chalconisomerasen bekannt.



**Abbildung 2.2:** A: Bänderdarstellung der pflanzlichen CHI von *Medicago sativa* (PDB: 1eyq). B: Koordination des (2S)-Naringenins im aktiven Zentrum<sup>32</sup>.

Eine Theorie zur Entwicklung des pflanzlichen Sekundärstoffwechsels basiert auf der Hypothese, dass jedes Enzym aus dem Sekundärstoffwechsel aus einem Vorläuferprotein aus dem Primärstoffwechsel entstand<sup>33</sup>. Doch bis vor kurzem waren nur homologe CHI-Proteine aus Pflanzen bekannt und ein Bezug zum Primärstoffwechsel fehlte vollkommen. Aufgrund ihrer dominanten Rolle in der Flavonoid-Biosynthese entstand so die Hypothese, dass die Anwesenheit eines CHI-Gens als Genmarker zur Klassifizierung für das Reich der Pflanzen angesehen werden könnte<sup>32</sup>. Gensheimer & Mushegian widerlegten jedoch mit einer Sequenzdatenbankanalyse und der Identifizierung homologer CHI-Proteinsequenzen in Pilzen und Bakterien diese Hypothese<sup>34</sup>. Bemerkenswerterweise besitzen diese Organismen jedoch entweder ein CHI-Gen oder ein CHS-Gen und die katalytisch aktiven Reste sind nicht konserviert, sodass davon ausgegangen werden kann, dass die Enzyme in diesen Organismen nicht für die Flavonoid-Biosynthese verantwortlich sind. Dies ließ die Frage nach ihrer physiologischen Rolle weiterhin unbeantwortet<sup>34</sup>. Erst 2012 konnte für die pflanzliche CHI ein Bezug zum Primärstoffwechsel hergestellt werden<sup>35</sup>. Ngaki et al. gelang die Aufklärung Fettsäure-bindender Proteine (*fatty acid binding protein*; FAP) aus *Arabidopsis thaliana* mit einem CHI-fold<sup>35</sup>. Die physiologische Rolle dieser Proteine in Pflanzen ist noch nicht vollständig aufgeklärt, doch zeigen Mutanten signifikante Einschränkungen der Reproduktivität, was ein wichtiges Merkmal für Enzyme aus dem Primärmetabolismus ist. Gemeinsam mit der

Chalconsynthase<sup>27</sup>, einer Polyketidsynthase, setzt dies die Flavonoidbiosynthese in eine direkte evolutionäre Beziehung zu der Fettsäurebiosynthese<sup>35</sup>.

### 2.3 Reaktionsmechanismus der pflanzlichen Chalconisomerasen

Das aktive Zentrum der pflanzlichen CHI befindet sich zwischen den Helices  $\alpha 4$ ,  $\alpha 6$  und den  $\beta$ -Strängen  $\beta 3a$  und  $\beta 3b$  (Abbildung 2.2B). Es weist eine eher hydrophobe Umgebung mit wenigen polaren Wechselwirkungen zwischen dem Substrat Naringeninchalcon und dem Enzym auf. Die Koordinierung beruht hauptsächlich auf Wasserstoffbrücken der Reste Asn113 und Thr190 mit der 4'-Hydroxylgruppe des Substrats. Die Carbonylfunktion wird über ein Wassermolekül von Tyr106 gebunden. Jez et al. nehmen an, dass die Katalyse über dieses Wassermolekül verläuft<sup>32, 36, 37</sup>. Der von Jez et al. postulierte Reaktionsmechanismus geht davon aus, dass die 2-Hydroxylgruppe, mit einem  $pK_a \approx 7 - 8$ , in Wasser als Phenolat vorliegt. Dies greift dann in einer intramolekularen oxa-Michael-Addition die Kohlenstoff-Kohlenstoff-Doppelbindung in  $\beta$ -Position zur Carbonylgruppe an. Das daraufhin gebildete Enolat wird über das katalytisch aktive Wassermolekül, das als generelle Säure agiert und so eine Oxyanion-Bindetasche bildet, durch vorübergehende Protonierung stabilisiert. Über Tautomerie kommt es dann zur Produktbildung des (2S)-Naringenins. Die pflanzliche CHI bildet ausschließlich das (S)-Enantiomer, das das alleinige physiologisch aktive Enantiomer darstellt<sup>32</sup>.

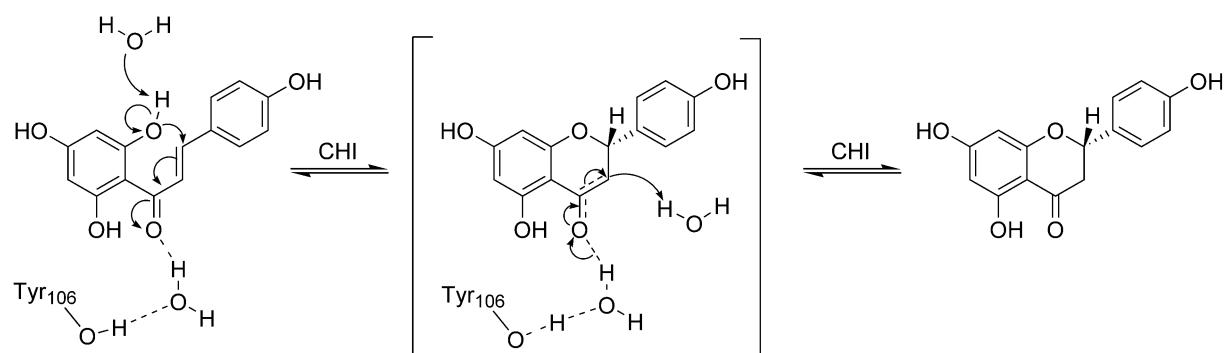


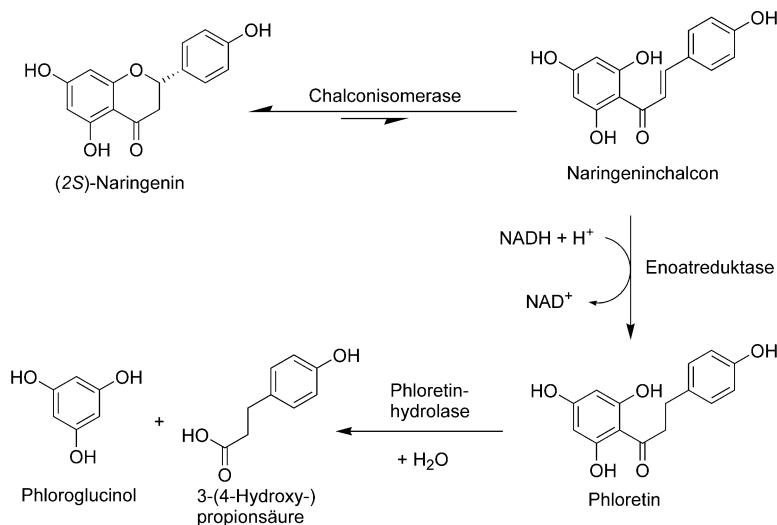
Abbildung 2.3: Postulierter Reaktionsmechanismus der pflanzlichen Chalconisomerase<sup>32</sup>.

Zur Untersuchung des Reaktionsmechanismus wurde von Jez et al. das Tyr106 zu einem Phenylalanin mutiert. Es stellte sich heraus, dass die Reaktionsgeschwindigkeit zwar 100-fach reduziert, jedoch die Restaktivität noch bedeutend größer war, als die spontane, nicht-enzymatische Zyklisierung des Naringeninchalcons in Wasser. Jez et al. führen dies auf eine Konformations-vermittelte Katalyse, ähnlich dem Katalysemechanismus der Chorismat-Mutase<sup>38</sup> oder dem Prinzip von katalytischen Antikörpern<sup>39</sup>, zurück. Das Substrat wird dabei durch die Form des aktiven Zentrums in einen Übergangszustand ähnliche Konformation gebracht, sodass die Energiebarriere zur Reaktion katalytisch herabgesetzt wird.

## 2.4 Degradation der Flavonoide durch die intestinale Mikroflora

Über die tägliche Nahrung nehmen wir eine Vielzahl von pflanzlichen Sekundärstoffen auf. Der Ernährungsbericht des Jahres 2008 der Deutschen Gesellschaft für Ernährung e.V. (DGE) geht von schätzungsweise 1,5 g pro Tag, bei einer vegetarischen Ernährungsweise sogar von einer deutlich höheren Aufnahme, aus<sup>40</sup>. Aufgrund der hohen Bioverfügbarkeit der Flavonoide beträgt ihr Anteil in unserer Nahrung ca. 50 – 200 mg täglich<sup>40</sup>. Nur ein geringer Anteil der Flavonoide wird im Dünndarm über Glucosetransporter<sup>41</sup> oder als Aglykon über die Zellmembran resorbiert. Auch wenn bereits seit 1965 die Beteiligung der intestinalen Mikroflora an der Umsetzung der Flavonoide bekannt war, war die Identifizierung der beteiligten Mikroorganismen eine langwierige Arbeit. 1987 wurde von Bokkenheuser et al. bewiesen, dass die obligat anaeroben Bakterien der Gattungen *Eubacterium*, *Clostridium* und *Bacteroides* der intestinalen Mikroflora an der Deglykosylierung der Flavonoide<sup>42</sup> beteiligt sind. *Eubacterium ramulus* ist hier von besonderer wissenschaftlicher Bedeutung, da es mit Quercetin-3-glucosid als einziger C-Quelle überleben kann<sup>43</sup>. Schneider et al. isolierten und quantifizierten 1999 das Auftreten dieses wichtigen Darmbakteriums im humanen Fäces<sup>44</sup>. In nachfolgenden Studien zu diesem Bakterium bestimmte die Gruppe um Blaut weitere degradierbare Flavonoide und einige ihrer Abbauprodukte<sup>45-47</sup>. Zur Untersuchung des Degradationsweges der Flavanone wurde Naringenin, das in der Natur am häufigsten vorkommende Flavanon, als alleinige Kohlenstoffquelle benutzt. Über das dabei entstehende Zwischenprodukt Phloretin konnten Schneider et al. vor allem einen Anstieg der Phloroglucinol- und 3-*para*-Hydroxypropionsäure-Konzentration detektieren<sup>45</sup>. Dies lässt folglich darauf schließen, dass die Degradation mit einer Aufspaltung des heterozyklischen C-Rings beginnt. Phloroglucinol wird folgend von einer Phloroglucinol-Reduktase zu Dihydrophloroglucinol reduziert und dieses kann zu Essigsäure und Buttersäure katabolisiert werden<sup>48</sup>. Eine Degradation von Naringin, dem 7-O-Rhamnoglukosid des Naringenins, konnte nicht beobachtet werden. Zwar wurden im Zellextrakt von *E. ramulus* verschiedene  $\alpha$ - und  $\beta$ -Glucosidasen nachgewiesen, jedoch keine  $\alpha$ -Rhamnosidasen<sup>45</sup>, sodass Schneider et al. von einer vorherigen Abspaltung der Zuckerreste durch andere Darmbakterien oder unspezifischer humaner Hydrolasen im Darm ausgehen. Demgegenüber zeigte sich das Bakterium tolerant gegenüber zusätzlicher 3'-Hydroxylierung und ist auch in der Lage Eriodictyol abzubauen. Schneider et al. konnten ebenfalls zeigen, dass die Carbonylfunktion für den Abbau essentiell ist, da Catechine nicht umgesetzt wurden<sup>45</sup>. Basierend auf den detektierten Abbauprodukten postulierten Schneider et al. einen Degradationsweg für das Flavanon Naringenin ohne Kenntnis der verantwortlichen Enzyme<sup>45</sup> (Abbildung 2.4). Zur Verifizierung des postulierten Flavanon-Degradationsweges isolierte die Gruppe um Blaut in der nachfolgenden Zeit einige der katalysierenden Enzyme aus *E. ramulus*. Als erstes Enzym konnten Schöfer et al. die Phloretinhydrolase isolieren und charakterisieren<sup>49</sup>. Im Jahr 2004 gelang dann Herles et al. die Isolierung der ersten bakteriellen

Chalconisomerase aus *E. ramulus*<sup>50</sup>. Auch wenn es Herles nur möglich war 15 Aminosäuren der N-terminalen Sequenz zu identifizieren, zeigten ihre biochemischen Untersuchungen doch bereits große Unterschiede zur pflanzlichen Chalconisomerase.



**Abbildung 2.4:** Degradationsweg von (2S)-Naringenin durch *Eubacterium ramulus* (postuliert von Schneider et al.<sup>45</sup>. CHI identifiziert von Herles et al.<sup>50</sup>, Phloretinhydrolase identifiziert von Schoefer et al.<sup>49</sup>, Enoatreduktase identifiziert und gesamter Degradationsweg bewiesen durch Gall et al.<sup>51</sup>).

Zur Identifizierung der vollständigen Gensequenz der bakteriellen CHI wurde das gesamte Genom von *E. ramulus* sequenziert<sup>51</sup>. Über eine anschließende Suche der von Herles et al.<sup>50</sup> bestimmten N-terminalen Sequenz konnte die bakterielle CHI von Gall et al.<sup>51</sup> identifiziert werden. Jedoch fehlten zu diesem Zeitpunkt immer noch Hinweise zur Identität des Enzyms, das den zweiten Schritt katalysiert. Ausgehend von der Architektur des Substrates schien eine Enoatreduktase naheliegend. Über eine Suche nach dem spezifischen Nukleotid-Bindemotiv für Enoatreduktasen GXGXXG(X)<sub>17</sub>E konnte eine anaerobe Enoatreduktase im Genom von *E. ramulus* identifiziert werden<sup>51</sup>. Durch die simultane Expression beider rekombinanter Enzyme in einem *E.-coli*-Stamm konnte der oben postulierte Abbauweg bewiesen werden<sup>51</sup>.

Die Sequenzidentität der bakteriellen CHI zur pflanzlichen Chalconisomerase oder anderen charakterisierten Proteine beträgt weniger als 10 %, sodass eine strukturelle Untersuchung dieses Enzyms für zukünftige biotechnologische Anwendungen erforderlich erschien.

### 3. Ergebnisse zur Chalconisomerase aus *Eubacterium ramulus*

#### 3.1 De novo Strukturlösung der bakteriellen CHI mittels SIRAS

Da schon die Proteinsequenz keine Homologie zu bekannten Sequenzen aufwies, waren keine homologen Proteinstrukturen bekannt und eine Strukturlösung über *Molecular Replacement* kam somit nicht in Betracht. Stattdessen wurden sowohl native als auch mit Seleno-Methionin derivatisierte<sup>52</sup> Proteinkristalle der CHI hergestellt. Die Diffraktionsqualität der Kristalle konnte mittels *streak seeding* mit Katzenschnurrhaaren<sup>53</sup> signifikant verbessert werden. Es wurden jeweils Datensätze unter Kryobedingungen an der Absorptionskante des Selens am Synchrotron BESSY II<sup>54</sup> (Berlin, Deutschland) aufgenommen. Eine Lösung des Phasenproblems war somit anhand der *single isomorphous replacement with anomalous scattering*-Methode (SIRAS) möglich. 54 Selenpositionen konnten identifiziert werden, bei 9 möglichen Selenmodifikationen innerhalb der Proteinsequenz entspricht dies 6 Polypeptidketten in der asymmetrischen Einheit.

#### 3.2 Strukturanalyse der bakteriellen CHI aus *E. ramulus*

Die bakterielle CHI kristallisiert in der orthorhombisch innenzentrierten Raumgruppe  $I2_12_12_1$ . Das Strukturmodell der nativen CHI konnte bei einer Auflösung von 1,8 Å zu einem finalen R/R<sub>free</sub>-Faktor von 13,4 %/15,4 % verfeinert werden. Die asymmetrische Einheit beinhaltet sechs Monomere die eine  $D_3$ -Symmetrie, also zwei Trimeren aus drei Dimeren, bilden (siehe Publikation I, Abbildung 1). PISA<sup>55</sup> Analyse und SAXS Experimente (ausgewertet von Dr. A. Tuukkanen, Gruppe Svergun, BioSAXS, EMBL, Außenstation Hamburg, DESY) verifizierten die hexamere Quartärstruktur als native Form in Lösung. Verantwortlich für die Oligomerisierung ist unter anderem ein wichtiger C-terminaler Sequenzabschnitt (267 – 282), dessen letzte vier Aminosäuren (279 – 282) in den Kern des anderen Monomers gebunden sind und so die Oberfläche der Dimerisierungsinteraktion drastisch erhöhen (687 Å<sup>2</sup> von 2102 Å<sup>2</sup>). Die Oberfläche einer Trimerisierungsinteraktion beträgt hingegen nur 910 Å<sup>2</sup>. Strukturell kann die bakterielle CHI in die Superfamilie der dimeren  $\alpha+\beta$  barrel-Proteine eingeordnet werden. Die Struktur des Monomers kann in zwei Domänen mit sehr ähnlicher Faltung unterteilt werden, die katalytische Domäne (1 – 143) und die *solvent exposed* Domäne (144 – 266; siehe Publikation II, Abbildung 1). Die katalytische Domäne ist sowohl an der Dimer- als auch an der Trimerbildung beteiligt, während die *solvent exposed* Domäne mit Ausnahme des oben erwähnten C-Terminus vollständig offen zum Solvens liegt. Während die Elektronendichte für den gesamten N- und C-Terminus deutlich erkennbar ist, war eine Identifizierung der Reste 108 – 130 nicht möglich. Diese Aminosäuren bilden eine für diese Faltungsklasse außergewöhnliche Lid-Struktur, welche

aufgrund ihrer hohen Flexibilität nur in der geschlossenen Konformation in der Elektronendichte interpretierbar ist. Näheres hierzu ist in den Kapiteln 3.3 und 3.4 zu finden.

Beide Domänen haben eine Ferredoxin-ähnliche Faltung basierend auf zwei antiparallelen  $\beta$ -Faltblättern mit 4 bzw. 5  $\beta$ -Strängen und zwei auf ihnen liegenden großen  $\alpha$ -Helices<sup>56</sup>. Das  $\beta$ -Faltblatt der katalytischen Domäne hat eine  $2 \uparrow 3 \downarrow 1 \uparrow 4 \downarrow$  Anordnung der Stränge, die  $\beta$ -Stränge in der *solvent exposed* Domäne präsentieren einen Aufbau mit  $5 \uparrow 7 \downarrow 8 \uparrow 6 \downarrow 9 \uparrow$ . Die  $\beta$ -Faltblätter beider Domänen formieren sich zu einem übereinander gelagerten  $\beta$ -Faltblatt-Sandwich, dessen  $\beta$ -Faltblätter in der Ebene um 90° gegeneinander rotiert sind. Die bereits oben erwähnten vier letzten Reste des C-Terminus werden zwischen den beiden  $\beta$ -Faltblättern koordiniert und zwingen das  $\beta$ -Faltblatt-Sandwich so eine V-Form einzunehmen (siehe Publikation II, SI-Abbildung 3).

Weitergehende von H. Kratzat in ihrer Bachelor-Arbeit durchgeführte Untersuchungen<sup>57</sup> zeigen, dass sowohl die Deletion des Lids als auch die Mutation des sich in einer *cis*-Peptidbindung befindenden Pro250 zu Alanin keine signifikanten Auswirkungen auf die thermische Stabilität des Enzyms haben. Eine Deletion des C-Terminus (Aminosäurereste Tyr278 – Arg282) hingegen führt zu einer inaktiven Variante mit deutlich verringelter thermischer Stabilität<sup>57</sup>.

### 3.3 Aktives Zentrum und Reaktionsmechanismus

Da wie bereits oben erwähnt keine homologen Proteinsequenzen zur bakteriellen CHI bekannt waren, gab es keine Informationen über das aktive Zentrum und potenziell katalytisch relevante Aminosäurereste. Zur Identifizierung des aktiven Zentrums wurde die CHI mit dem Substrat Naringeninchalcon cokristallisiert. Unter Kryobedingungen konnte am Synchrotron BESSY II (Berlin, Deutschland) ein Datensatz mit einer Auflösung von 2,0 Å aufgenommen werden. In dem über *Molecular Replacement* gelösten Modell des Ligandkomplexes befindet sich ein Trimer (Monomere A, C und E) mit geschlossenen Lid-Konformationen und mit je einem Molekül (2S)-Naringenin komplexiert. Die Elektronendichte für den Lid ist in der geschlossenen Form deutlich erkennbar. Das andere Trimer (Monomere B, D und F) weist hingegen offene Formen des Lids auf. Hier sind im Eingangsbereich zum aktiven Zentrum aufgrund der artifiziell hohen Naringeninchalcon-Konzentration bei der Kristallisation mehrere unterschiedliche Konformationen von Naringenin und Naringeninchalcon in einer komplizierten Fehlordnung gebunden (siehe Publikation II, SI-Abbildung 5).

Das Substrat bindet in der katalytischen Domäne zwischen den  $\beta$ -Strängen 1 und 3 und den  $\alpha$ -Helices 1, 3 und 4. Mit Distanzen zwischen 3,2 – 2,5 Å zum Liganden übernehmen die Reste Gln40, Gln69, Thr71, Asp79 und Gln101 die Koordinierung (siehe Publikation II, Abbildung 2). Für die Untersuchung der enzymatischen Aktivität konnten drei mögliche Aminosäurereste His33, His73 und Tyr48 identifiziert werden, die in einem katalytisch relevanten Abstand zum Substrat lagen. Durch Mutagenese zu His33Ala, His73Ala und zu Tyr48Phe sollte über einen darauffolgenden Verlust der enzymatischen Aktivität der katalytisch aktive Rest ermittelt werden. Während die Mutante His73Ala keine Einschränkung in der Aktivität zeigte, war die Aktivität der Tyr48Phe um etwa ein Drittel gegenüber dem Wildtyp reduziert. Der größte Aktivitätsverlust, mit 2000-fach geringerer Aktivität, wurde jedoch für die His33Ala Mutante gemessen. Da diese Mutante jedoch immer noch eine geringe Aktivität zeigte, wurde die Hypothese aufgestellt, dass ein Wassermolekül den durch die eingeführte Mutation gewonnenen Platz einnehmen und die Katalyse vermitteln kann. Aufgrund dessen wurde noch die Mutante His33Gln hergestellt. Der Glutaminrest weist vergleichbare sterische Ansprüche wie Histidin auf, kann jedoch keine Säure-Base-vermittelte Reaktion katalysieren. Tatsächlich zeigte diese Mutante eine 8,3-fach geringere Aktivität im Vergleich zur His33Ala Mutation. Eine Wiederherstellung der katalytischen Aktivität durch die Mutation zu einer protischen Aminosäure wie Glutamat (His33Glu) war jedoch nicht möglich. Basierend auf diesen Ergebnissen wurde ein Säure-Base-vermittelter Reaktionsmechanismus für den Abbau von Naringenin zu Naringeninchalcon postuliert. Initiiert wird die Reaktion durch die Deprotonierung am C3. Das daraufhin gebildete Enolat als Zwischenprodukt wird von Tyr48 und Thr71, die als Oxyanion-Bindetasche fungieren, stabilisiert. Die nachfolgende Protonierung am O2' begünstigt die Ringöffnung des heterozyklischen C-Ringes und das Chalcon wird gebildet. Um diese Hypothese zu beweisen, wurde die Reaktion in  $^2\text{H}_2\text{O}$  durchgeführt und das gebildete Produkt in Kooperation mit J. Dickerhoff von der Arbeitsgruppe Weisz (Analytische Biochemie, Universität Greifswald) via  $^1\text{H-NMR}$  untersucht. Die Ergebnisse belegen eindeutig, dass die Protonierung/Deprotonierung in *pro-S*-Konfiguration am C3 erfolgt. Dies stimmt mit den Erkenntnissen aus der Struktur mit gebundenen (2S)-Naringenin überein, da das axiale H-Atom am C3 direkt auf das His33Ne zeigt (siehe Publikation II, Abbildung 3).

### 3.4 Bedeutung des Lids für die bakterielle CHI

In allen gelösten Strukturen mit offener Lid-Konformation (ohne Substratbindung) hat jedes CHI Monomer eine Solvens-exponierte Kluft. Bei geschlossenen Lid-Konformationen, ob mit oder ohne Substratbindung, ist das aktive Zentrum vom Solvens abgeschirmt. Dies könnte den Vorteil haben, dass das Wasser-instabile Produkt Naringeninchalcon vor der Selbstisomerisierung geschützt wird<sup>58</sup>.

Ein Vergleich der offenen mit der geschlossenen Lid-Konformation zeigt, dass die Konformationsänderung ausschließlich auf den Bereich des Lids begrenzt ist und keine Veränderungen im Rückgrat des Proteins geschehen. Lediglich eine Änderung von Seitenkettenkonformationen ist zu beobachten. So bewirkt die Schließung des Lids, dass das C $\alpha$  des Arg125 in einen so dichten Abstand zum aromatischen Ring des Phe135 kommt, dass dieser um 90° rotieren muss. Dies bewirkt nachfolgend eine Verschiebung des Glu91 aufgrund dessen einer Salzbrückenbildung mit Arg125 erleichtert wird. Ein weiterer wichtiger Aminosäure-Rest zur Bestimmung der Lid-Konformation ist Glu131. Dieser Rest bildet in der offenen Konformationen eine starke Salzbrücke zu Lys253 aus (2,8 Å). Kommt es zur Bildung der geschlossenen Form ist dieser Rest Solvens-exponiert und nicht mehr in der Elektronendichte identifizierbar. Diese Konformationsänderungen können auch in den Strukturen mit geschlossenen Lid-Konformationen ohne Substratbindung beobachtet werden und sind somit nicht auf die Substratbindung zurückzuführen, sondern beruhen ausschließlich auf der Lid-Schließung. Zur Überprüfung dieser Hypothese wurden in Kollaboration mit der Gruppe Svergun (BioSAXS, EMBL, Außenstation Hamburg, DESY) *small angle X-ray scattering* (SAXS) Experimente durchgeführt. Die Streukurve des nativen Enzyms ohne Substratanbindung konnte nur an ein Modell unter der Annahme einer zufälligen Zusammensetzung offener und geschlossener Lid-Konformationen angepasst werden (siehe Publikation II, Abbildung 4). Dies zeigt, dass in Lösung ein dynamisches Gleichgewicht zwischen offenen und geschlossenen Lid-Konformationen vorliegt. Schlussfolgernd kann demnach von einer sehr geringen Energiebarriere zur Schließung des Lids ausgegangen werden. Eine vorausgehende Substratbindung scheint nicht essentiell zu sein. Darüber hinaus konnte gezeigt werden, dass bei Substratanbindung eine sehr viel kompaktere, globulärere Quartärstruktur vorliegt.

### 3.5 Suche nach Faltungshomologen zur bakteriellen CHI

Zur Untersuchung ob trotz fehlender Proteinsequenzhomologie eine konservierte Faltungshomologie zwischen der pflanzlichen CHI von *Medicago sativa* (PDB: 1eyq)<sup>32</sup> und der bakteriellen CHI besteht, wurde eine Überlagerung der Sekundärstrukturelemente vorgenommen. Da beide Strukturmodelle auf einem antiparallelen  $\beta$ -Faltblatt mit zwei darauf liegenden großen  $\alpha$ -Helices basieren, ist eine Überlagerung auf den ersten Blick möglich. Dieses  $\beta$ -Faltblatt- $\alpha$ -Helix-Arrangement ist jedoch ein sehr häufiges Motiv der Proteinfaltung und bei genauer Betrachtung der Topologien der beiden Proteine ist zu erkennen, dass die Verbindungen zwischen den einzelnen  $\beta$ -Strängen sehr unterschiedlich sind. Dies macht es sehr unwahrscheinlich, dass diese beiden Enzyme einen engen evolutionären Hintergrund besitzen. Dementsprechend konnte auch zu dem postulierten

Vorläuferprotein der pflanzlichen CHIs, dem Fettsäure-bindenden-Protein (FAP) von *Arabidopsis thaliana* (PDB: 4doi)<sup>35</sup>, keine Faltungshomologie abgeleitet werden.

Für eine weitere Suche zur Einordnung der bakteriellen CHI in einen evolutionären Kontext wurde eine Datenbanksuche zur Detektion von Proteinen mit derselben Anordnung der Sekundärstruktur-elemente<sup>59</sup>, also faltungshomologen Proteinen, durchgeführt. Den besten Treffer stellte eine Häm-abhängige Chloritdismutase von *Dechloromonas aromatica* (PDB: 3q08)<sup>60</sup> dar, deren Aufgabe die Detoxifizierung von  $\text{ClO}_2^-$  ist. Die Struktur eines Monomers, dieses offensichtlich nicht funktionell verwandten Enzyms, weist eine sehr ähnliche Faltung mit zwei Ferredoxin-ähnlichen Domänen auf. Dementsprechend lassen sich die katalytischen Domänen dieser beiden Enzyme sehr gut überlagern. Unterschiedlich ist die bereits oben erwähnte Einlagerung des C-Terminus bei der bakteriellen CHI in das andere Monomer zur Dimerisierung. Aufgrund dessen werden bei der bakteriellen CHI die  $\beta$ -Faltblätter V-förmig auseinander gedrückt. Da dieses Strukturphänomen bei der Chloritdismutase nicht auftritt, liegen die  $\beta$ -Faltblätter hier annähernd parallel aufeinander. Dies führt dazu, dass bei einer Überlagerung der gesamten Tertiärstruktur die anderen Domänen jeweils gegeneinander verdreht sind. Aufgrund dieser Ergebnisse ist die Hypothese naheliegend, dass beide Proteine durch unterschiedliche Fusionsarten von zwei *ferredoxin-like* Domänen entstanden sind. Durch weiterführende Datenbankanalyse konnte ein mögliches Vorläuferprotein tatsächlich gefunden werden. Das mit Stress verbundene Protein SP1 von *Populus tremula* (PDB: 1tr0)<sup>61</sup> besteht aus 108 Aminosäureresten und besitzt eine Ferredoxin-ähnliche Faltung. Zwei SP1 Moleküle generieren über ihre  $\beta$ -Faltblätter als Kontaktfläche starke Wechselwirkungen miteinander und bilden ein stabiles Dimer mit einer Schmelztemperatur  $T_m$  von 107 °C. Eine Überlagerung der Sekundärstrukturelemente des Dimers von SP1 mit der bakteriellen CHI ergab einen r.m.s.d. für Ca's von 2,8 Å. Es ist naheliegend zu postulieren, dass sich die Dimerisierung der *ferredoxin-like* Domänen in der Evolution als eine besonders stabile Proteininfaltung herausstellte und konvergent zueinander unterschiedliche Fusionsarten entstanden. Dies entspricht auch der Tatsache, dass in der Superfamilie der dimeren  $\alpha+\beta$ -barrel Proteine, eine Unterkategorie der Ferredoxin-ähnlichen Faltungsfamilie, sowohl fusionierte als auch nicht fusionierte Vertreter vorzufinden sind<sup>62</sup>.

#### 4. Einleitung zu (*R*)-selektiven Amin-Transaminasen

Die große Bedeutung der Herstellung enantiomerenreiner Wirkstoffe wurde Anfang der 1960er Jahre durch den Arzneimittelskandal um das Medikament Contergan auf tragische Weise demonstriert. Die Verwendung des racemischen Wirkstoffes Thalidomid führte zu schweren Missbildungen Neugeborener aufgrund der teratogenen Wirkung des (*S*)-Enantiomers<sup>63</sup>. Um enantiomeren abhängige Nebenwirkungen zu vermeiden, sind selektive Methoden zur Herstellung nur eines Enantiomers favorisiert. Eine Möglichkeit hierzu ist die asymmetrische Synthese über Metall-basierte Katalyse<sup>64, 65</sup>. Diese Methoden haben jedoch meist den Nachteil, dass bei hohen Drücken, hohen Temperaturen und in organischen Lösungsmitteln gearbeitet werden muss. Zusätzlich sind die verwendeten Metallkomplexe meist sehr teuer<sup>66</sup>. Enzyme hingegen arbeiten im wässrigen Milieu, unter moderaten Temperaturbedingungen und stellen aufgrund ihrer exzellenten Enantioselektivität eine alternative Möglichkeit dar<sup>67, 68</sup>. Insbesondere das Interesse an enantioselektiven Amin-Synthesen ist bei der pharmazeutischen Industrie innerhalb des letzten Jahrzehnts stark gestiegen<sup>69-71</sup>.

Die am weitesten verbreiteten Enzyme, die Aminogruppen überführen können, sind Aminotransferasen, die vom Cofaktor Pyridoxal-5'-phosphat (PLP) abhängig sind. In den sieben Faltungsklassen PLP-abhängiger Enzyme kommen Amin-Transaminasen (EC 2.6.1) in der Faltungsklasse I und IV vor. Der Mechanismus der Transaminierung kann in zwei Reaktionshälften unterteilt werden: Zu Beginn erfolgt die Überführung der Aminogruppe von einem Aminodonator auf das PLP unter Bildung von Pyridoxamin-5'-phosphat (PMP) und anschließend die Überführung der Aminogruppe in der zweiten Hälfte des Reaktionsmechanismusses vom PMP auf den Aminoakzeptor<sup>72, 73</sup>.

Der Mechanismus (Abbildung 4.1) beginnt ausgehend von einem internen Aldimin, welches vom PLP und dem katalytischen Lysin gebildet wird. Durch die Bindung des Aminodonors kommt es zu einer Transaldimierungs-Reaktion und ein externes Aldimin, bestehend aus PLP und Aminodonator, entsteht. Das katalytische Lysin abstrahiert folgend ein Proton am C $\alpha$ , die daraufhin entstehende negative Ladung am C $\alpha$  wird über das delokalisierte  $\pi$ -Elektronensystems des PLPs stabilisiert<sup>74</sup> und ein chinoides Intermediat entsteht. Diese als Elektronensenke bezeichnete Eigenschaft des PLPs<sup>75</sup> wird in Transaminasen noch durch die Koordinierung des Pyridin-Stickstoffes mittels eines konservierten, negativ geladenen Aminosäurerestes (Aspartat oder Glutamat) begünstigt. Anschließend kommt es zur Protonierung am C4' durch das katalytische Lysin und zur Bildung des Ketimins. Die Hydrolyse des Ketimins schließt die erste Halbreaktion der Transaminierung ab und der Aminodonator verlässt das

Enzym als entsprechendes Keton bzw.  $\alpha$ -Ketosäure unter Bildung von PMP. Die Überführung der Aminofunktion vom PMP auf einen Aminoakzeptor verläuft entgegengesetzt und schließt den katalytischen Zyklus der Transaminierungsreaktion ab<sup>72, 73</sup>. Dieser Mehrsubstrat-Reaktionstyp, bei dem erst ein Substrat umgesetzt wird, daraufhin der Cofaktor verändert vorliegt und anschließend das zweite Substrat umgesetzt wird, nennt man Ping-Pong-Bi-Bi-Mechanismus<sup>76-78</sup>.

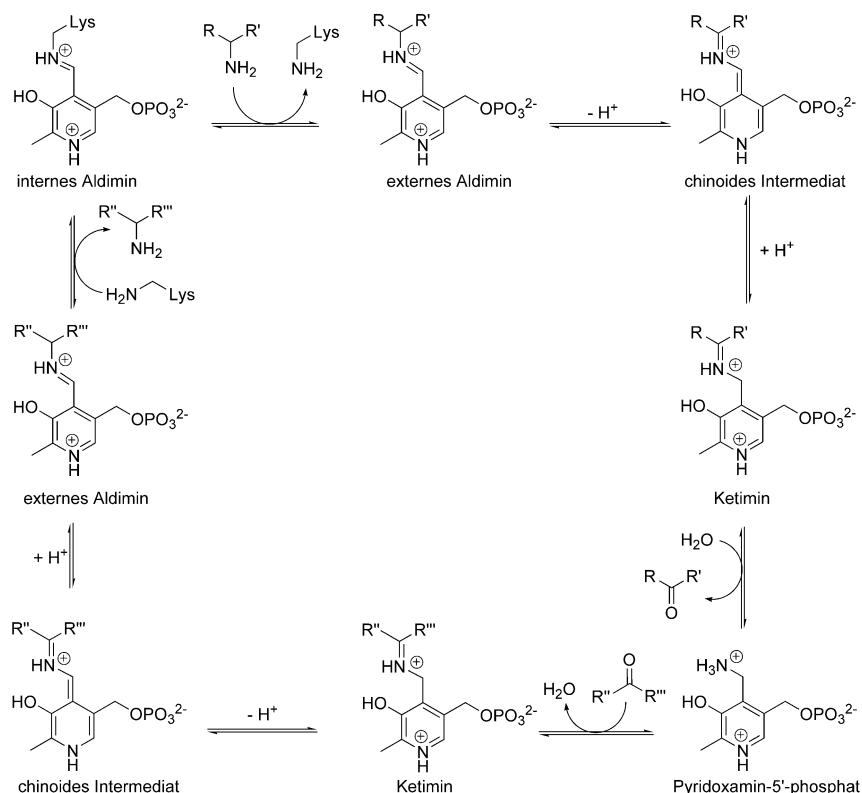
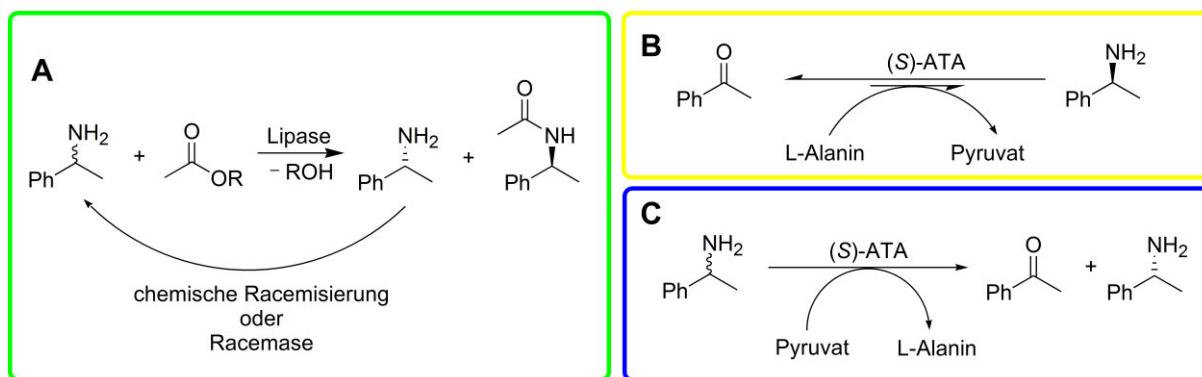


Abbildung 4.1: Mechanismus der PLP-abhängigen Transaminierung<sup>72</sup>.

Aminotransferasen können basierend auf ihren Substratspektrum in drei Klassen unterteilt werden:  $\alpha$ -Aminosäure-Transaminasen,  $\omega$ -Transaminasen und Amin-Transaminasen (ATA). So sind  $\alpha$ -Aminosäure-Transaminasen auf eine Carboxylgruppe in  $\alpha$ -Position zur Carbonylgruppe angewiesen, während  $\omega$ -Transaminasen Substrate mit mindestens einem C-Atom zwischen der Carbonylfunktion und Carboxylgruppe umsetzen können<sup>69, 79-81</sup>. Amin-Transaminasen hingegen sind nicht auf die Anwesenheit einer Carboxylgruppe angewiesen und erweitern somit das Spektrum zugänglicher Amine<sup>82, 83</sup>. Während die Strukturen der  $\alpha$ -Aminosäure-Transaminasen und  $\omega$ -Transaminasen bereits seit 1980<sup>84</sup> bzw. 1997<sup>85</sup> bekannt sind, sind die der Amin-Transaminasen erst seit kurzem bekannt. Die Identifizierung von zwei Bakterienstämmen mit (*S*)-selektiven Amin-Transaminasen gelang Shin & Kim im Jahr 1996 über Anreicherungskulturen in Minimal-Medien<sup>86, 87</sup>. Jedoch konnte die Gensequenz der ersten (*S*)-ATA von *Vibrio fluvialis* erst im Jahr 2003 durch Shin & Kim identifiziert werden<sup>88</sup>. Die Aufklärung der Struktur von *Vibrio fluvialis* gelang Park & Jang im Jahre 2011<sup>89</sup>.

Bis zur Identifizierung der (*S*)-ATAs war die enantiomerenreine Herstellung von Aminen nur über Protease- oder Lipase-vermittelte kinetische Racematspaltungen möglich, bei denen nur ein Enantiomer selektiv acetyliert wird<sup>90, 91</sup> (Abbildung 4.2A). Allerdings sind hohe Enantiomerenreinheiten nur bei einer vollständiger Umsetzung möglich<sup>67</sup> und eine Ausbeute über 50 % des gewünschten Enantiomers kann nur mittels gleichzeitiger Racemisierung des unerwünschten Enantiomers erreicht werden.

Die Identifizierung der (*S*)-ATAs ermöglichte nun die Durchführung asymmetrischer Synthesen, bei denen theoretisch eine 100 %ige Ausbeute des gewünschten Enantiomers mit exzellenter Enantiomerenreinheit ( $\%ee_p > 99\%$ ) möglich ist<sup>67, 92, 93</sup> (Abbildung 4.2B). Zwar liegt das Gleichgewicht der Reaktion stark auf der Seite der Edukte doch durch intelligente Eliminierungsverfahren des Coproduktes Pyruvat kann das Gleichgewicht zur Produktbildung verschoben werden<sup>94, 95</sup>. Aufgrund dieser Vorteile werden (*S*)-ATAs schon seit einiger Zeit erfolgreich bei der Herstellung von (*S*)-Aminen in der pharmazeutischen Industrie angewendet<sup>69, 96</sup>. Ein weiterer Vorteil der asymmetrischen Synthese gegenüber der kinetischen Racematspaltung ist, dass schon bei geringen Umsatzmengen eine hohe Enantiomerenreinheit des gewünschten Enantiomers möglich ist. Bis vor kurzem war die Herstellung des korrespondierenden (*R*)-Enantiomers jedoch nur über kinetische Racematspaltungen mithilfe von (*S*)-ATAs möglich (Abbildung 4.2C). Die maximale Ausbeute von 50 % des erwünschten Enantiomers und anschließende chemische Aminierung des Ketons mit nachfolgender Racematspaltung senkte die Gesamteffektivität eines derartigen Prozesses jedoch erheblich herab.

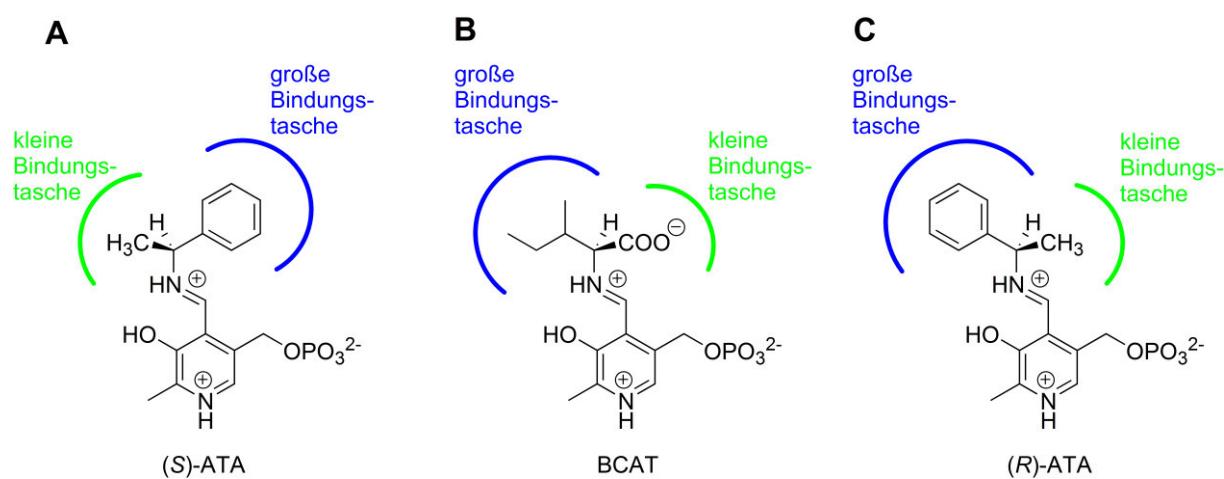


**Abbildung 4.2: Unterschiedliche Möglichkeiten zur enantiomerenreinen Herstellung von Aminen. A:** Kinetische Racematspaltung mit einer Lipase<sup>91</sup>. **B:** Prinzip einer asymmetrischen Synthese<sup>67</sup>. **C:** Prinzip einer kinetischen Racematspaltung mithilfe einer (*S*)-ATA zur Herstellung eines (*R*)-Amins<sup>67</sup>.

Eine bessere Effektivität verspricht eine asymmetrische Synthese (Abbildung 4.2B). Die Voraussetzung für eine asymmetrische Synthese ist jedoch die Verfügbarkeit eines Enzyms mit der entsprechend geforderten Enantiopräferenz. Dies war allerdings für die Herstellung von (*R*)-Aminen lange nicht gegeben. Zu diesem Zeitpunkt gab es lediglich die Kenntnis von zwei isolierten

Bakterienstämmen zur Herstellung von (*R*)-Aminen, jedoch ohne Identifizierung und Kenntnis der Sequenz der verantwortlichen Enzyme<sup>97,98</sup>.

Für die Identifizierung weiterer Sequenzen von (*R*)-ATAs wurde der klassische Weg der Anreicherungskulturen von Mikroorganismen in Minimalmedien als zu langwierig erachtet. Höhne et al. entwickelten deshalb im Jahr 2010 einen rationalen Ansatz über eine *in silico* Datenbankanalyse<sup>82</sup>. Ausgehend hierfür war eine vorhergehende Analyse der Strukturen bereits bekannter Aminotransferasen. Die Amin-Transaminasen der Faltungsklasse I haben ohne Ausnahme eine Präferenz für das (*S*)-Enantiomer. Die Faltungsklasse IV weist hingegen die Besonderheit auf, dass während der Evolution innerhalb dieser Proteinfaltung bereits eine Umkehr der Enantiopräferenz stattgefunden hat, denn sowohl D-Aminosäuren-Aminotransferasen (*D-amino acid aminotransferase*; D-ATA) als auch die (*S*)-selektiven Verzweigtkettigen-Aminosäure-Aminotransferasen (*branched-chain amino acid aminotransferase*; BCAT) sind in der Faltungsklasse IV anzufinden. Dies lässt darauf schließen, dass diese Art der Proteinfaltung eine hohe strukturelle Flexibilität des aktiven Zentrums erlaubt. Aus diesem Grund haben Höhne et al. die Faltungsklasse IV als Grundlage für ihre weitere Analyse gewählt. Bei genauerer Betrachtung der Architektur der aktiven Zentren der BCATs und der D-ATAs, stellt man fest, dass, wie auch schon bei den (*S*)-ATA beobachtet<sup>99,100</sup>, diese in eine große und eine kleine Bindungstasche unterteilt werden können (Abbildung 4.3A und B).



**Abbildung 4.3: Architektur der aktiven Zentren unterschiedlicher Vertreter der Transaminasen im Vergleich.** A: Faltungsklasse I mit der (*S*)-ATA B: Faltungsklasse IV mit BCAT. Beide Enzymklassen weisen eine Unterteilung des aktiven Zentrums in eine kleine und große Bindungstasche auf. Man beachte die vertauschte Anordnung der Bindungstaschen. C: Postulierte aktives Zentrum einer (*R*)-ATA nach Höhne et al.<sup>82</sup>.

Ein bemerkenswerter Unterschied der Faltungsklasse I und IV ist die gegensätzliche Anordnung der Bindungstaschen. Höhne et al. postulierten, dass durch die Inversion des Stereozentrums, aufgrund der Änderung der Prioritäten nach der Cahn-Ingold-Prelog-Regel<sup>101</sup>, die Architektur einer potentiellen (*R*)-ATA eine ähnliche Anordnung der kleinen und großen Bindungstasche wie eine BCAT haben

müsste (Abbildung 4.3C). Im Gegensatz zu jener Enzymklasse müssten allerdings beide Bindungstaschen einer potentiellen (*R*)-ATA einen eher hydrophoben Charakter aufweisen.

Ausgehend von diesen Überlegungen entwickelten Höhne et al. einen Algorithmus für die Proteinsequenz-basierte Datenbankanalyse, der Sequenzen mit den spezifischen Sequenzmotiven, die für die Substraterkennung bei BCATs und D-ATAs verantwortlich sind, ausschloss. Hierdurch gelang zwar die Identifizierung und biochemische Charakterisierung von 17 Proteinsequenzen (*R*)-selektiver Amin-Transaminasen<sup>82</sup>, doch fehlten immer noch strukturelle Informationen zum Aufbau des aktiven Zentrums.

Ausgehend von der bereits oben erwähnten hydrophob geprägten Umgebung im aktiven Zentrum einer (*R*)-ATA stellt sich ferner die Frage, wie diese die Koordination des polaren Aminodonors D-Alanin vermitteln. Diese Fähigkeit zur Erkennung und Bindung zweier unterschiedlicher Substrate nennt man *dual substrate recognition* und konnte für Transaminasen der Faltungsklasse I bereits aufgeklärt werden. Unabhängig voneinander beobachteten Cronin & Kirsch<sup>102</sup> und Hayashi<sup>103</sup>, dass die Mutation eines am Eingangstunnel befindlichen Arginins, die Transaminaseaktivität der Aspartat-Aminotransferase (L-AspAT) von *E. coli* drastisch herabsetzt. Onuffer und Kirsch<sup>104</sup> postulierten daraufhin, dass dieses Arginin die Koordinierung einer Carboxylgruppe vermitteln kann, indem es eine Konformation einnimmt, in der es direkt ins aktive Zentrum zeigt<sup>104</sup>. Bindet das Enzym nun ein hydrophobes Substrat wie Tyrosin, nimmt das Arginin eine alternative Konformation ein, in der es sich aus dem aktiven Zentrum herausdreht. Diese Hypothese konnte wenige Zeit später von Almo et al. über Kristallstrukturen der L-AspAT\_Arg292Ala<sup>105</sup> und von Malashkevich et al. mittels Strukturen der L-AspAT mit kovalent gebundenen Inhibitorkomplexen<sup>106</sup> verifiziert werden. Ob die duale Substraterkennung in (*R*)-ATAs, die in der Faltungsklasse IV eingeordnet sind, ähnlich verläuft, ist noch nicht hinreichend untersucht. Vor kurzem publizierten zwei Gruppen unabhängig voneinander Untersuchungen zur dualen Substraterkennung bei (*R*)-ATAs, deren Ergebnissen sich allerdings widersprechen. Während Sayer et al. die Kristallstruktur eines Inhibitorkomplex der (*R*)-ATA aus *Nectria haematococca*<sup>107</sup> untersuchten und keinen Hinweis auf ein *switching* Arginin fanden, zeigen die Ergebnisse von Łyskowski et al., dass die Amin-Transaminaseaktivität des Enzyms aus *Aspergillus terreus*<sup>108</sup> signifikant durch die Mutation Arg128Ala beeinträchtigt wird.

## 5. Ergebnisse zur (*R*)-selektiven Amin-Transaminase aus *Aspergillus fumigatus*

### 5.1 Kristallisation und Röntgendiffraktion der (*R*)-selektiven Amin-Transaminase

Die Expression und Reinigung des Enzyms wurde von L. Skalden aus dem Arbeitskreis Biotechnologie und Enzymkatalyse der Universität Greifswald durchgeführt. Hochauflösende Kristalle der (*R*)-ATA aus *Aspergillus fumigatus*, die durch batch-Kristallisation in einen 1,5-mL-Reaktionsgefäß entstanden, wurden unter Kryobedingungen am Synchrotron BESSY II<sup>54</sup> (Berlin, Deutschland) vermessen. Bei einer Wellenlänge von 0,9184 Å beugten die Kristalle Röntgenstrahlen bis zu einer Auflösung von 1,27 Å. Da zu dem Zeitpunkt der Datenaufnahme nur Strukturen von homologen Proteinen mit einer Sequenzidentität von <27 % bekannt waren, wurde ein weiterer Datensatz bei einer Wellenlänge von 1,7712 Å aufgenommen, um die Strukturlösung über die anomale Dispersion der im Protein enthaltenden Schwefelatome (*Sulphur single-wavelength anomalous diffraction; SSAD*) zu ermöglichen. Aufgrund der Detektorgeometrie war die Auflösung dieses Datensatzes auf 1,84 Å begrenzt.

Die Datenprozessierung ergab, dass das Protein in der orthorhombisch zentrierten Raumgruppe  $C222_1$  kristallisiert (siehe Publikation III, Abbildung 1 und Tabelle 1). Die Selbstrotationsfunktion impliziert, dass ein Dimer die asymmetrische Einheit bildet.

Das Strukturmodell besteht aus zwei Polypeptidketten mit insgesamt 639 Aminosäuren, zwei PLP-Molekülen, vier Kaliumionen, vier Chlorid-Ionen und zwei Glycerolmolekülen. Bei einer Auflösung von 1,27 Å wurde das Strukturmodell bis zu einem  $R/R_{\text{free}}$ -Faktor von 10,3 %/12,7 % verfeinert.

### 5.2 Strukturanalyse der (*R*)-selektiven Amin-Transaminase

Das Strukturmodell des Monomers der (*R*)-ATA aus *Aspergillus fumigatus* kann in eine kleine Domäne (N-terminus – Pro144) mit  $\alpha/\beta$ -Struktur, einen Interdomänen-Loop (Tyr145 – Met149) und eine große Domäne (Ala150 – C-Terminus) mit einer *pseudo-barrel* Faltung unterteilt werden. Das Sekundärstrukturalignment mit einem Vertreter der Verzweigtkettigen-Aminosäure-Aminotransferasen (BCAT von *Thermus thermophilus*, PDB: 1wrv, r.m.s.d. für  $\text{C}\alpha$ 's: 1.8 Å) und der D-Aminosäuren-Amin-Transaminasen (D-ATA von *Bacillus sp.* YM-1, PDB: 3lqs, r.m.s.d. für  $\text{C}\alpha$ 's: 2.0 Å) bestätigte die Hypothese von Höhne et al. und ordnet die (*R*)-ATA in die Faltungsklasse IV der PLP-abhängigen Enzyme ein (siehe Publikation IV, Abbildung 3).

Ein wesentlicher Unterschied in Hinblick auf den strukturellen Vergleich mit den BCATs und D-ATAs stellt eine nur in den (*R*)-ATAs vorhandene 20 Aminosäuren lange  $\alpha$ -Helix am N-Terminus dar. Diese Helix scheint für die (*R*)-ATAs eine essentielle Bedeutung der Löslichkeitsvermittlung und somit für die Proteinstabilität zu haben. Eine von L. Skalden durchgeführte Deletion dieser Helix von der (*R*)-ATA aus *Neosartorya fischeri* (96 % Sequenzidentität zur (*R*)-ATA aus *Aspergillus fumigatus*) führte zur Expression unlöslichen Proteins.

Durch die Beteiligung beider Monomere an der Bildung des aktiven Zentrums ist die kleinste katalytische Einheit der (*R*)-ATAs ein Homodimer. Der sogenannte *active site loop* (Gly121\* - Asn135\*; Aminosäurereste, die mit einem Stern markiert werden, gehören zur anderen Untereinheit) wird jeweils vom anderen Monomer zur Verfügung gestellt und ist an der Bildung der großen Bindungstasche sowie der Beschränkung des Eingangstunnels zum aktiven Zentrum beteiligt. Die Koordinierung des Cofaktors PLP im aktiven Zentrum wird über Aminosäurereste, die hochkonserviert für die gesamte Faltungsklasse IV der PLP-abhängigen Enzyme sind, vermittelt (siehe Publikation IV, Abbildung 7). Das PLP konnte in der Struktur des Holoenzyms in zwei Konformationen gefunden werden. Zum einen kovalent verknüpft als internes Aldimin mit dem katalytisch verantwortlichen Lys179 und als freies PLP bzw. PMP. Diese beiden Konformationen unterscheiden sich um 19° im Diederwinkel der C5-C5A-Bindung am Pyridoxalring. Die Bindung des PLPs erfolgt somit hauptsächlich über die Koordinierung des N1s und der Phosphatgruppe.

Die Erklärung der (*R*)-Enantioselektivität und Unterteilung des aktiven Zentrums in die große und kleine Bindungstasche erfolgten über *in silico* Docking-Experimente (durchgeführt von L. Skalden). Gedockt wurden die Substrat-Addukte von (*R*)- $\alpha$ -Phenylethylamin und PLP als auch mit (*S*)- $\alpha$ -Phenylethylamin und PLP (siehe Publikation IV, Abbildung 5). Trotz unterschiedlicher Chiralität wurden beide Substrat-Addukte in eine ähnliche Orientierung gedockt. Die Methyl-Gruppen definieren die kleine Bindungstasche, die von den Aminosäuren Val60, Phe113 und Ile146 gebildet wird, während der Benzylring von His53\*, Tyr58, Arg126\*, Val148 und Trp183 koordiniert wird und dies die große Bindungstasche definiert. Die (*R*)-Enantioselektivität lässt sich anhand der Tatsache zeigen, dass nur bei dem (*R*)- $\alpha$ -MBA-PLP Addukt eine Abstraktion des Wasserstoffatoms am C $\alpha$ -Atom von der *si*-Seite<sup>109</sup> zur Initialisierung der Deaminierung möglich ist. Nur in dieser Konformation befindet sich das Wasserstoffatom in einem sinnvollen Abstand zum katalytischen Lysin (2,8 Å).

Ein Vergleich der (*R*)-ATA mit BCATs und D-ATAs (siehe Publikation IV, Abbildung 7) zeigt eine hohe Konservierung des PLP-Bindemotivs. Außer dem katalytischen Lysin und den für die PLP-Bindung verantwortlichen Aminosäurereste gibt es jedoch keine weiteren Sequenzübereinstimmungen. Wie in der Einleitung bereits dargelegt worden ist, wurde postuliert, dass das aktive Zentrum der (*R*)-ATA aufgrund der Inversion des Stereozentrums ähnlicher zu den BCATs als zu den D-ATAs sein müsste.

Dies kann mit Hilfe der Struktur bestätigt werden (siehe Publikation IV, Abbildung 8). Während bei BCATs und (*R*)-ATAs das Volumen der kleinen Bindungstasche durch größere aromatische Aminosäuren reduziert wird, sind bei den D-ATAs kleine hydrophobe Reste vorherrschend. Zur Bindung der meist hydrophoben Substrate ist die große Bindungstasche der BCATs und (*R*)-ATA mit aromatischen Aminosäuren aufgebaut, in Gegensatz zu den bei D-ATAs vorherrschend polaren Aminosäuren zur Koordinierung der Carboxylgruppe.

Zeitgleich zu der hier diskutierten Struktur wurde die Struktur der (*R*)-ATA von *Aspergillus terreus* veröffentlicht<sup>108</sup> (PDB: 4ce5; 71,5 % Sequenzidentität zu der (*R*)-ATA von *Aspergillus fumigatus*). Ein Vergleich der beiden Strukturen zeigt wenig Unterschiede in ihrer 3D-Struktur (r.m.s.d für Cα's von 0,57 Å) und auch die beiden PLP-Konformationen sind in 4ce5 erkennbar. Die einzige Abweichung ist ein anderer Hauptkettenverlauf des dem Solvens zugewandten Bereiches des *active site loops*. Auf der Grundlage dieser Strukturen konnten erste Informationen mit annähernd atomarer Auflösung zur Funktionsweise des Substraterkennungsmechanismus von (*R*)-selektiven Amin-Transaminasen erhalten werden.

### 5.3 Untersuchungen zur dualen Substraterkennung der (*R*)-selektiven Amin-Transaminase

Um einen tieferen Einblick und Verständnis für die Mechanismen der Substraterkennung und insbesondere des dualen Substraterkennung-Mechanismus zu bekommen, wurde ein Inhibitorkomplex der (*R*)-ATA mit Gabaculin<sup>110</sup>, einem häufig verwendeten Suizid-Inhibitor für PLP-abhängige Enzyme<sup>111</sup>, hergestellt. Die Kristalle der (*R*)-ATA wurden zuerst mit PLP, zur vollständigen Besetzung des Cofaktors im Protein und anschließend mit Gabaculin getränkt und vermessen. Es konnte ein Datensatz mit einer Auflösung von 1,6 Å am Synchrotron BESSY II<sup>54</sup> (Berlin, Deutschland) aufgenommen werden. Dieser wurde mittels *Molecular Replacement* und dem Apoenzym der oben gelösten Struktur als Template gelöst (siehe Publikation V, Tabellen 1 und 2).

Die im Abschnitt 5.2 beschriebene Segmentierung des aktiven Zentrums der (*R*)-ATA konnte durch den kovalenten Inhibitorkomplex *meta*-Carboxyphenylpyridoxaminphosphat (*mCPP*), das Addukt aus PLP und Gabaculin, verifiziert werden. Die Koordinierung der polaren Carboxylgruppe des *mCPPs* wird über ein Wassermolekül von His53\* und Tyr58 aus der großen Bindungstasche und mit Hilfe eines weiteren Wassermoleküls über Arg126\* aus dem *active site loop* vermittelt (siehe Publikation V, Abbildung 4). Eine Besonderheit dieses Strukturmodells der (*R*)-ATA ist, dass bedingt durch andere Kristallisationsbedingungen und einer anderen Raumgruppe (*P3<sub>1</sub>21*) unterschiedliche Konformationen des *active site loops* vorliegen. Während Monomer B in einer geschlossenen Konformation vorliegt, dominiert in Monomer A mit einem Besetzungsgrad von 0,7 eine durch

Kristallkontakte stabilisierte offene Form des *active site loops*. Nur wenige Reste dieses *loops* (Gly121 – Arg126; Leu133 – Asn135) liegen mit einer geringen Besetzung (0,3) in der geschlossenen Form vor.

Der gebundene Inhibitor konnte in beiden Monomeren identifiziert werden. In der offenen Konformation ist der Benzoesäure-Rest jedoch nur mit einer Besetzung von 0,8 vorzufinden. Zusätzlich sind alternative Konformationen durch Rotationsbewegungen der Carboxylgruppe sichtbar. In der geschlossenen Konformation liegt nur eine *mCPP*-Konformation mit voller Besetzung vor. Dies ist darauf zurückzuführen, dass durch die Schließung des *active site loops* die Koordinierung der Carboxylgruppe des *mCPPs* durch Arg126 über ein Wassermolekül vermittelt wird, sodass nur noch eine distinkte Konformation erlaubt ist, während bei der offenen Konformation des *loops* weniger Interaktionen vorherrschen und dadurch größere Rotationsfreiheiten erlaubt sind.

Der Vergleich mit der kürzlich von Sayer et al. publizierten Inhibitorstruktur von *Nectria haematococca*<sup>107</sup> (PDB: 4cmf) zeigt bemerkenswerte Unterschiede. So zeigt der Inhibitorkomplex in der Struktur 4cmf im Gegensatz zu der deutlich abgeknickten *mCPP*-Konformation im Enzym von *Aspergillus fumigatus* eine annähernd planare Anordnung des *m*-Carboxylphenylamins zum Pyridinring. Die Differenz der Diederwinkel um C4-C4A-N9-C10 beträgt ca. 17° (siehe Publikation V, Abbildung 6). Die gebogene *mCPP*-Konformation führt zu einer deutlich verkürzten Distanz der Carboxylatgruppe zum Arg126 von nur 3,8 Å im Vergleich zu 5,5 Å bei 4cmf.

Zeitgleich zu dieser Strukturanalyse wurde von L. Skalden mittels Mutagenese die Bedeutung des Arg126 zur Akzeptanz unterschiedlicher Aminoakzeptoren untersucht. Die Fähigkeit der Arg126Ala Variante Pyruvat als Aminoakzeptor umzusetzen sank auf 3 % der Aktivität im Vergleich zum Wildtyp und beweist somit die signifikante Bedeutung des Arg126 für die duale Substraterkennung. Diese Annahme wird auch von den Ergebnissen von Łyskowski et al.<sup>108</sup> gestützt. Des Weiteren konnte eine Inversion der Präferenz zu den einzelnen Aminoakzeptoren beobachtet werden. Während der Wildtyp Pyruvat Pentanal vorzieht, ist die Aktivität der Arg126Ala Variante mit Pentanal als Aminoakzeptor zwar auch herabgesetzt, jedoch noch signifikant höher als mit Pyruvat und zeigt somit deutlich die höhere Präferenz der Arg126Ala Variante für unpolare Substrate (siehe Publikation V, Abbildung 5).

Die Verknüpfung struktureller mit Enzym-kinetischen Untersuchungen zum dualen Substraterkennungsmechanismus erbrachte wichtige Erkenntnisse zum Verständnis der Umsetzung carboxylierter Substrate. Diese Informationen sind essentielle Voraussetzungen für weitere Experimente zur Erweiterung des Substratspektrums der (*R*)-selektiven Amin-Transaminasen für deren Anwendung in der Synthese pharmazeutisch wichtiger Amine.

## 6. Literatur

1. Gupta, S. K.; Prakash, J.; Srivastava, S., Validation of traditional claim of tulsi, *Ocimum sanctum* Linn. as a medical plant. *Indian J. Exp. Biol.* **2002**, *40*, 765-773.
2. Kelm, M. A.; Nair, M. G.; Strasburg, G. M.; DeWitt, D. L., Antioxidant and cyclooxygenase inhibitory phenolic compounds from *Ocimum sanctum* Linn. *Phytomedicine* **2000**, *7* (1), 7-13.
3. Szent-Gyorgyi Nagyrapolt, A., *On oxidation, fermentation, vitamins, health and disease*. **1939**.
4. Cook, N. C. & Samman, S., Flavonoids—Chemistry, metabolism, cardioprotective effects, and dietary sources. *J. Nutr. Biochem.* **1996**, *7* (2), 66-76.
5. Tahara, S., A journey of twenty-five years through the ecological biochemistry of flavonoids. *Biosci. Biotechnol. Biochem.* **2007**, *71* (6), 1387-1404.
6. Das, S. & Rosazza, J. P. N., Microbial and enzymatic transformations of flavonoids. *J. Nat. Prod.* **2006**, *69* (3), 499-508.
7. Cabrita, L.; Fossen, T.; Andersen, Ø. M., Colour and stability of the six common anthocyanidin 3-glucosides in aqueous solutions. *Food Chem.* **2000**, *68* (1), 101-107.
8. Mateus, N. & Freitas, V., Anthocyanins as food colorants. Winefield, C.; Davies, K.; Gould, K., Eds. Springer New York: **2009**; pp 284-304.
9. Marín, F. R.; Frutos, M. J.; Pérez-Alvarez, J. A.; Martínez-Sánchez, F.; Del Río, J. A., Flavonoids as nutraceuticals: structural related antioxidant properties and their role on ascorbic acid preservation. *Stud. Nat. Prod. Chem.* **2002**, *26*, 741-778.
10. Formica, J. V. & Regelson, W., Review of the biology of quercetin and related bioflavonoids. *Food Chem. Toxicol.* **1995**, *33* (12), 1061-1080.
11. da Silva, E. L.; Piskula, M. K.; Yamamoto, N.; Moon, J.-H.; Terao, J., Quercetin metabolites inhibit copper ion-induced lipid peroxidation in rat plasma. *FEBS Lett.* **1998**, *430* (3), 405-408.
12. Di Carlo, G.; Mascolo, N.; Izzo, A. A.; Capasso, F., Flavonoids: Old and new aspects of a class of natural therapeutic drugs. *Life Sci.* **1999**, *65* (4), 337-353.
13. Kahraman, A.; Erkasap, N.; Köken, T.; Serteser, M.; Aktepe, F.; Erkasap, S., The antioxidative and antihistaminic properties of quercetin in ethanol-induced gastric lesions. *Toxicology* **2003**, *183* (1–3), 133-142.
14. Caltagirone, S.; Rossi, C.; Poggi, A.; Ranelli, F. O.; Natali, P. G.; Brunetti, M.; Aiello, F. B.; Piantelli, M., Flavonoids apigenin and quercetin inhibit melanoma growth and metastatic potential. *Int. J. Cancer* **2000**, *87* (4), 595-600.
15. Notoya, M.; Tsukamoto, Y.; Nishimura, H.; Woo, J.-T.; Nagai, K.; Lee, I.-S.; Hagiwara, H., Quercetin, a flavonoid, inhibits the proliferation, differentiation, and mineralization of osteoblasts in vitro. *Eur. J. Pharmacol.* **2004**, *485* (1–3), 89-96.
16. De Conti Lourenço, R.; da Silva Melo, P.; de Almeida, A., Flavonoids as Antifungal Agents. In *Antifungal Metabolites from Plants*, Razzaghi-Abyaneh, M. & Rai, M., Eds. Springer Berlin Heidelberg: **2013**; pp 283-300.
17. Vessal, M.; Hemmati, M.; Vasei, M., Antidiabetic effects of quercetin in streptozocin-induced diabetic rats. *Comp. Biochem. Physiol. C: Toxicol. Pharmacol.* **2003**, *135* (3), 357-364.
18. Alguer, Y.; Meng, C.; Terán, W.; Krell, T.; Ramos, J. L.; Gallegos, M.-T.; Zhang, X., Crystal structures of multidrug binding protein TtgR in complex with antibiotics and plant antimicrobials. *J. Mol. Biol.* **2007**, *369* (3), 829-840.
19. Kumar, S. & Pandey, A. K., Chemistry and biological activities of flavonoids: an overview. *The Scientific World Journal* **2013**, *2013*, 16.
20. Santell, R. C.; Chang, Y. C.; Nair, M. G.; Helferich, W. G., Dietary genistein exerts estrogenic effects upon the uterus, mammary gland and the hypothalamic/pituitary axis in rats. *J. Nutr.* **1997**, *127* (2), 263-269.

21. Cos, P.; Calomme, M.; Pieters, L.; Vlietinck, A. J.; Vanden Berghe, D., Structure-activity relationship of flavonoids as antioxidant and pro-oxidant compounds. In *Studies in Natural Products Chemistry*, Atta-ur, R., Ed. Elsevier: **2000**; Vol. Volume 22, Part C, pp 307-341.
22. Hodek, P.; Trefil, P.; Stiborová, M., Flavonoids-potent and versatile biologically active compounds interacting with cytochromes P450. *Chem. Biol. Interact.* **2002**, *139* (1), 1-21.
23. Zhai, S.; Dai, R.; Friedman, F. K.; Vestal, R. E., Comparative inhibition of human cytochromes P450 1A1 and 1A2 by flavonoids. *Drug Metab. Dispos.* **1998**, *26* (10), 989-992.
24. Doostdar, H.; Burke, M. D.; Mayer, R. T., Bioflavonoids: selective substrates and inhibitors for cytochrome P450 CYP1A and CYP1B1. *Toxicology* **2000**, *144* (1-3), 31-38.
25. Heller, W. & Hahlbrock, K., Highly purified "flavanone synthase" from parsley catalyzes the formation of naringenin chalcone. *Arch. Biochem. Biophys.* **1980**, *200* (2), 617-619.
26. Ferrer, J.-L.; Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P., Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nat. Struct. Mol. Biol.* **1999**, *6* (8), 775-784.
27. Austin, M. B. & Noel, J. P., The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* **2003**, *20* (1), 79-110.
28. Jung, W.; Yu, O.; Lau, S.-M. C.; O'Keefe, D. P.; Odell, J.; Fader, G.; McGonigle, B., Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nat. Biotechnol.* **2000**, *18* (2), 208-212.
29. Moustafa, E. & Wong, E., Purification and properties of chalcone-flavanone isomerase from soya bean seed. *Phytochem.* **1967**, *6* (5), 625-632.
30. Dixon, R. A.; Richard Blyden, E.; Robbins, M. P.; Van Tunen, A. J.; Mol, J. N., Comparative biochemistry of chalcone isomerases. *Phytochem.* **1988**, *27* (9), 2801-2808.
31. Kimura, Y.; Aoki, T.; Ayabe, S.-i., Chalcone isomerase isozymes with different substrate specificities towards 6'-Hydroxy- and 6'-deoxychalcones in cultured cells of *Glycyrhiza echinata*, a leguminous plant producing 5-deoxyflavonoids. *Plant Cell Physiol.* **2001**, *42* (10), 1169-1173.
32. Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P., Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nat. Struct. Mol. Biol.* **2000**, *7* (9), 786-791.
33. Weng, J.-K.; Philippe, R. N.; Noel, J. P., The rise of chemodiversity in plants. *Science* **2012**, *336* (6089), 1667-1670.
34. Gensheimer, M. & Mushegian, A., Chalcone isomerase family and fold: No longer unique to plants. *Protein Sci.* **2004**, *13* (2), 540-544.
35. Ngaki, M. N.; Louie, G. V.; Philippe, R. N.; Manning, G.; Pojer, F.; Bowman, M. E.; Li, L.; Larsen, E.; Wurttele, E. S.; Noel, J. P., Evolution of the chalcone-isomerase fold from fatty-acid binding to stereospecific catalysis. *Nature* **2012**, *485* (7399), 530-533.
36. Jez, J. M.; Bowman, M. E.; Noel, J. P., Role of hydrogen bonds in the reaction mechanism of chalcone isomerase. *Biochemistry* **2002**, *41* (16), 5168-5176.
37. Jez, J. M. & Noel, J. P., Reaction mechanism of chalcone isomerase. *J. Biol. Chem.* **2002**, *277* (2), 1361-1369.
38. Copley, S. D. & Knowles, J. R., The uncatalyzed Claisen rearrangement of chorismate to prephenate prefers a transition state of chair-like geometry. *J. Am. Chem. Soc.* **1985**, *107* (18), 5306-5308.
39. Tramontano, A.; Janda, K.; Lerner, R., Catalytic antibodies. *Science* **1986**, *234* (4783), 1566-1570.
40. Deutsche Gesellschaft für Ernährung, *Ernährungsbericht 2008*. 335-346: Bonn, **2008**.
41. Cunningham, P.; Afzal-Ahmed, I.; Naftalin, R. J., Docking studies show that D-glucose and quercetin slide through the transporter GLUT1. *J. Biol. Chem.* **2006**, *281* (9), 5797-5803.
42. Bokkenheuser, V. D.; Shackleton, C. H.; Winter, J., Hydrolysis of dietary flavonoid glycosides by strains of intestinal bacteroides from humans. *Biochem. J.* **1987**, *248* (3), 953-956.
43. Schneider, H.; Schwierz, A.; Collins, M. D.; Blaut, M., Anaerobic transformation of quercetin-3-glucoside by bacteria from the human intestinal tract. *Arch. Microbiol.* **1999**, *171* (2), 81-91.

44. Simmering, R.; Kleessen, B.; Blaut, M., Quantification of the flavonoid-degrading bacterium *Eubacterium ramulus* in human fecal samples with a species-specific oligonucleotide hybridization probe. *Appl. Environ. Microbiol.* **1999**, *65* (8), 3705-3709.
45. Schneider, H. & Blaut, M., Anaerobic degradation of flavonoids by *Eubacterium ramulus*. *Arch. Microbiol.* **2000**, *173* (1), 71-75.
46. Braune, A.; Gütschow, M.; Engst, W.; Blaut, M., Degradation of quercetin and luteolin by *Eubacterium ramulus*. *Appl. Environ. Microbiol.* **2001**, *67* (12), 5558-5567.
47. Schoefer, L.; Mohan, R.; Braune, A.; Birringer, M.; Blaut, M., Anaerobic C-ring cleavage of genistein and daidzein by *Eubacterium ramulus*. *FEMS Microbiol. Lett.* **2002**, *208* (2), 197-202.
48. Haddock, J. D. & Ferry, J. G., Purification and properties of phloroglucinol reductase from *Eubacterium oxidoreducens* G-41. *J. Biol. Chem.* **1989**, *264* (8), 4423-4427.
49. Schoefer, L.; Braune, A.; Blaut, M., Cloning and expression of a phloretin hydrolase gene from *Eubacterium ramulus* and characterization of the recombinant enzyme. *Appl. Environ. Microbiol.* **2004**, *70* (10), 6131-6137.
50. Herles, C.; Braune, A.; Blaut, M., First bacterial chalcone isomerase isolated from *Eubacterium ramulus*. *Arch. Microbiol.* **2004**, *181* (6), 428-434.
51. Gall, M.; Thomsen, M.; Peters, C.; Pavlidis, I. V.; Jonczyk, P.; Grünert, P. P.; Beutel, S.; Schepers, T.; Gross, E.; Backes, M.; Geißler, T.; Ley, J. P.; Hilmer, J.-M.; Krammer, G.; Palm, G. J.; Hinrichs, W.; Bornscheuer, U. T., Enzymatic conversion of flavonoids using bacterial chalcone isomerase and enoate reductase. *Angew. Chem. Int. Ed.* **2014**, *53* (5), 1439-1442.
52. Doublie, S., Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **1997**, *276*, 523-530.
53. Bergfors, T., Seeds to crystals. *J. Struct. Biol.* **2003**, *142* (1), 66-76.
54. Mueller, U.; Darowski, N.; Fuchs, M. R.; Forster, R.; Hellwig, M.; Paithankar, K. S.; Puhringer, S.; Steffien, M.; Zocher, G.; Weiss, M. S., Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J. Synchrotron Radiat.* **2012**, *19* (3), 442-449.
55. Krissinel, E. & Henrick, K., Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372* (3), 774-797.
56. Abendroth, J.; Rice, A. E.; McLuskey, K.; Bagdasarian, M.; Hol, W. G. J., The crystal structure of the periplasmic domain of the type II secretion system protein EpsM from *Vibrio cholerae*: the simplest version of the ferredoxin fold. *J. Mol. Biol.* **2004**, *338* (3), 585-596.
57. Kratzat, H. Detaillierte Struktur-Untersuchungen zur Chalconisomerase aus *Eubakterium ramulus*. Bachelor-Arbeit, Universität Greifswald, Greifswald, **2014**.
58. N.M. Mol, J.; P. Robbinst, M.; A. Dixon, R.; Veltkamp, E., Spontaneous and enzymic rearrangement of naringenin chalcone to flavanone. *Phytochem.* **1985**, *24* (10), 2267-2269.
59. Dietmann, S.; Park, J.; Notredame, C.; Heger, A.; Lappe, M.; Holm, L., A fully automatic evolutionary classification of protein folds: Dali Domain Dictionary version 3. *Nucleic Acids Res.* **2001**, *29* (1), 55-57.
60. de Geus, D. C.; Thomassen, E. A. J.; Hagedoorn, P.-L.; Pannu, N. S.; van Duijn, E.; Abrahams, J. P., Crystal Structure of chlorite dismutase, a detoxifying enzyme producing molecular oxygen. *J. Mol. Biol.* **2009**, *387* (1), 192-206.
61. Dgany, O.; Gonzalez, A.; Sofer, O.; Wang, W.; Zolotnitsky, G.; Wolf, A.; Shoham, Y.; Altman, A.; Wolf, S. G.; Shoseyov, O.; Almog, O., The structural basis of the thermostability of SP1, a novel plant (*Populus tremula*) boiling stable protein. *J. Biol. Chem.* **2004**, *279* (49), 51516-51523.
62. Fox, N. K.; Brenner, S. E.; Chandonia, J.-M., SCOPe: structural classification of proteins—extended, integrating SCOP and ASTRAL data and classification of new structures. *Nucleic Acids Res.* **2014**, *42* (D1), D304-D309.
63. Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H., Identification of a primary target of thalidomide teratogenicity. *Science* **2010**, *327* (5971), 1345-1350.
64. Hansen, K. B.; Balsells, J.; Dreher, S.; Hsiao, Y.; Kubryk, M.; Palucki, M.; Rivera, N.; Steinhuebel, D.; Armstrong, J. D.; Askin, D.; Grabowski, E. J. J., First generation process for

- the preparation of the DPP-IV inhibitor sitagliptin. *Org. Process Res. Dev.* **2005**, *9* (5), 634-639.
65. Noyori, R. & Hashiguchi, S., Asymmetric transfer hydrogenation catalyzed by chiral ruthenium complexes. *Acc. Chem. Res.* **1997**, *30* (2), 97-102.
66. Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J., Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **2010**, *329* (5989), 305-309.
67. Höhne, M. & Bornscheuer, U. T., Application of transaminases in organic synthesis. In *Enzymes in Organic Synthesis*, May, O., Gröger, H., Drauz, W., Ed. Ed. Wiley-VCH: Weinheim: **2012**; pp 779-820.
68. Malik, M. S.; Park, E.-S.; Shin, J.-S., Features and technical applications of  $\omega$ -transaminases. *Appl. Microbiol. Biotechnol.* **2012**, *94* (5), 1163-1171.
69. Midelfort, K. S.; Kumar, R.; Han, S.; Karmilowicz, M. J.; McConnell, K.; Gehlhaar, D. K.; Mistry, A.; Chang, J. S.; Anderson, M.; Villalobos, A.; Minshull, J.; Govindarajan, S.; Wong, J. W., Redesigning and characterizing the substrate specificity and activity of *Vibrio fluvialis* aminotransferase for the synthesis of imagabalin. *Protein Eng. Des. Sel.* **2013**, *26* (1), 25-33.
70. Ogawa, K.; Koyama, Y.; Ohashi, I.; Sato, I.; Hirama, M., Total synthesis of a protected aglycon of the kedarcidin chromophore. *Angew. Chem. Int. Ed.* **2009**, *48* (6), 1110-1113.
71. Sutin, L.; Andersson, S.; Bergquist, L.; Castro, V. M.; Danielsson, E.; James, S.; Henriksson, M.; Johansson, L.; Kaiser, C.; Flyrén, K.; Williams, M., Oxazolones as potent inhibitors of 11- $\beta$ -hydroxysteroid dehydrogenase type 1. *Bioorg. Med. Chem. Lett.* **2007**, *17* (17), 4837-4840.
72. Eliot, A. & Kirsch, J., Pyridoxal phosphate enzymes: Mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* **2004**, *73*, 383 - 415.
73. Jansonius, J. N., Structure, evolution and action of vitamin B6-dependent enzymes. *Curr. Opin. Struct. Biol.* **1998**, *8* (6), 759-769.
74. Dunathan, H. C., Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc. Natl. Acad. Sci. U S A.* **1966**, *55* (4), 712-716.
75. Christen, P. & Mehta, P. K., From cofactor to enzymes. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Chem. Rec.* **2001**, *1* (6), 436-447.
76. Cleland, W. W., The kinetics of enzyme-catalyzed reactions with two or more substrates or products: II. Inhibition: Nomenclature and theory. *Biochim. Biophys. Acta* **1963**, *67* (0), 173-187.
77. Cleland, W. W., The kinetics of enzyme-catalyzed reactions with two or more substrates or products: I. Nomenclature and rate equations. *Biochim. Biophys. Acta* **1963**, *67* (0), 104-137.
78. Henson, C. P. & Cleland, W. W., Kinetic studies of glutamic oxaloacetic transaminase Isozymes. *Biochemistry* **1964**, *3* (3), 338-345.
79. Humble, M. S.; Cassimjee, K. E.; Håkansson, M.; Kimbung, Y. R.; Walse, B.; Abedi, V.; Federsel, H.-J.; Berglund, P.; Logan, D. T., Crystal structures of the *Chromobacterium violaceum*  $\omega$ -transaminase reveal major structural rearrangements upon binding of coenzyme PLP. *FEBS J.* **2012**, *279* (5), 779-792.
80. Sayer, C.; Isupov, M. N.; Westlake, A.; Littlechild, J. A., Structural studies of *Pseudomonas* and *Chromobacterium*  $\omega$ -aminotransferases provide insights into their differing substrate specificity. *Acta Cryst. D* **2013**, *69* (4), 564-576.
81. Steffen-Munsberg, F.; Vickers, C.; Thontowi, A.; Schätzle, S.; Tumlirsch, T.; Svedendahl Humble, M.; Land, H.; Berglund, P.; Bornscheuer, U. T.; Höhne, M., Connecting unexplored protein crystal structures to enzymatic function. *ChemCatChem* **2013**, *5* (1), 150-153.
82. Höhne, M.; Schätzle, S.; Jochens, H.; Robins, K.; Bornscheuer, U. T., Rational assignment of key motifs for function guides *in silico* enzyme identification. *Nat. Chem. Biol.* **2010**, *6* (11), 807-813.
83. Schätzle, S.; Steffen-Munsberg, F.; Thontowi, A.; Höhne, M.; Robins, K.; Bornscheuer, U. T., Enzymatic asymmetric synthesis of enantiomerically pure aliphatic, aromatic and

- arylaliphatic amines with (*R*)-selective amine transaminases. *Adv. Synth. Catal.* **2011**, *353* (13), 2439-2445.
84. Ford, G. C.; Eichele, G.; Jansonius, J. N., Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proc. Natl. Acad. Sci. U S A.* **1980**, *77* (5), 2559-2563.
85. Shen, B. W.; Hennig, M.; Hohenester, E.; Jansonius, J. N.; Schirmer, T., Crystal structure of human recombinant ornithine aminotransferase. *J. Mol. Biol.* **1998**, *277* (1), 81-102.
86. Shin, J.-S. & Kim, B.-G., Kinetic resolution of  $\alpha$ -methylbenzylamine with  $\omega$ -transaminase screened from soil microorganisms: Application of a biphasic system to overcome product inhibition. *Biotechnol. Bioeng.* **1997**, *55* (2), 348-358.
87. Shin, J.-S. & Kim, B.-G., Asymmetric synthesis of chiral amines with  $\omega$ -transaminase. *Biotechnol. Bioeng.* **1999**, *65* (2), 206-211.
88. Shin, J. S.; Yun, H.; Jang, J. W.; Park, I.; Kim, B. G., Purification, characterization, and molecular cloning of a novel amine:pyruvate transaminase from *Vibrio fluvialis* JS17. *Appl. Microbiol. Biotechnol.* **2003**, *61* (5-6), 463-471.
89. Park, H. H. & Jang, T., PDB: 3NUI, **2011**.
90. Jaeger, K. E.; Liebeton, K.; Zonta, A.; Schimossek, K.; Reetz, M. T., Biotechnological application of *Pseudomonas aeruginosa* lipase: efficient kinetic resolution of amines and alcohols. *Appl. Microbiol. Biotechnol.* **1996**, *46* (2), 99-105.
91. Reetz, M. T. & Dreisbach, C., Highly efficient lipase-catalyzed kinetic resolution of chiral amines. *CHIMIA* **1994**, *48* (12), 570-570.
92. Blaser, H.-U., The chiral switch of (*S*)-metolachlor: a personal account of an industrial odyssey in asymmetric catalysis. *Adv. Synth. Catal.* **2002**, *344* (1), 17-31.
93. Mathew, S. & Yun, H.,  $\omega$ -Transaminases for the production of optically pure amines and unnatural amino acids. *ACS Catal.* **2012**, *2* (6), 993-1001.
94. Höhne, M.; Kühl, S.; Robins, K.; Bornscheuer, U. T., Efficient asymmetric synthesis of chiral amines by combining transaminase and pyruvate decarboxylase. *ChemBioChem* **2008**, *9* (3), 363-365.
95. Matcham, G.; Bhatia, M.; Lang, W.; Lewis, C.; Nelson, R.; Wang, A.; Wu, W., Enzyme and reaction engineering in biocatalysis: synthesis of (*S*)-methoxyisopropylamine (= (*S*)-1-methoxypropan-2-amine). *CHIMIA* **1999**, *53* (12), 584-589.
96. Koszelewski, D.; Pressnitz, D.; Clay, D.; Kroutil, W., Deracemization of mexiletine biocatalyzed by  $\omega$ -transaminases. *Org. Lett.* **2009**, *11* (21), 4810-4812.
97. Iwasaki, A.; Yamada, Y.; Kizaki, N.; Ikenaka, Y.; Hasegawa, J., Microbial synthesis of chiral amines by (*R*)-specific transamination with *Arthrobacter sp.* KNK168. *Appl. Microbiol. Biotechnol.* **2006**, *69* (5), 499-505.
98. Koszelewski, D.; Lavandera, I.; Clay, D.; Rozzell, D.; Kroutil, W., Asymmetric synthesis of optically pure pharmacologically relevant amines employing  $\omega$ -transaminases. *Adv. Synth. Catal.* **2008**, *350* (17), 2761-2766.
99. Cho, B.-K.; Park, H.-Y.; Seo, J.-H.; Kim, J.; Kang, T.-J.; Lee, B.-S.; Kim, B.-G., Redesigning the substrate specificity of  $\omega$ -aminotransferase for the kinetic resolution of aliphatic chiral amines. *Biotechnol. Bioeng.* **2008**, *99* (2), 275-284.
100. Shin, J.-S. & Kim, B.-G., Exploring the active site of amine:pyruvate aminotransferase on the basis of the substrate structure-reactivity relationship: How the enzyme controls substrate specificity and stereoselectivity. *J. Org. Chem.* **2002**, *67* (9), 2848-2853.
101. Cahn, R. S.; Ingold, C.; Prelog, V., Specification of molecular chirality. *Angew. Chem. Int. Ed.* **1966**, *5* (4), 385-415.
102. Cronin, C. N. & Kirsch, J. F., Role of arginine-292 in the substrate specificity of aspartate aminotransferase as examined by site-directed mutagenesis. *Biochemistry* **1988**, *27* (12), 4572-4579.
103. Hayashi, H.; Kuramitsu, S.; Inoue, Y.; Morino, Y.; Kagamiyama, H., [Arg292  $\rightarrow$  val] or [Arg292  $\rightarrow$  leu] mutation enhances the reactivity of *Escherichia coli* aspartate aminotransferase with aromatic amino acids. *Biochem. Biophys. Res. Commun.* **1989**, *159* (1), 337-342.

- 
104. Onuffer, J. J. & Kirsch, J. F., Redesign of the substrate specificity of *Escherichia coli* aspartate aminotransferase to that of *Escherichia coli* tyrosine aminotransferase by homology modeling and site-directed mutagenesis. *Protein Sci.* **1995**, *4* (9), 1750-1757.
  105. Almo, S. C.; Smith, D. L.; Danishefsky, A. T.; Ringe, D., The structural basis for the altered substrate specificity of the R292D active site mutant of aspartate aminotransferase from *E. coli*. *Protein Eng.* **1994**, *7* (3), 405-412.
  106. Malashkevich, V. N.; Onuffer, J. J.; Kirsch, J. F.; Jansonius, J. N., Alternating arginine-modulated substrate specificity in an engineered tyrosine aminotransferase. *Nat. Struct. Mol. Biol.* **1995**, *2* (7), 548-553.
  107. Sayer, C.; Martinez-Torres, R. J.; Richter, N.; Isupov, M. N.; Hailes, H. C.; Littlechild, J. A.; Ward, J. M., The substrate specificity, enantioselectivity and structure of the (*R*)-selective amine:pyruvate transaminase from *Nectria haematococca*. *FEBS J.* **2014**, *281* (9), 2240-2253.
  108. Łyskowski, A.; Gruber, C.; Steinkellner, G.; Schürmann, M.; Schwab, H.; Gruber, K.; Steiner, K., Crystal structure of an (*R*)-selective  $\omega$ -transaminase from *Aspergillus terreus*. *PLoS ONE* **2014**, *9* (1), e87350.
  109. Hanson, K. R., Applications of the sequence rule. I. Naming the paired ligands g,g at a tetrahedral atom Xggij. II. Naming the two faces of a trigonal atom Yghi. *J. Am. Chem. Soc.* **1966**, *88* (12), 2731-2742.
  110. Kobayashi, K.; Miyazawa, S.; Endo, A., Isolation and inhibitory activity of gabaculine, a new potent inhibitor of  $\gamma$ -aminobutyrate aminotransferase produced by a *Streptomyces*. *FEBS Lett.* **1977**, *76* (2), 207-210.
  111. Rando, R. R., Mechanism of the irreversible inhibition of  $\gamma$ -aminobutyric acid- $\alpha$ -ketoglutaric acid transaminase by the neurotoxin gabaculine. *Biochemistry* **1977**, *16* (21), 4604-4610

## 7. Publikationsliste

Im Rahmen dieser Arbeit diskutierte Publikationen und Anteile der jeweiligen Autoren:

- [I] M. Gall, M. Thomsen, C. Peters, I.V. Pavlidis, P. Jonczyk, P.P. Grünert, S. Beutel, T. Scheper, E. Gross, M. Backes, T. Geißler, J.P. Ley, J.M. Hilmer, G. Krammer, G.J. Palm, W. Hinrichs, U.T. Bornscheuer. Enzymatic conversion of flavonoids using bacterial chalcone isomerase and enoate reductase. *Angew. Chem. Int. Ed.* **2014**, 53 (5), 1439-1442. Enzymatische Umsetzung von Flavonoiden mit einer bakteriellen Chalconisomerase und einer Enoatreduktase, *Angew. Chem.* **2014**, 126 (5), 1463-1466.

M. Gall: Identifizierung und Klonierung der Gene, Unterstützung bei der Aufreinigung und Charakterisierung der CHI, Verfassen des Manuskriptes; M. Thomsen: Aufreinigung und Charakterisierung der CHI, Proteinkristallisation, Daten-Sammlung und -Prozessierung, Strukturlösung, Refinement und Analyse der Struktur, Verfassen des Manuskriptes, deutsche Übersetzung; C. Peters: Untersuchungen und Klonierung der ERED, Durchführung Biokatalysen alternativer Substrate, Verfassen des Manuskriptes; I.V. Pavlidis: Optimierung der Expression der ERED, Verfassen des Manuskriptes; E. Gross, T. Geißler: Projektkoordinierung, Analytik der alternativen Substrate; M. Backes: chemische Synthese des Naringeninchalcons; G.J. Palm: Unterstützung bei der Datensammlung und -Prozessierung, Strukturlösung, Refinement; U.T. Bornscheuer: Verfassen des Manuskriptes; T. Scheper, J.P. Ley, J.M. Hilmer, G. Krammer, U.T. Bornscheuer: Initiierung des Projektes. Alle Autoren waren an der Korrektur und Diskussion des Manuskriptes beteiligt.

- [II] M. Thomsen, A. Tuukkanen, J. Dickerhoff, G.J. Palm, H. Kratzat, D. Svergun, K. Weisz, U.T. Bornscheuer, W. Hinrichs. Structure and catalytic mechanism of the evolutionarily unique bacterial chalcone isomerase. *submitted*

M. Thomsen: Mutagenese, Expression, Aufreinigung, Bestimmung kinetischer Parameter; Proteinkristallisation, Daten-Sammlung und -Prozessierung, Strukturlösung, Refinement und Analyse der Struktur, Korrektur und Diskussion des Manuskriptes ; G.J. Palm: Unterstützung Daten-Sammlung und -Prozessierung, Strukturlösung, Refinement und Analyse der Struktur; H. Kratzat: Klonierung CHI<sub>ΔC-term</sub> und T<sub>m</sub>-Bestimmung; A. Tuukkanen, D. Svergun: Durchführung und Analyse der SAXS-Experimente; J. Dickerhoff, K. Weisz: Durchführung der NMR-Experimente; U.T. Bornscheuer: Initiierung des Projektes, Korrektur und Diskussion des Manuskriptes; W. Hinrichs: Initiierung des Projektes, Analyse der Struktur, Verfassen und Diskussion des Manuskriptes.

- [III] M. Thomsen\*, L. Skalden\*, G.J. Palm, M. Höhne, U.T. Bornscheuer, W. Hinrichs. Crystallization and preliminary X-ray diffraction studies of the (R)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Cryst. F* **2013**, 69 (12), 1415-1417.

M. Thomsen: Proteinkristallisation, Daten-Sammlung, -Prozessierung und -Auswertung, Verfassen und Diskussion des Manuskriptes; L. Skalden: Protein-Expression und -Reinigung, Verfassen und Diskussion des Manuskriptes; G.J. Palm: Unterstützung bei der

Datensammlung; M. Höhne, U.T. Bornscheuer und W. Hinrichs: Korrektur und Diskussion des Manuskriptes.

- [IV] M. Thomsen\*, L. Skalden\*, G.J. Palm, M. Höhne, U.T. Bornscheuer, W. Hinrichs. Crystallographic characterization of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Cryst. D* **2014**, *70* (4), 1086-1093.

M. Thomsen: Proteinkristallisation, Daten-Sammlung und -Prozessierung, Strukturlösung, Refinement und Analyse der Struktur, Verfassen und Diskussion des Manuskriptes; L. Skalden: Mutagenese, Protein-Expression und Reinigung, Docking-Versuche, Verfassen und Diskussion des Manuskriptes; G.J. Palm: Unterstützung bei der Datensammlung; M. Höhne, U.T. Bornscheuer und W. Hinrichs: Korrektur und Diskussion des Manuskriptes.

- [V] L. Skalden\*, M. Thomsen\*, M. Höhne, U.T. Bornscheuer, W. Hinrichs. Structural and biochemical characterization of the dual substrate recognition of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*. *FEBS J.*, *online*. doi: 10.1111/febs.13149.

L. Skalden: Mutagenese, Protein-Expression und Reinigung, Bestimmung der Enzymaktivitäten, Verfassen des Manuskriptes; M. Thomsen: Proteinkristallisation, Soaken des Inhibitors, Daten-Sammlung und -Prozessierung, Strukturlösung, Refinement und Analyse der Struktur, Verfassen und Diskussion des Manuskriptes; U.T. Bornscheuer und W. Hinrichs: Initiierung des Projektes; Alle Autoren waren an der Diskussion der Daten und Korrektur des Manuskriptes beteiligt.

\*gleichberechtigte Erstautoren

# **Publikation I**

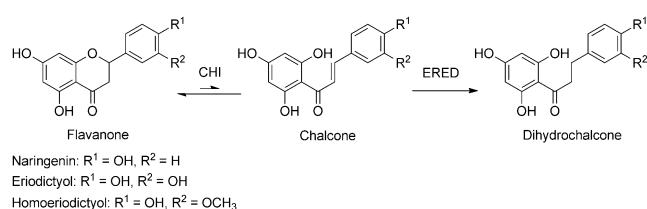
# Enzymatic Conversion of Flavonoids using Bacterial Chalcone Isomerase and Enoate Reductase<sup>\*\*</sup>

Mechthild Gall, Maren Thomsen, Christin Peters, Ioannis V. Pavlidis, Patrick Jonczyk, Philipp P. Grünert, Sascha Beutel, Thomas Scheper, Egon Gross, Michael Backes, Torsten Geißler, Jakob P. Ley,\* Jens-Michael Hilmer, Gerhard Krammer, Gottfried J. Palm, Winfried Hinrichs, and Uwe T. Bornscheuer\*

**Abstract:** Flavonoids are a large group of plant secondary metabolites with a variety of biological properties and are therefore of interest to many scientists, as they can lead to industrially interesting intermediates. The anaerobic gut bacterium *Eubacterium ramulus* can catabolize flavonoids, but until now, the pathway has not been experimentally confirmed. In the present work, a chalcone isomerase (CHI) and an enoate reductase (ERED) could be identified through whole genome sequencing and gene motif search. These two enzymes were successfully cloned and expressed in *Escherichia coli* in their active form, even under aerobic conditions. The catabolic pathway of *E. ramulus* was confirmed by biotransformations of flavanones into dihydrochalcones. The engineered *E. coli* strain that expresses both enzymes was used for the conversion of several flavanones, underlining the applicability of this biocatalytic cascade reaction.

**F**lavonoids are polyphenolic compounds that occur in plants; they are involved in plant coloration and act as biochemical sensing molecules, but also have important roles

as antioxidants<sup>[1]</sup> and flavor enhancers.<sup>[2]</sup> Their biosynthesis in plants is well documented in the literature, and the basic scaffold is formed from malonyl and coumaroyl precursors in the presence of 4-coumaryl-CoA ligase and a chalcone synthase. In the next step, chalcone isomerase (CHI) catalyzes the stereospecific formation of the tricyclic flavanones such as naringenin (Scheme 1). These are further functional-



**Scheme 1.** Postulated pathway for the degradation of flavonoids. CHI=chalcone isomerase, ERED=enoate reductase. The equilibrium lies strongly on the side of flavanone formation.

ized by other enzymes, such as hydroxylases or glycosidases, to create the diverse family of flavonoids. For plant-derived CHIs, their evolutionary origin,<sup>[3]</sup> reaction mechanism,<sup>[4]</sup> and structures<sup>[5]</sup> have been extensively studied. More recently, it was found that CHI can also be found in some gammaproteobacteria and ascomycetes, but their physiological role remains to be determined.<sup>[6]</sup> One exception is the anaerobic gut microorganism *Eubacterium ramulus*. It has been reported that this strain is able to degrade a range of flavonoids, including naringenin-7-neohesperidoside, a glycosylated derivative of naringenin.<sup>[7]</sup> The authors suggested a possible degradation pathway, but did not present experimental verification of key intermediate steps. Furthermore, phloretin ( $R^1 = OH, R^2 = H$ ; Scheme 1) is hydrolyzed in *E. ramulus* to 3-(4-hydroxyphenyl)propanoic acid and phloroglucinol, which is an undesirable pathway, as phloretin, as a degradation product of naringenin, is a very important flavor compound.<sup>[8]</sup> Although phloretin can in principle be chemically produced by reduction of the corresponding chalcone or by a Friedel-Crafts-type acylation of phenol with dihydrocinnamic acid,<sup>[9]</sup> these chemically synthesized products cannot be used as natural flavoring substances according to European flavor legislation. Therefore, an enzymatic or fermentative process for phloretin production would hold additional value for the flavor and fragrance

[\*] Dipl.-Biol. M. Gall, Dipl.-Biochem. C. Peters, Dr. I. V. Pavlidis, Prof. Dr. U. T. Bornscheuer  
Institute of Biochemistry  
Department of Biotechnology & Enzyme Catalysis  
Greifswald University  
Felix-Hausdorff-Strasse 4, 17487 Greifswald (Germany)  
E-mail: uwe.bornscheuer@uni-greifswald.de  
Homepage: <http://biotech.uni-greifswald.de>

Dipl.-Biochem. M. Thomsen, Dr. G. J. Palm, Prof. Dr. W. Hinrichs  
Institute of Biochemistry, Department of Structural Biology  
Greifswald University  
Felix-Hausdorff-Strasse 4, 17487 Greifswald (Germany)

M. Sc. P. Jonczyk, M. Sc. P. P. Grünert, Dr. S. Beutel,  
Prof. Dr. T. Scheper  
Institute of Technical Chemistry  
Gottfried Wilhelm Leibniz University of Hannover  
Callinstrasse 5, 30167 Hannover (Germany)

E. Gross, Dr. M. Backes, Dr. T. Geißler, Dr. J. P. Ley, Dr. J.-M. Hilmer,  
Dr. G. Krammer  
Symrise, P.O. Box 1253, 37603 Holzminden (Germany)  
E-mail: jakob.ley@symrise.com

[\*\*] We thank the “Bundesministerium für Bildung und Forschung” for financial support within the “Biokatalyse 2021” cluster (FKZ: 0315365 and 031A109). M.T. thanks the “Landesgraduiertenkolleg of Mecklenburg-Vorpommern” for financial support.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201306952>.

industry. Alternatively, glycosides of phloretin, such as phlorizin, can be obtained from *Malus* ssp.<sup>[10]</sup> and are subsequently hydrolyzed to yield the aglycon phloretin.<sup>[11]</sup> However, this established process is time- and cost-intensive and also depends on the seasonal availability of the starting material. To avoid the seasonal impact and to simultaneously reduce production costs and waste streams, a biotechnological route for this important flavor compound is of high interest.

The cultivation of the strictly anaerobic *E. ramulus* DSM 16296 for the setup of an industrially useful bioprocess is an obstacle, as the growth is very slow, and whole-cell biotransformations lead to further degradation of phloretin.<sup>[12]</sup> These problems can be overcome by the heterologous expression of enzymes in microorganisms. This technique is already in use for the production of flavonoids from various substrates, such as L-phenylalanine, tyrosine, or cinnamic acid, but only for genes derived from plants.<sup>[13]</sup> In our present study, we came across several challenges, including the identification of the genes from the *E. ramulus* genome that encode the enzymes that are involved in the biocatalytic transformation of naringenin into phloretin, and the functional expression of the proteins in a recombinant host.

Herles and co-workers<sup>[14]</sup> determined a fragment of 15 amino acids at the N terminus of the CHI from *E. ramulus* DSM 16296 that is proposed to be involved in the conversion of naringenin into phloretin. This peptide sequence was used in the present study for the identification of the whole gene. Therefore, we sequenced the entire genome of *E. ramulus* DSM 16296 and used the resulting contigs to identify the encoding gene for the CHI. Only one hit was found, which aligned twelve consecutive residues of the published sequence directly after the starting methionine. This open reading frame (ORF; GenBank Accession Number: KF154734) corresponds to a protein of 32.5 kDa, which correlates well with the size that Herles and co-workers determined by SDS-PAGE for one subunit of the enzyme.<sup>[14]</sup> Interestingly, a BLAST<sup>[15]</sup> search did not identify any other CHIs from plant or other sources with a sequence identity of  $\geq 10\%$  compared to this bacterial CHI; hence, this protein is unique (Supporting Information, Figure S1).

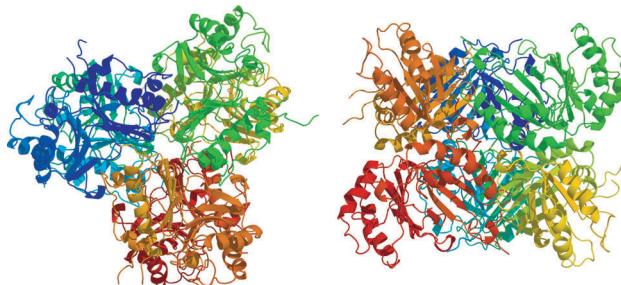
For the enoate reductase (ERED), no gene information was available. For this reason, we aligned 34 sequences of known EREDs and found a conserved motif that served as a basis to also identify the ERED-encoding gene in the *E. ramulus* DSM 16296 genome. Again, we found only one ORF (GenBank Accession Number: KF154735), its amino acid sequence shared low identity ( $\leq 29\%$ ; Figure S2) with known EREDs.<sup>[16]</sup> However, all of these other EREDs have been only insufficiently characterized.

The genes that encode the bacterial CHI and ERED were cloned into common pET vectors, and functional expression in *E. coli* Rosetta was performed. We were pleased to find that the CHI from *E. ramulus* was expressed in soluble form under aerobic conditions in the *E. coli* Rosetta (DE3) strain (Figure S3). Slight modifications of the purification protocol described by Herles and co-workers<sup>[14]</sup> provided us with pure recombinant CHI for biochemical characterization and subsequent crystallographic analysis (Figure S4). The same purification protocol was also applied to the wild-type protein

from *E. ramulus*; however, lower specific activities were found (Tables S1 and S2). The characterization of CHI revealed that the recombinant CHI maintains more than 90% of its activity at pH 6.4–7.6 (Figure S5), whereas the optimal operational temperature is 45°C (Figure S6). The recombinant CHI also exhibited satisfactory stability, maintaining more than 50% of its initial activity after incubation at 41°C for six hours (Figure S7).

The kinetic constants that were determined for the recombinant CHI underline the high activity of the enzyme (Figure S8). Although the  $K_m$  value of the recombinant bacterial CHI (36.9 μM) lies between those of the plant CHIs, which vary from 2 μM to 112 μM for the CHI from *Glycine max*<sup>[17]</sup> and *Medicago sativa*,<sup>[4]</sup> the turnover number is significantly higher. The  $k_{cat}$  values that have been reported in the literature for plant CHIs vary from 186 s<sup>-1</sup><sup>[4]</sup> to 833 s<sup>-1</sup><sup>[18]</sup>, whereas a value of 4483 s<sup>-1</sup> was determined for the recombinant CHI from *E. ramulus*. Hence, the recombinantly expressed bacterial CHI shows a catalytic efficiency of  $1.2 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ , which is 75-times higher than the plant one from *Medicago sativa*, whose catalytic efficiency is  $1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ <sup>[4]</sup>. When we compared the recombinantly expressed CHI with the native CHI that was produced in *E. ramulus*, the kinetic values differed. Herles and co-workers determined the  $K_m$  to be 42.7 μM with a  $k_{cat}$  value of 2300 s<sup>-1</sup>, which resulted in a catalytic efficiency of  $0.5 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ ,<sup>[14]</sup> which is significantly lower than the one of the recombinant enzyme; this difference is mainly due to the different turnover numbers. The higher catalytic activity of the recombinant CHI can be attributed to the higher purity or to the formation of different multimers.

Analysis of the X-ray crystal structure (PDB code: 3zph) revealed that the recombinant CHI forms hexamers as trimers of three dimer units (Figure 1), whereas Herles and



**Figure 1.** Crystal structure (left: top view; right: turned by 90°) of a hexamer with  $D_3$  symmetry of the recombinant CHI from *E. ramulus*; PDB code: 3zph.

co-workers stated that only tetramers were observed in *E. ramulus*.<sup>[14]</sup> Regardless of this observation, the structure of this bacterial CHI differs significantly from those of plant CHIs; the protein–protein interactions of a dimer unit of the recombinant bacterial CHI are very strong owing to the incorporation of the C terminus of one protein chain into the second protein chain, whereas the plant enzymes mainly occur as monomers.<sup>[5]</sup> Moreover, in contrast to plant CHIs, an internal symmetry exists in the tertiary structure of a mono-

mer, as it can be divided into two domains with similar folds, which are rotated by 90° against each other.

In contrast to the expression of the CHI, expression of the ERED was more challenging. As the ERED gene with the native nucleotide sequence showed only a low expression level in *E. coli* Rosetta, a synthetic codon-optimized gene of ERED (sERED) was cloned into vector pET22b, which led to overexpression of soluble sERED at 20°C under aerobic conditions. The protein had the expected molecular weight of 70 kDa, as determined by SDS-PAGE, which correlates well with the theoretical size of 75 kDa. Unfortunately, only low activity was observed for the conversion of the naringenin chalcone into phloretin. As the enzyme originates from a strictly anaerobic bacterium, we performed the expression under anaerobic conditions. Under nitrogen atmosphere, the highest expression level was observed at 25°C. After ensuring that all steps (cell harvest, lysis, and biocatalysis) were performed under nitrogen atmosphere, significant enzymatic activity in the reduction of the chalcone to phloretin could be monitored. Furthermore, a crude cell extract that contained the sERED could be added to (purified) CHI, and conversion from naringenin into phloretin was observed. Pleasingly, the undesired enzymatic degradation of phloretin that was observed with *E. ramulus* did not take place with the *E. coli* system, thus enabling the synthesis of the target product.

The high oxygen sensitivity of the ERED from *E. ramulus* might be explained by its sequence similarity to a 2,4-dienoyl-CoA reductase and the anaerobic enoate reductase from *Clostridium* sp.<sup>[16a]</sup> These two enzymes are described as multi-domain proteins with a barrel domain that is related to the “Old Yellow Enzyme” with strong sequence conservation in a core region of approximately 40 amino acids. The second domain of this ERED is also redox-active and related to glutathione reductases with an iron–sulfur cluster, and could therefore be the oxygen-sensitive part of the protein. In this sequence, the four cysteines that are required for the iron–sulfur cluster were found in the typical motif CXXCXXC(X)<sub>22</sub>C, a motif that also appears in the ERED from *E. ramulus*, starting at C361 (Figure S2).<sup>[14b]</sup> It should be noted that the sequence motif GXGXXG(X)<sub>17</sub>E for the NADH and FAD binding site in glutathione reductase<sup>[19]</sup> is found twice in the sequence of *E. ramulus* ERED.

To facilitate the industrial application of this bioprocess, we envisaged the simultaneous expression of both enzymes in an *E. coli* strain. To achieve this, the CHI gene was cloned with a ribosome binding site directly behind the sERED gene on the pET22b vector. With this new construct (pET22b\_sERED/CHI), the soluble and active expression of both enzymes under anaerobic cultivation conditions was possible. Biocatalysis with the crude cell extract allowed the production of 50 μM phloretin after one hour under anaerobic conditions. Under aerobic conditions, both enzymes can be solubly expressed, although a significant amount of sERED is produced in its insoluble form (Figure S9). However, the amount of soluble sERED that was produced with this cultivation method is large enough to push the equilibrium towards the production of phloretin.

To evaluate the substrate specificity of this system and thus its versatility, *E. coli* BL21(DE3) cells with the construct

pET22b\_sERED/CHI were cultivated aerobically (see the Supporting Information), and biocatalytic transformations of three structurally similar flavanones were performed by whole cells under nitrogen atmosphere, as the aerobic process led to lower and reversible activity. High conversions were observed for naringenin, eriodictyol, and homoeriodictyol (Scheme 1) within a short period of time (Table 1). The

**Table 1:** Conversion [%] of flavanones into their respective dihydrochalcones.

Substrate	Reaction time		
	1 h	2 h	17 h
Naringenin	69 ± 1 %	86 ± 1 %	93 ± 1 %
Eriodictyol	46 ± 2 %	51 ± 4 %	72 ± 2 %
Homoeriodictyol	47 ± 3 %	52 ± 1 %	63 ± 2 %

hydroxyl group in the *para* position of the phenyl ring appears to be crucial for CHI activity; docking experiments of naringenin at the active site of the CHI showed that this hydroxyl group participates in a hydrogen-bonding network with Asp79 and Gln101, which is crucial for the right orientation of the substrate. On the other hand, these results show that the CHI seems to be tolerant towards substituents at the *meta* position of the phenol ring. Even though naringenin is the preferred substrate, as it only bears a hydrogen atom at this position, it seems that the active site of CHI can accommodate bulkier substituents, such as hydroxyl (eriodictyol) or methoxy (homoeriodictyol) groups.

In summary, the results of our experiments confirmed the metabolic pathway that was proposed by Herles and co-workers<sup>[14]</sup> for the degradation of flavonoids by the identification and successful recombinant expression of the chalcone isomerase and an enoate reductase from the anaerobic bacterium *Eubacterium ramulus*. The engineered *E. coli* strain that expresses both enzymes can be used for the conversion of several flavanones, which underlines the applicability of the biocatalytic system that was developed in this study.

Received: August 7, 2013

Revised: November 14, 2013

Published online: December 20, 2013

**Keywords:** chalcone isomerase · enoate reductase · enzyme biocatalysis · *Eubacterium ramulus* · flavonoids

- [1] A. Crozier, I. B. Jaganath, M. N. Clifford, *Nat. Prod. Rep.* **2009**, 26, 1001–1043.
- [2] a) J. P. Ley, G. Krammer, G. Reinders, I. L. Gatfield, H.-J. Bertram, *J. Agric. Food Chem.* **2005**, 53, 6061–6066; b) S. Scharbert, T. Hofmann, *J. Agric. Food Chem.* **2005**, 53, 5377–5384; c) M. S. J. Simmonds, *Phytochemistry* **2001**, 56, 245–252.
- [3] a) M. N. Ngaki, G. V. Louie, R. N. Philippe, G. Manning, F. Pojer, M. E. Bowman, L. Li, E. Larsen, E. S. Wurtele, J. P. Noel, *Nature* **2012**, 485, 530–533; b) W. A. Peer, *Nat. Chem. Biol.* **2012**, 8, 607–608.
- [4] J. M. Jez, J. P. Noel, *J. Biol. Chem.* **2002**, 277, 1361–1369.

- [5] J. M. Jez, M. E. Bowman, R. A. Dixon, J. P. Noel, *Nat. Struct. Biol.* **2000**, *7*, 786–791.
- [6] M. Gensheimer, A. Mushegian, *Protein Sci.* **2004**, *13*, 540–544.
- [7] H. Schneider, M. Blaut, *Arch. Microbiol.* **2000**, *173*, 71–75.
- [8] “Flavor suppression and enhancement”: J. P. Ley, K. Reichelt, G. Krammer in *Food Flavors. Chemical, Sensory and Technological Properties* (Ed.: H. H. Jelen), CRC Press, Boca Raton, USA, **2012**, pp. 155–175.
- [9] V. Siddaiah, C. V. Rao, S. Venkateswarlu, G. V. Subbaraju, *Tetrahedron* **2006**, *62*, 841–846.
- [10] F. Will, H. Zessner, H. Becker, H. Dietrich, *Food Sci. Technol. LEB* **2007**, *40*, 1344–1351.
- [11] J. Ley, J.-M. Hilmer, I. L. Gatfield, European Patent 2017272, **2008**.
- [12] P. Jonczyk, A. Schmidt, I. Bice, M. Gall, E. Gross, J.-M. Hilmer, U. Bornscheuer, S. Beutel, T. Schepers, *Chem. Ing. Tech.* **2011**, *83*, 2147–2152.
- [13] a) M. Koffas, E. Leonard, Y. Yan, J. Chemler, Patent WO/2006/010117, **2006**; b) C. Schmidt-Dannert, K. Watts, Patent WO/2005/084305, **2005**.
- [14] C. Herles, A. Braune, M. Blaut, *Arch. Microbiol.* **2004**, *181*, 428–434.
- [15] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. Madden, *BMC Bioinf.* **2009**, *10*, 421.
- [16] a) R. E. Williams, N. C. Bruce, *Microbiology* **2002**, *148*, 1607–1614; b) F. Rohdich, A. Wiese, R. Feicht, H. Simon, A. Bacher, *J. Biol. Chem.* **2001**, *276*, 5779–5787.
- [17] L. Ralston, S. Subramanian, M. Matsuno, O. Yu, *Plant Physiol.* **2005**, *137*, 1375–1388.
- [18] R. A. Bednar, J. R. Hadcock, *J. Biol. Chem.* **1988**, *263*, 9582–9588.
- [19] O. Dym, D. Eisenberg, *Protein Sci.* **2001**, *10*, 1712–1728.



## Supporting Information

© Wiley-VCH 2014

69451 Weinheim, Germany

### Enzymatic Conversion of Flavonoids using Bacterial Chalcone Isomerase and Enoate Reductase\*\*

*Mechthild Gall, Maren Thomsen, Christin Peters, Ioannis V. Pavlidis, Patrick Jonczyk, Philipp P. Grünert, Sascha Beutel, Thomas Scheper, Egon Gross, Michael Backes, Torsten Geißler, Jakob P. Ley,\* Jens-Michael Hilmer, Gerhard Krammer, Gottfried J. Palm, Winfried Hinrichs, and Uwe T. Bornscheuer\**

anie\_201306952\_sm\_miscellaneous\_information.pdf

## **Supporting Information**

### **Table of Contents**

<b>Experimental section</b>	<b>2</b>
Chemicals and materials	2
Identification of the open reading frames in the <i>E. ramulus</i> genome encoding for the CHI and the ERED	2
Cloning experiments	2
Expression of chalcone isomerase	3
Chalcone isomerase purification	3
Activity assay for the characterization of the recombinant chalcone isomerase	3
Standard procedure for cultivation and expression	3
Biocatalysis and HPLC analysis	4
SDS-PAGE analysis	4
<b>Results</b>	<b>5</b>
Figure SI-1	5
Figure SI-2	6
Figure SI-3	7
Figure SI-4	7
Figure SI-5	7
Figure SI-6	8
Figure SI-7	8
Figure SI-8	9
Figure SI-9	9
Table SI-1	10
Table SI-2	10
<b>Reference</b>	<b>10</b>

## **Experimental section**

### ***Chemicals and materials***

(*R,S*)-Naringenin (95%) and phloretin (99%) were purchased from Sigma (Steinheim, Germany). Eriodictyol (>90%) and homoeriodictyol (>90%) were purchased from Carl Roth (Karlsruhe, Germany). Naringenin chalcone (>95%) was provided by Symrise (Holzminden, Germany). All other chemicals were purchased from Fluka (Buchs, Switzerland), Sigma (Steinheim, Germany), Merck (Darmstadt, Germany), VWR (Hannover, Germany) and Carl Roth (Karlsruhe, Germany). Restriction enzymes and polymerases were obtained from New England Biolabs GmbH (NEB, Frankfurt am Main, Germany). Primers were from Invitrogen™ (life technologies™ GmbH, Darmstadt, Germany). GATC Biotech AG (Konstanz, Germany) performed sequence analysis, while the complete genome sequencing of *Eubacterium ramulus* DSM 16296 was performed by TuTech Innovation GmbH (Hamburg, Germany).

### ***Identification of the open reading frames in the *E. ramulus* genome encoding for the CHI and the ERED***

The 114 contigs of the genome sequencing of *E. ramulus* DSM 16296 were translated and six amino acid sequence bulks were prepared. For the identification of the open reading frame of the CHI gene, the 15 amino acid sequence of the N-terminus presented of Herles and coworkers<sup>[1]</sup> (ADFKFEPMRSLIAWK) was used in order to find a match. No match was found for the whole sequence; however, one match was found containing the 12 first consecutive residues and this had the correct position for a relating open reading frame on DNA level. This open reading frame was designated as the CHI of *E. ramulus* DSM 16296 and the gene was deposited at GenBank (GenBank code: KF154734).

For the ERED, no amino acid sequence was known for this microorganism, so we made an amino acid sequence alignment of 34 published sequences of known EREDS and a conserved motif was identified (NXRXDXXGG). After comparing the translated amino acid sequences of the genome, one ORF was identified containing the conserved motif of the EREDS. This ORF was aligned with the known EREDS and exhibited high homology, indicating that it is a putative ERED. After the activity was certified using a codon-optimized gene, the gene derived from the genome was deposited at GenBank (GenBank code: KF154735).

### ***Cloning experiments***

The amplification of the CHI gene was initialized by PCR using Pfu plus polymerase, genomic DNA of *E. ramulus* DSM 16296 as a template and the following primers:

Forward primer: 5'- CTA ATC GGA TCC GGT ACC ATG GCA GAT TTC AAA TTC GAA CCA ATG - 3'

Reverse primer: 5'- TCA GTA GCG GCC GCT TAT CTC ATG GTG ATG TAT CCA CGA TAA TT - 3'

The resulting DNA fragment (certified by agarose electrophoresis) was cloned into pCR2.1-TOPO using TOPO TA Cloning® (Invitrogen, Carlsbad, California, USA). This vector was subsequently digested with *Nco*I and *Not*I and the CHI gene was ligated into pET28b digested with the same enzymes. The correct sequence was certified by sequence analysis and the resulting plasmid was named pET28b\_CHI.

The amplification of the ERED gene was initialized by PCR using Pfu plus polymerase, genomic DNA of *E. ramulus* DSM 16296 as a template and the following primers:

Forward primer: 5'- GAT CCT CGA GAT GGC AGA AAA AAA TCA GTA TTT TCC ACA- 3'

Reverse primer: 5'- GAT CAA GCT TAG ATA ATT TCC ATT GCT GCG GTC CA- 3'

The resulting DNA fragment (certified by agarose electrophoresis) was cloned into pCR2.1-TOPO using TOPO TA Cloning® (Invitrogen, Carlsbad, California, USA). In order to subclone

the ERED gene in pET52b, a restriction site for *KpnI* was added before its N-terminus, using the following primers:

Forward primer: 5'- GTG TGA TGG GTA CCT GCA GAA TTC GCC- 3'

Reverse primer: 5'- GAT CAA GCT TAG ATA ATT TCC ATT GCT GCG GTC CA- 3'

Subsequently, the ERED gene was isolated from this vector by digestion with *KpnI* and *HindIII* and it was ligated into pET52b digested with the same enzymes. The correct sequence was certified by sequence analysis and the resulting plasmid was named pET52b\_ERED. As the native gene showed only low expression levels in the *E. coli* Rosetta strain, a codon-optimized gene for expression in *E. coli* cloned in pET22b was purchased from Genscript (Piscataway, USA), which was named pET22b\_sERED.

The cloning of the CHI gene behind the sERED gene in pET22b was initialized by the amplification of the CHI gene with the ribosome-binding site from the pET28 construct, with oligonucleotides supplemented with restriction sites for *BamHI* at the N-terminus and *Xhol* at the C-terminus. The primers were the following:

Forward primer (*BamHI*): 5'- GTC TAG GAT CCA GAA ATA ATT TTG TTT AAC TTT AAG AAG GAG A -3'

Reverse primer (*Xhol*): 5'- CTA GTT ATT GCT CAG CGG -3'.

The resulting fragment (certified by agarose electrophoresis) was digested with *BamHI* and *Xhol* and ligated into pET22b\_sERED digested with the same enzymes. The resulting plasmid was named pET22b\_sERED/CHI.

### ***Expression of the chalcone isomerase***

Expression of the CHI gene in *E. coli* Rosetta (DE3) with pET28b\_CHI was initiated by inoculating 500 mL LB<sub>Kan</sub> (1:100 v/v) and incubated at 37°C under shaking (180 rpm) until OD<sub>600</sub> of 1.0. The protein expression was then induced with 0.1 mM IPTG and the cultivation was continued at 20°C for protein expression. After 21 h the cells were harvested by centrifugation (4000 xg, 30 min, 4°C). The pellets were frozen at – 20°C or used directly for protein purification.

### ***Chalcone isomerase purification***

The purification protocol of CHI was based on work by Herles and coworkers.<sup>[1]</sup> The cell pellet was resuspended in sodium phosphate buffer (50 mM, pH 6.8) and sonicated with a Sonopuls sonicator (Bandelin, Germany) on ice. After centrifugation the supernatant was used for protein purification. The first step was a DEAE-Sepharose-column (GE Healthcare) equilibrated with sodium phosphate buffer (50 mM, pH 6.8). Unbound proteins were washed away with three column volumes. Bound proteins were eluted by a linear gradient of 0 to 0.15 M KCl in the buffer, followed by two column volumes of buffer containing 1 M KCl. The flow-rate was 4 mL min<sup>-1</sup>. The fractions that contain chalcone isomerase were determined by the activity assay and combined. Ammonium sulfate was added to the CHI fraction until a final concentration of 1.5 M was reached. This solution was injected on a butyl sepharose column (GE Healthcare) equilibrated with buffer containing 1.5 M ammonium sulfate and the unbound protein was eluted with two column volumes. The bound protein was eluted with a decreasing linear gradient of 1.5–0.3 M ammonium sulfate, followed by one column volume of 0.3 M ammonium sulfate and a final decreasing, linear gradient of 0.3–0.0 M ammonium sulfate. The flow-rate was 4 mL min<sup>-1</sup>. The fractions that contained chalcone isomerase were pooled and injected on a Superdex200 column (HiLoad 16/60, Amersham Pharmacia Biotech) for gel filtration. The elution of the chalcone isomerase was performed with sodium phosphate buffer (50 mM, pH 6.8) at a flow rate of 0.3 mL min<sup>-1</sup>. On average 4-6 mg of pure protein were obtained from 250 mL cultivation volume.

### ***Activity assay for the characterization of the recombinant chalcone isomerase***

The chalcone isomerase activity was evaluated spectrophotometrically by the conversion of naringenin chalcone to naringenin at 368 nm ( $\epsilon = 29,068 \text{ mM}^{-1} \text{ cm}^{-1}$ ) using a 2 mm cuvette. In a standard procedure, 36 µM naringenin chalcone, 1 mg bovine serum albumin (BSA) and an

appropriate amount of enzyme were mixed in a total volume of 1 mL sodium phosphate buffer (50 mM, pH 6.4) to initiate the reaction. Proper arrangements were applied to the protocol in order to identify the pH optimum and to determine the kinetic constants. For determining the temperature activity the buffer with BSA was pre-incubated for 30 min at the individual temperature and the measurement was performed in a tempered cuvette holder. For measuring the temperature stability the enzyme was incubated at the individual temperature and the measurement was performed at room temperature.

#### **Standard procedure for cultivation and expression**

The vector pET22b\_sERED/CHI with the T7 promoter was used for the simultaneous expression of CHI and ERED in *E. coli* BL21(DE3) both in aerobic and anaerobic fermentations. 35 mL LB<sub>Amp</sub> as overnight pre-culture were inoculated using an over-day culture (1:100 v/v) of the respective variant and incubated at 37°C under shaking (180 rpm). 700 mL LB<sub>Amp</sub> were inoculated in a New Brunswick Scientific (Edison, NJ, USA) fermenter using the pre-culture (1:20 v/v) of the respective variant and incubated at 37°C and oxygen saturation of 50% under mechanical stirring (150 rpm). In the case of anaerobic fermentation, nitrogen was supplied and the temperature was decreased to 25°C when OD<sub>600</sub> of 1 mL<sup>-1</sup> was reached. In the case of the aerobic fermentation, only temperature was decreased to 20°C when OD<sub>600</sub> of 2 mL<sup>-1</sup> was reached. Then, the production of the enzymes was induced by the addition of IPTG (0.1 mM). The cultures were incubated at the specified temperature for another 20 h to express the protein and then harvested by centrifugation (6000 xg, 15 min, 4°C) under nitrogen atmosphere. The cells were disrupted with FastPrep with a standard protocol under nitrogen atmosphere. For instance, a sample of 500 µL resuspended cells was added to an Eppendorf tube containing 200 µL glass beads (0.1-0.11 µm) and homogenized in a FastPrep®-24 (MP Biomedicals, USA) for 40 sec with 4 m s<sup>-1</sup>. The cell debris and the glass beads were removed by centrifugation (10500 xg, 10 min, 4°C). The supernatant was transferred to a separate tube (soluble fraction). The pellet was washed from soluble proteins with 1 mL of the same buffer, vortexed and then centrifuged again. The supernatant was discarded and then the pellet was resuspended in 500 µL buffer. The sample was left for 5-10 sec for the glass beads to precipitate before taking sample of the insoluble fragment. Both fractions (soluble and insoluble) were kept at 4°C until use.

#### **Biocatalysis and HPLC analysis**

In a total volume of 1 mL, 480 µL of the soluble fraction were mixed with 440 µL sodium phosphate buffer (50 mM, pH 7.5). To initiate the reaction, 0.5 mg BSA, 2 mM NADH and 150 µM naringenin were added. In the case of whole cell biotransformation, 100 OD of cells were resuspended in 980 µL of sodium phosphate buffer (50 mM, pH 7.5) and 150 µM naringenin were added. The reaction took place at 23°C under air or N<sub>2</sub> atmosphere. Samples were withdrawn at selected time intervals and the reactants were extracted with ethyl acetate. The organic solvent containing the reactants was then evaporated with a SpeedVac and the samples were kept at -20°C until analysed by HPLC. To analyse them, the reactants were resuspended in 200 µL methanol. The analysis was performed with a Zorbax SB-C18 column (4.6 mm \* 50 mm, 1.6 µm, Agilent), at 40°C at a solvent flow of 1.3 mL min<sup>-1</sup> using acetonitrile (solvent A) and water with 0.1% trifluoroacetic acid (solvent B) and the compounds were monitored at 280 nm. The gradient program was as follows: 10% solvent A for the initial 5 min, then increase to 25% in the next 10 min, kept for 5 min; increased to 100% solvent A in 3 min, kept for another 3 min; drop solvent A concentration to 10 instantly and run for another 14 min to equilibrate the column. Retention time: 16.5 min naringenin, 17.1 min phloretin. A similar protocol was used for the other substrates.

#### **SDS-PAGE analysis**

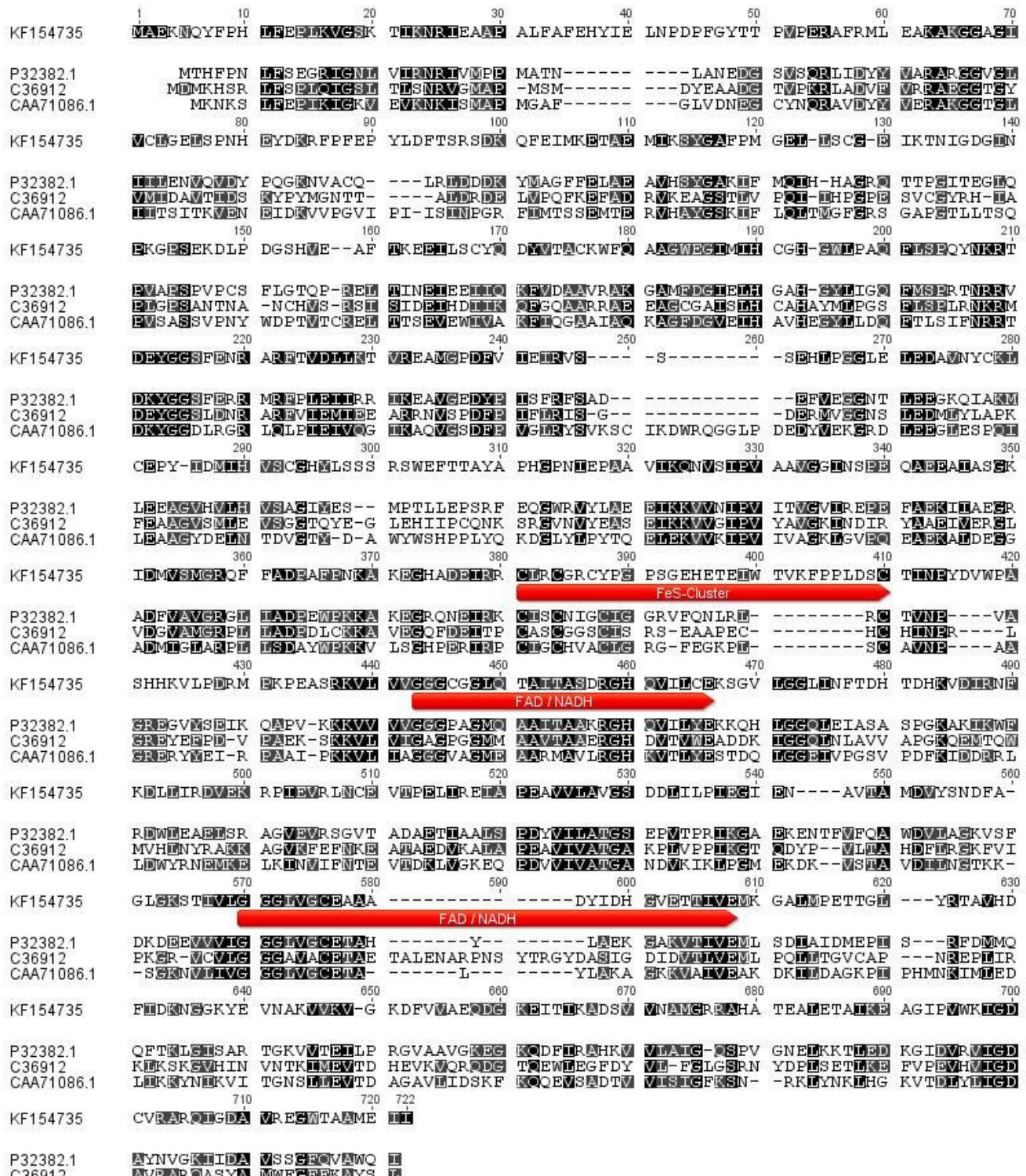
Cultivation samples were removed before induction and in specific time intervals until harvesting, resuspended in 500 µL buffer and homogenized with FastPrep as described

previously to get the soluble and the insoluble fragment. A SDS-PAGE gel was prepared with 10% acrylamide/bis-acrylamide as a resolving gel and 4% acrylamide/bis-acrylamide for the stacking gel. The protein samples run in the gel for about 80 min using 25 mA per gel, 170 V. The gels were then stained with Coomassie Brilliant Blue G-250 or with silver staining and then destained until elimination of the background color.

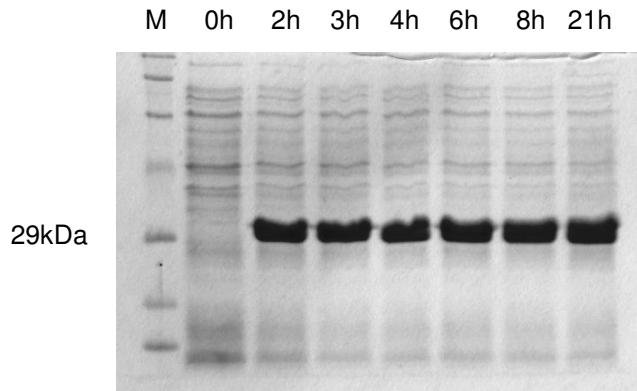
## Results

KF154734	<b>MADFKFEPMR</b>	SLIYVDCVSE	DYRPK <b>LQRWI</b>	<b>YKVHIPDSIS</b>	QFEP <b>YVTKYA</b>	FYPSFP <b>I</b> PPQ
AAA32766.1	<b>MSSSN</b>	ACASPSPFPA	<b>VTKLHVDSVT</b>	<b>FVPSVKSEAS</b>	<b>SNPLFLGGAG</b>	VR---GLDIO
Q9ZWR1.1	<b>MNPS-</b>	-----PS	<b>VTELQVENVT</b>	<b>FTPSVQPEGS</b>	<b>TKSHFLGGAG</b>	ER---GLEIE
AES63019.1	<b>MA T P-</b>	-----S	<b>VTS LAIESIV</b>	<b>FPPTMKAEGS</b>	<b>TNNFFLGGAG</b>	VR---GIOIQ
AAB41524.1	<b>MAAS-</b>	-----	<b>ITAITVENLE</b>	<b>YPAVVTSEVT</b>	<b>GKSYFLGGAG</b>	ER---GLTIE
	70	80	90	100	110	120
KF154734	<b>CDRFGYARMQ</b>	<b>LTEHHWLVSD</b>	<b>LDPRLIEIKAI</b>	<b>AETFPMDVLV</b>	<b>WOGQIPAAAH</b>	TDAQIDSDGD
AAA32766.1	<b>GKFVIFIVIG</b>	<b>VYLEGNAPPS</b>	<b>LSVKWKGKTT</b>	<b>EE-LTESIPF</b>	<b>FREIVVTGAFE</b>	KFIKVTMKLP
Q9ZWR1.1	<b>GKFVKFTAIG</b>	<b>VYLEDDAVPL</b>	<b>LAGWKWKGKTA</b>	<b>EE-LTESVFE</b>	<b>FRDVVTGPFE</b>	KFMKVTMILP
AES63019.1	<b>DKFVKFTAIG</b>	<b>VYLQDIAIPY</b>	<b>LAALKWKWKGP</b>	<b>HK-LTESVPF</b>	<b>FMDIVVTGPFE</b>	KFMRVTMIRP
AAB41524.1	<b>GNFIKFTAIG</b>	<b>VYLEDIAVAS</b>	<b>LAALKWKWKSS</b>	<b>EE-LLETLD</b>	<b>YRDIIISGPFE</b>	KLIRGSKIRE
	130	140	150	160	170	180
KF154734	AGNAARKSNN	<b>AEGNPFTFAF</b>	<b>LPMWWWEKDLK</b>	<b>GKGRTIEDGA</b>	N YRFNMTIGF	<b>PEGVDKAEGE</b>
AAA32766.1	LTGQOYSSEKV	TENCAVAIWKQ	<b>LGLYTDCDEAK</b>	<b>AVEKFLEIFK</b>	-----EETF	PPGSS-----
Q9ZWR1.1	LTGAQYSSEKV	<b>AENCIAIWKQ</b>	<b>FGIYTDAEAK</b>	<b>AIEKFTEVFK</b>	-----DEIF	PPGSS-----
AES63019.1	LTGQEYSNKV	SENCAVAIWKS	<b>LGIYTNEEAK</b>	<b>AIKKFVSVF</b>	-----DET	PPGSS-----
AAB41524.1	LSGPEYSRKV	MENCVAHLKS	<b>VGTYGDAEAE</b>	<b>AMOKFAEAFK</b>	-----PVNF	PPGAS-----
	190	200	210	220	230	240
KF154734	KWLFEKVVPL	<b>LQAAPPECTRV</b>	<b>LASAVKKDIN</b>	<b>GCVMDWVLEI</b>	WFENQSGWYK	VMVDDMKALE
AAA32766.1	-----IL	FAISPT--GS	<b>LTVAFSKDDS</b>	-----	-----	<b>IPETGIAVI</b>
Q9ZWR1.1	-----IL	FTQSP----GS	<b>LTISFSKDGS</b>	-----	-----	-IPKDGVAVI
AES63019.1	-----IL	FTVSPKGGLGS	<b>LTISFSKDGS</b>	-----	-----	-IPEVE TAVI
AAB41524.1	-----VF	YRQSPD--GI	<b>LGLSFSPDTKS</b>	-----	-----	-IPEKEAALI
	250	260	270	280	290	297
KF154734	KPSWAQQDAF	PFIKPYHNVC	SAAVADYTPS	NNLANYRGYI	TMR*	
AAA32766.1	ENKLLEAEAVL	<b>ESIIGKNGVS</b>	<b>PGTRLSVAER</b>	<b>LSQLMMKKNKD</b>	<b>EKEVSDHSLE</b>	<b>EKLAKEN</b>
Q9ZWR1.1	ESNLLSSEAVL	<b>ESMIGKNGVS</b>	<b>PAAKKSLAER</b>	<b>LSALLNVTS</b>	KMK	
AES63019.1	ENKLLSOAVL	<b>ESMIGAHGVS</b>	<b>PAAKQSLASR</b>	<b>LSKLFKEGGN</b>	ANN	
AAB41524.1	ENKAVSSAVL	ETMIGEHAVS	<b>PDLKRCLAA</b>	<b>LPALLNEGAF</b>	KIGN	

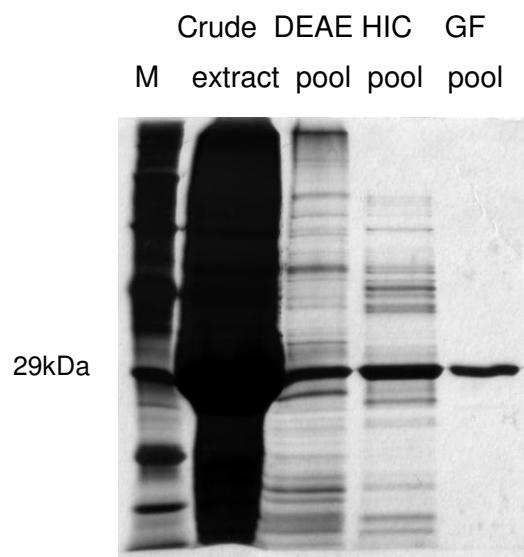
**Figure SI-1:** Multiple sequence alignment of four amino acid sequences of CHIs from plants to *E. ramulus* CHI (KF154734). Residues shared by at least three proteins are marked by contrast inversion. It should be noted that the CHI from *E. ramulus* has less than 10 % identity to the plant enzymes. AAA32766.1, chalcone isomerase [*Arabidopsis thaliana*]; Q9ZWR1.1, chalcone flavonone isomerase [*Citrus sinensis*]; AES63019.1, chalcone-flavonone isomerase [*Medicago truncatula*]; AAB41524.1, chalcone isomerase [*Medicago sativa*].



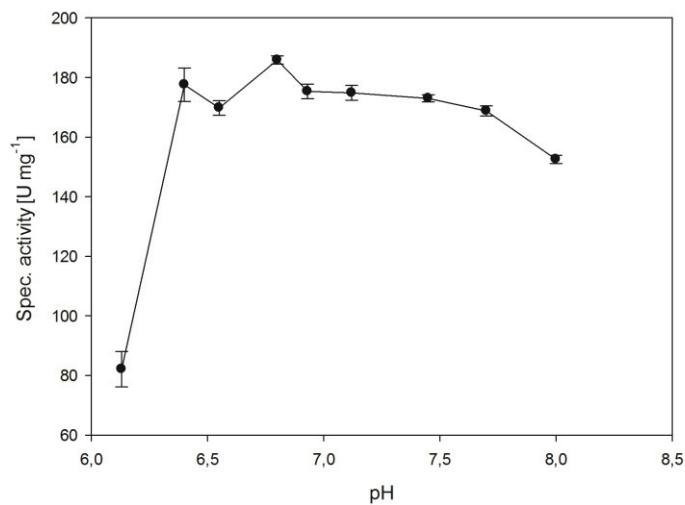
**Figure SI-2:** Multiple sequence alignment of three amino acid sequences of putative proteins similar to ERED from *E. ramulus* (KF154735). Residues shared by at least three proteins are marked by contrast inversion. The motifs for the binding of Fe/S cluster, FAD and NADH are annotated. The ERED from *E. ramulus* shows identity between 15 and 29% compared to the other enoate reductases. C36912, NADH flavinoxidase [*Eubacterium* sp. strain VPI 12708]; CAA71086.1, 2-enoate reductase [*Clostridium tyrobutyricum*]; P32382.1, NADH acceptor oxidoreductase [*Thermoanaerobium brockii*].



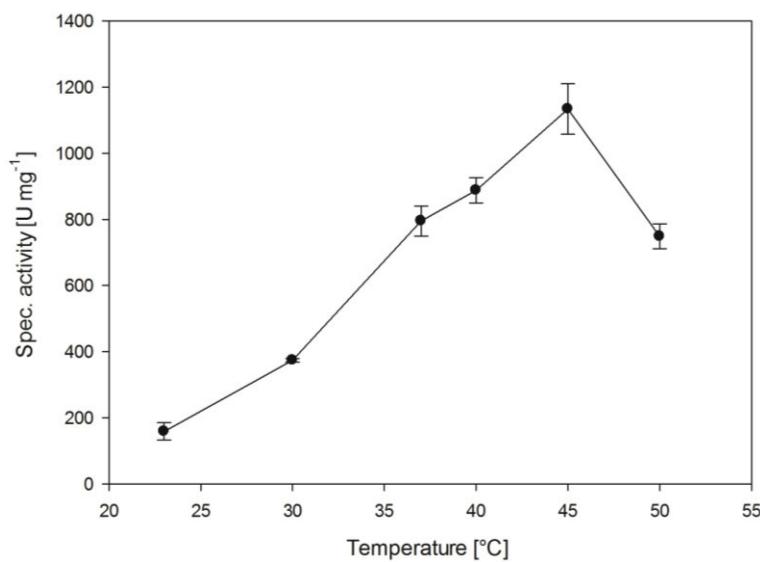
**Figure SI-3:** SDS-PAGE (Coomassie Blue stained) of the soluble fragment of *E. coli* Rosetta after induction of CHI expression with 0.1 mM IPTG.



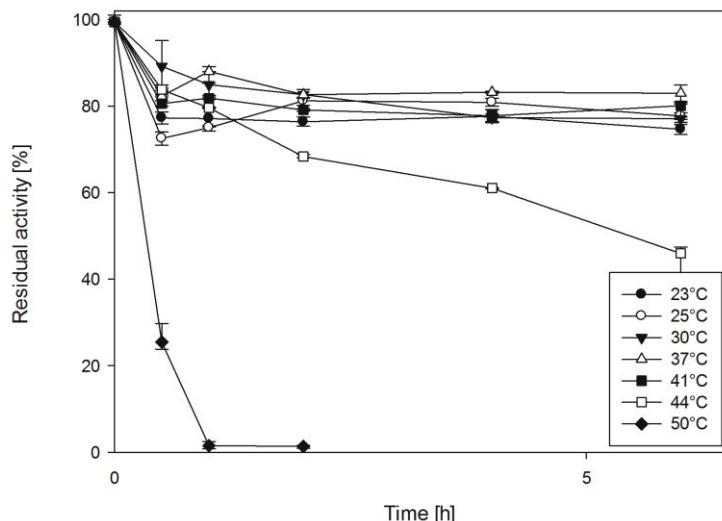
**Figure SI-4:** SDS-PAGE (silver stained) of the purification of recombinant CHI expressed in *E. coli* Rosetta. From left to right: marker; crude extract; DEAE pool; HIC pool; Gel filtration pool.



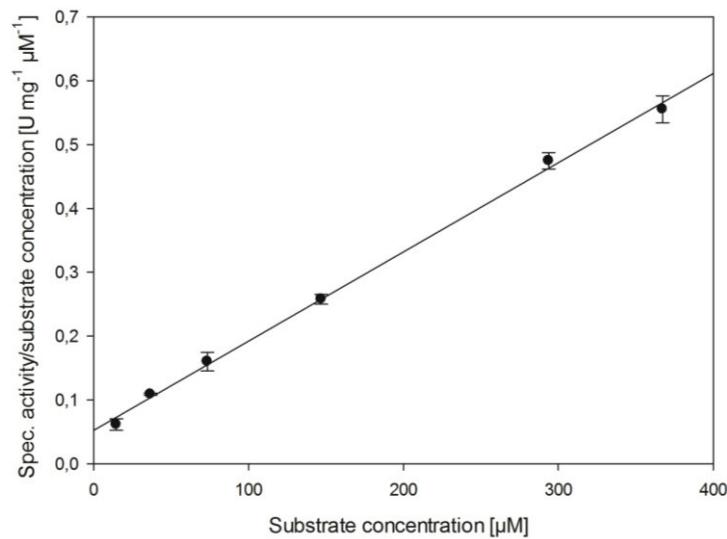
**Figure SI-5:** pH profile of the recombinant CHI using 36  $\mu$ M naringenin chalcone as substrate in sodium phosphate buffer (50 mM).



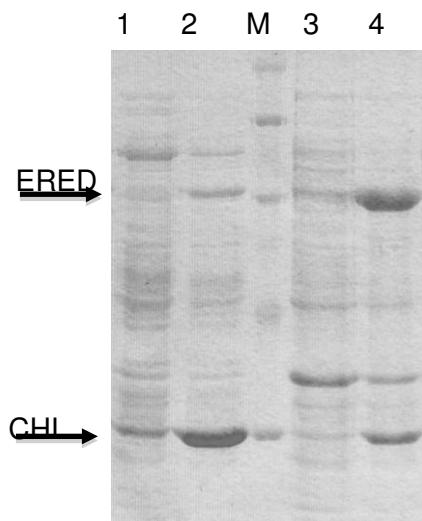
**Figure SI-6:** Temperature activity profile of the recombinant CHI using 36  $\mu$ M naringenin chalcone as substrate in sodium phosphate buffer (50 mM, pH 6.4).



**Figure SI-7:** Thermal stability of the recombinant CHI in sodium phosphate buffer (50 mM, pH 6.4).



**Figure SI-8:** Hanes-Woolf plot of the effect of naringenin chalcone concentration on the specific activity of recombinant CHI at 23°C in sodium phosphate buffer (50 mM, pH 6.4).



**Figure SI-9:** SDS-PAGE (Coomassie Blue stained) of the soluble and insoluble fragment of *E. coli* BL21 (DE3) with pET22\_sERED/CHI; 1) soluble at OD<sub>600</sub> 5 mL<sup>-1</sup>; 2) soluble before harvest; 3) insoluble at OD<sub>600</sub> 5 mL<sup>-1</sup>; 4) insoluble before harvest.

**Table SI-1:** Summary of the purification of the recombinant CHI expressed in *E. coli* Rosetta.

Step	Total protein [mg]	Spec. activity [U mg <sup>-1</sup> ]	Total activity [U]	Yield [%]	Purification factor
Crude extract	253.2	90.1	22806	100.0	1.00
DEAE	79.7	76.4	6090	26.7	0.85
HIC	23.4	180.7	4222	18.5	2.01
Gel filtration	6.2	112.5	697	3.1	1.25

**Table SI-2:** Summary of the purification of the CHI expressed in the original host *Eubacterium ramulus*.

Step	Total protein [mg]	Spec. activity [U mg <sup>-1</sup> ]	Total activity [U]	Yield [%]	Purification factor
Crude extract	552.8	0.02	8.2	100.0	1.0
DEAE	62.8	0.13	8.0	97.0	8.5
HIC	7.8	0.71	5.5	66.4	47.2
Gelfiltration	0.04	35.9	1.3	15.6	2402

## Reference

- [1] C. Herles, A. Braune, M. Blaut, *Arch. Microbiol.* **2004**, *181*, 428-434.

# **Publikation II**

# Structure and catalytic mechanism of the evolutionarily unique bacterial chalcone isomerase

Maren Thomsen<sup>†</sup>, Anne Tuukkanen<sup>‡</sup>, Jonathan Dickerhoff<sup>†</sup>, Gottfried J. Palm<sup>†</sup>, Hanna Kratzat<sup>†</sup>, Dmitri I. Svergun<sup>‡</sup>, Klaus Weisz<sup>†</sup>, Uwe T. Bornscheuer<sup>†</sup> & Winfried Hinrichs<sup>†</sup>

<sup>†</sup>Institute of Biochemistry, University of Greifswald, Felix-Hausdorff-Str. 4, 17489 Greifswald, Germany. <sup>‡</sup>European Molecular Biology Laboratory, Hamburg Outstation, EMBL c/o DESY, Notkestr. 85, 22603 Hamburg, Germany.

**ABSTRACT:** Flavonoids represent a large class of secondary metabolites produced by plants. These polyphenolic compounds are well-known for their anti-oxidative abilities, are antimicrobial phytoalexins, responsible for flower pigmentation to attract pollinators, and, besides other properties, are also specific bacterial regulators governing expression of *Rhizobium* genes involved in root nodulation<sup>1</sup>. The bacterial chalcone isomerase (CHI) from *Eubacterium ramulus* catalyses the first step of a flavanone degradation pathway by ring opening of (2S)-naringenin to form naringenin chalcone. Structural biology and enzymology of plant CHI is well documented<sup>2,3</sup> whereas the existence of bacterial CHIs was elucidated only recently<sup>4,5</sup>. This first determination of the protein structure of a bacterial CHI provides detailed structural insights into the key step of the flavonoid degradation pathway. The active site could be confirmed by co-crystallization with the substrate (2S)-naringenin. The stereochemistry of the proposed mechanism for the isomerase reaction was verified by a specific <sup>1</sup>H/<sup>2</sup>H isotope exchange observed by <sup>1</sup>H NMR experiments and further supported by mutagenesis studies. The active site is shielded by a flexible lid whose varying structure<sup>2</sup> could be modelled in different states of the catalytic cycle using small-angle X-ray scattering data together with the crystallographic structures. Comparison of bacterial CHI with the plant enzyme from *Medicago sativa* reveals unrelated folds, suggesting that the enzyme activity evolved convergent from different ancestor proteins. Despite the lack of any functional relationship, the tertiary structure of the bacterial CHI shows similarities to the ferredoxin-like fold of a chlorite dismutase and the stress-related protein SP1.

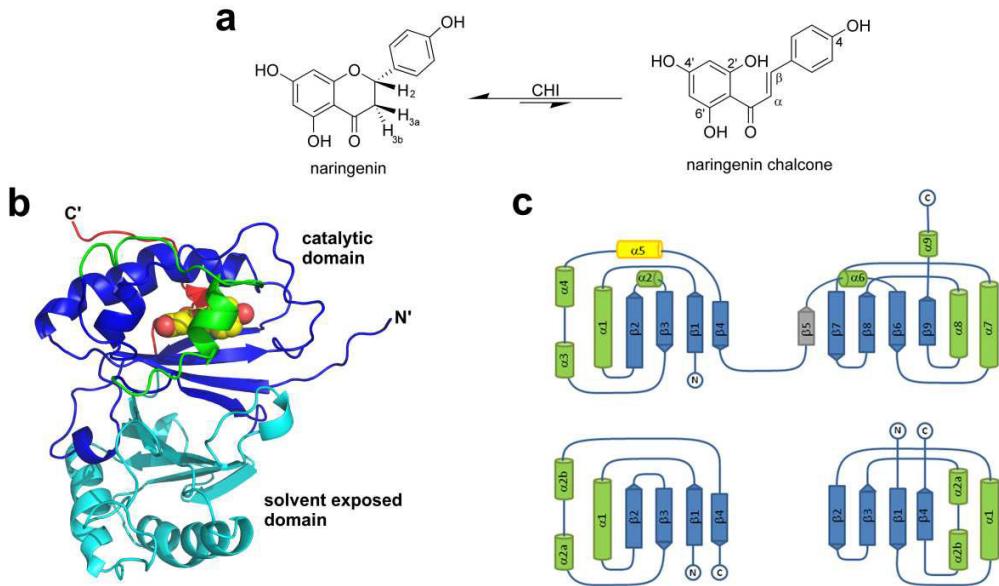
## INTRODUCTION

The first step in the flavonoid biosynthesis is the formation of 4,2',4',6'-tetrahydroxychalcone from the precursors malonyl- and coumaroyl-CoA catalyzed by the chalcone synthase<sup>6</sup>. Afterwards, the plant chalcone isomerase (CHI) catalyzes stereospecific intramolecular cyclisation by an oxa-Michael addition<sup>7</sup> to (2S)-naringenin. This builds the skeleton for a variety of flavonoids and can be further modified to various bioactive compounds. The three-dimensional structure of CHI from plants (*Medicago sativa*)<sup>2</sup>, their evolutionary history<sup>3</sup>, and the reaction mechanism have been investigated in detail<sup>8,9</sup>. A flavonoid degrading gut bacterium was first observed<sup>5</sup> in 1999 and later identified as *Eubacterium ramulus*<sup>10,11</sup>. The enzymes involved in naringenin degradation were purified and characterized as a chalcone isomerase<sup>4</sup>, an enoate reductase (ERED)<sup>12</sup>, and a phloretin hydrolase<sup>13</sup>. Based on these results a degradation pathway for naringenin was postulated (Fig. S1). Whole-genome-sequencing of *Eubacterium ramulus* identified the chalcone isomerase gene (GenBank accession code: KF154734). A PSI-BLAST<sup>14</sup> search of sequence databases revealed no identity compared to plant CHI or other polypeptide sequences (<10%). Thus, neither structural homology nor close evolutionary relationship was expected. In contrast, sequence

data base searches identified homologs of plant CHI in gammaproteobacteria and ascomycetes even though their physiological role remains unclear<sup>15</sup>. The naringenin catabolic pathway is a multistep enzymatic degradation cascade, but in vitro the enzyme catalyses more efficient ring closure of the chalcone, comparable to the flavanone synthesis step in plants (Fig. 1a). Recently, biocatalytical conversion of naringenin into phloretin was shown by using CHI and ERED recombinant expressed in *E. coli*<sup>12</sup>. Thus, the bacterial CHI is of general interest for biotechnological applications in stereospecific flavonoid syntheses and flavanone conversion.

## RESULTS AND DISCUSSION

We carried out X-ray crystal structure analyses of the bacterial CHI from *Eubacterium ramulus* and its (2S)-naringenin complex at 1.8 Å and 2.0 Å resolution, respectively. The bacterial CHI crystallized in the space group  $I_{2,2,2}$  with six monomeric subunits (A to F) in the asymmetric unit constituting a quaternary structure with D<sub>3</sub>

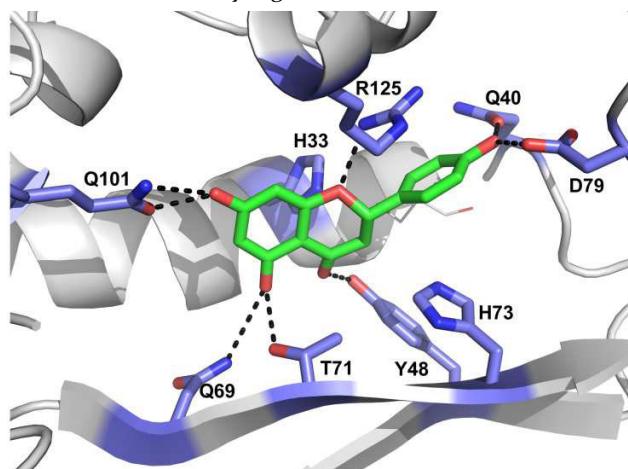


**Figure 1.** Isomerization products, tertiary structure and folding topology of the chalcone isomerase CHI. **a**, Molecular structures of (2S)-naringenin and naringenin chalcone. **b**, Tertiary structure of CHI. The active site with closed lid (green helix and ribbon) is shown with bound (2S)-naringenin (yellow and red spheres for carbon and oxygen atoms, respectively). The catalytic and solvent exposed domains are shown with dark and light blue secondary structure elements, respectively. Both domains have a long  $\alpha$ -helix kinked at a central Pro residue. **c**, Topology diagrams of CHI and the dimer of the stress related protein SP1 (PDB entry iTRo) showing the common fold motive with  $\beta$ -strands of the ferredoxin-fold sheets in blue,  $\alpha$ -helices in green, and the  $\alpha$ -helix of the lid structure in yellow.

symmetry (Fig. S2), which we verified by small-angle X-ray scattering in solution (SAXS) (Table S6). Each polypeptide chain has 282 amino-acid residues with a molecular weight of 32.4 kDa. The tertiary structure belongs to the superfamily of dimeric  $\alpha+\beta$  barrels<sup>16</sup> and is divided into a catalytic domain (residues 1 - 143) and a solvent exposed domain (residues 144 - 266) (Fig. 1b). These very similar domains have a bacterial ferredoxin-like fold<sup>17</sup>, both with anti-parallel  $\beta$ -sheets with 4 and 5  $\beta$ -strands, respectively, each with two  $\alpha$ -helices on top. The domains form a  $\beta$ -sheet sandwich with the  $\beta$ -sheets rotated against each other by about 90°. The catalytic domain is strongly involved in dimer as well as trimer formation, whereas the solvent exposed domain shows C-terminal interactions (residues 264 - 282) for dimer formation. The C-terminal residues 279 - 282 penetrate into the core of a neighbored monomer contributing to dimer formation by a network of hydrogen bonds (Fig. S3). Detailed crystal packing analyses of the protein interfaces with PISA<sup>8</sup> revealed that for each monomer the buried surface area upon dimer formation (2100 Å<sup>2</sup>) is much higher than upon trimer formation (910 Å<sup>2</sup>). Truncation of the last 5 residues lead to a variant (CHI<sub>ΔC-term</sub>) that has reduced thermal stability and almost no enzymatic activity ( $\Delta T_m = 9^\circ\text{C}$ , Table S5 and Fig. S9c). To investigate the evolutionary relation of CHIs from bacteria and plants with respect to tertiary structure, superposition of the secondary structure elements was performed using PDBeFold<sup>19</sup>. Both structures are based on an anti-parallel  $\beta$ -sheet with two  $\alpha$ -helices on top. At a first glance, the solvent exposed do-

main of the bacterial CHI and the plant enzyme from *Medicago sativa* (PDB entry iEYQ)<sup>3</sup> can be superimposed. However,  $\beta$ -sheet- $\alpha$ -helix arrangements are very common structural motifs in protein folding, but the topology diagrams show rather different connections between the individual  $\beta$ -strands (Fig. 1c and Fig. S4a-d). We conclude that it is unlikely that the enzymes have a close evolutionary link. This is consistent with the observation that the postulated ancestor protein of the plant CHI<sup>3</sup>, the fatty-acid-binding protein of *Arabidopsis thaliana* (PDB entry 4DOI), cannot be superimposed either. Structural homologs were identified by a database search based on secondary structure elements with the DALI-server<sup>20</sup> (Table S4). The best hit turned out to be a chlorite dismutase<sup>21</sup> (PDB entry 3Qo8) which is a haem-dependent enzyme responsible for detoxification of ClO<sub>2</sub><sup>-</sup>. These obviously functionally not related enzymes exhibit similar folds with two bacterial ferredoxin-like domains. The catalytic domains are superimposed with an overall root mean square deviation (r.m.s.d.) of 3.2 Å on  $\text{C}\alpha$ -atoms. However, the other domains are contorted to each other although the topology diagram shows identical  $\beta$ -strand connections (Fig. 1c). As mentioned above, a twofold related CHI monomer penetrates its C-terminus into the interface of the  $\beta$ -sheets and causes this contortion. It is tempting to conclude that in evolution two single ferredoxin-like domains were linked up. Further database searches with only one domain identified such a possible ancestor protein. The stress-related, thermo-stable protein SP1 from *Populus tremula* (PDB entry iTRo) consists

of 108 amino-acid residues and has a ferredoxin-like fold<sup>22</sup> (Fig. 1c). The superposition of the secondary structure elements with the catalytic domain of the bacterial CHI reveals a fit with an overall r.m.s.d. of 2.9 Å on C $\alpha$  atoms. The  $\beta$ -sheets of two SP1 molecules generate strong interaction for dimerization<sup>22</sup>. To identify the active site, co-crystallization with the naringenin chalcone was performed and resulted in the flavanone complex of CHI due to the equilibrium of this reaction. The hexameric quaternary structure shows three different dimers (A/B, C/D, and E/F) with respect to the active site. In each subunit of one trimer (chains A, C, E) (2S)-naringenin could be found in the active site shielded by a lid structure (amino-acid residues 109 – 130), which is part of the catalytic domain. In all subunits of the other trimeric half of the quaternary structure (chains B, D, F) the lid structures are not in the closed conformation and not observed in the electron density maps. The active sites with open entry tunnels are occupied by naringenin and naringenin chalcone due to artificially high concen-



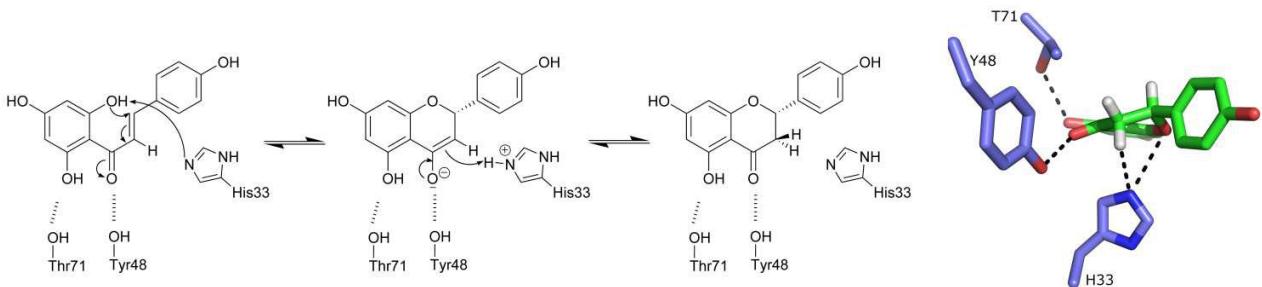
**Figure 2.** Active site and substrate recognition of CHI. (2S)-naringenin binding in the active site to amino-acid residues labelled with one letter code. Hydrogen bonds are shown as stippled lines.

tration used for co-crystallization. The electron density maps show a complicated disorder of substrate and product molecules in the active site and the entry tunnel. This is interpreted as superposition of snapshots of substrate binding and product release steps (Fig. S5a – f). The substrate binds in a sensible way between the  $\beta$ -sheet and  $\alpha$ -helices comparable to other examples of active sites in this superfamily, e.g. the chlorite dismutase<sup>21</sup>. The side chains of residues Q40, Y48, Q69, T71, D79, Q101, and R125 are responsible for substrate binding by hydrogen bonding within distances of 2.5 – 3.2 Å (Fig. 2). In the substrate-free enzyme T71 and Q101 are hydrogen bonded to one glycerol molecule of the cryo-protectant replacing the phenolic 4',6'-dihydroxy fragment of the chalcone (Fig. S6). The lid forms a short  $\beta$ -strand (residues 112–116) and  $\alpha$ -helix ( $\alpha$ 5, residues 120–126) and positions the chalcone in an appropriate conformation for flavanone formation. The side chain of R125, C-terminal of  $\alpha$ 5, contributes to lid closing by a salt bridge to E91. Simultaneously, at the entrance of the substrate binding cleft the salt

bridge of E131 (at the lid hinge) and K253, observed in the open conformation, is lost upon substrate binding, increasing the distance of the corresponding ammonium and carboxylate to about 12 Å. The guanidinium moiety of R125 positions the phenyl-ring of the substrate by  $\pi$ -stacking (3.1 – 3.4 Å) and with a hydrogen bond to the ring O1 of the flavanone (2.9 Å) (Fig. 2). The isomerization mechanism of plant CHI relies on a catalytic water molecule and the bound substrate is solvent accessible<sup>2,8,9</sup>. In contrast, in the active site of bacterial CHI no water molecule is observed around naringenin. Reversible naringenin to naringenin chalcone isomerization relies on a general acid-base mechanism initiated by the key residue H33 as a proton shuttle. H33N $\epsilon$ 2 is within 3.2 – 3.5 Å equally distant to both C<sub>3</sub> and O<sub>1</sub> of naringenin (Fig. 3). Thus, the flavanone-ring opening is initiated by deprotonating the methylene group of C<sub>3</sub> by the imidazole of H33. The enolate intermediate is stabilized by hydroxyl groups of Y48 and T71, which act as an “oxyanion hole”. Subsequent protonation of O<sub>1</sub> promotes opening of the central flavanone ring and forms the chalcone. Consequently, the Michael reaction<sup>7</sup> of ring closure will be initiated by deprotonating the phenolic O<sub>2'</sub> of the chalcone and finished by protonation of C<sub>3</sub> (Fig. 3). This proposed reaction mechanism was tested by an enzyme assay<sup>12</sup> performed in <sup>2</sup>H<sub>2</sub>O and subsequent product analysis by <sup>1</sup>H NMR spectroscopy. Naringenin chalcone was used as substrate to form (2S)-naringenin (Methods and Fig. S7). The experiments unambiguously show that the oxo-Michael addition<sup>7</sup> involves protonation of the axial pro-S position at C<sub>3</sub> (Fig. 3). This is consistent with the mutagenesis studies and the structure with bound (2S)-naringenin, as the axial H-atom at C<sub>3</sub> directly points to H33N $\epsilon$  (Fig. 3). Possible catalytic residues His33, Y48, H73 or R125 were investigated by mutagenesis studies. The enzymatic efficiency ( $k_{cat}/K_M$ ) of the wild-type bacterial CHI is unchanged for the Y48A and H73A variants (Table 1). The efficiency of H33A and H33Q drops by about 3 and 4 orders of magnitude, respectively. Probably, the H33A variant has residual activity caused by an imidazole-

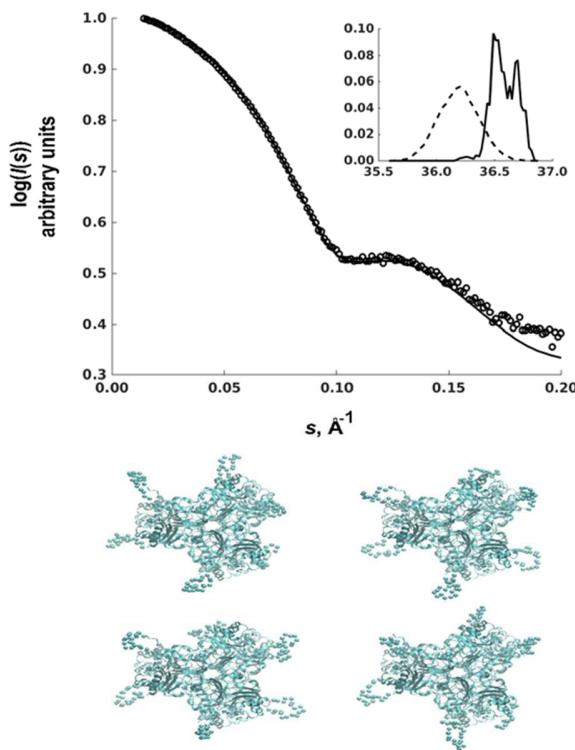
**Table 1. Kinetic parameters of CHI and its variants**

Enzyme	K <sub>M</sub>	k <sub>cat</sub>	k <sub>cat</sub> /K <sub>M</sub>
	[μM]	[s <sup>-1</sup> ]	[M <sup>-1</sup> × s <sup>-1</sup> ]
CHI	35.4 ± 3.5	1142 ± 29	3.2 × 10 <sup>7</sup>
CHI <sub>Δ</sub> lid	199.5 ± 23.3	1177 ± 63	5.9 × 10 <sup>6</sup>
CHI_R125A	56.7 ± 7.7	1379 ± 64	2.4 × 10 <sup>7</sup>
CHI_H73A	43.9 ± 3.5	1313 ± 28	2.8 × 10 <sup>7</sup>
CHI_Y48F	14.3 ± 1.1	784 ± 13	5.5 × 10 <sup>7</sup>
CHI_H33A	20.3 ± 1.6	0.6 ± 0.01	2.9 × 10 <sup>4</sup>
CHI_H33E	131.9 ± 16.8	0.2 ± 0.013	1.6 × 10 <sup>3</sup>
CHI_H33Q	79.8 ± 8.4	0.07 ± 0.003	0.9 × 10 <sup>3</sup>
plant CHI <sup>9</sup>	112 ± 28	186.3 ± 23	1.66 × 10 <sup>6</sup>
uncatalyzed <sup>9</sup>		0.78 × 10 <sup>-4</sup>	



**Figure 3.** Mechanism of reaction of the reversible Michael addition catalyzed by H33. The ring opening reaction forming the naringenin chalcone is initiated by imidazole H33 abstracting the proton H3b. The reversible reaction relies on the hydrogen bonding (stippled lines) pattern of H33 with C3 and O1. Hydrogen bonding of other active site residues contributes to polarization of the substrates.

replacing water molecule capable of proton transfer. This is in agreement with the almost inactive H33Q and H33E variants. The side-chains of Q33 and E33 have approximately the same sterical requirements as H33 and may interact in a hydrogen-bonding network, but Q33 is unable to mediate an acid-base mechanism. This is in line with the slightly increased efficiency of the H33E variant.



**Figure 4.** a, Ensemble optimization method (EOM) fit of the ligand-free CHI. The SAXS data of CHI (circles) are optimally fitted (solid line) by an ensemble of hexamers with some monomers in open and some in closed lid conformations. The intensity is shown as a function of momentum transfer ( $s = 4\pi \sin\theta/\lambda$ , where  $\lambda$  is the X-ray wavelength and  $2\theta$  is the scattering angle). Inset:  $R_g$  distribution of the random pool (dashed line) and of the selected ensemble (solid line). b, Examples of CHI conformations with the  $C\alpha$  positions of the lid residues represented as spheres using PyMOL graphics<sup>36</sup>.

Significant differences between the substrate-free CHI, its naringenin complex, and the lid deletion variant CHI\_Alid were revealed by SAXS data (Table S6 and Fig. 4, S10 – S13). The measured scattering profile of the substrate-free CHI hexamer could be optimally fitted with an ensemble of hexamer structures in which some of the monomers had partially or fully closed lids (the goodness-of-the-fit  $\chi = 1.4$ , Fig. 4). Two main structural populations with average  $R_g$  values of 36.5 and 36.7 Å with a continuous distribution between them were found. This is an indication of an equilibrium between the open and closed states and of a low energy barrier for the mechanism of lid closing. The finding is further supported by the closed lid conformation observed in one subunit of the substrate-free hexamer observed in a dataset at 2.8 Å resolution with similar crystallization buffer conditions (PDB entry 3ZPH). The overall structure of the substrate-bound CHI is more compact than that of the native protein which is observed as differences in the model-free parameters  $R_g$  and  $D_{max}$  derived from the SAXS data. There is also a change of the minimum of the SAXS profile of the naringenin bound complex about 0.1 Å<sup>-1</sup> indicating further a more globular structure compared to the substrate-free CHI (Fig. S1). A model of the naringenin bound hexamer constructed using the corresponding X-ray crystallographic structure of a trimer with all lids closed (PDB entry 4Do6) and a symmetry operation yields an excellent fit against SAXS data of the substrate-bound CHI. This suggests that the majority of the lids are in the closed state in the substrate-bound enzyme. The scattering profile of the lid deletion variant CHI\_Alid is consistent with the hexameric X-ray crystallographic structure of CHI which is missing electron density of lids (PDB entry 4C9S). Superposition of all monomers of CHI structures with and without bound substrate looks like an enzyme at work. The tertiary structure has a remarkable rigid polypeptide scaffold with the exception of large conformational changes of the lid structure. Only the side-chains of the catalytically active residues move for substrate binding and product release in the preformed active site (Fig. 3). Major advantage of lid closure is the 19-fold enhanced enzymatic efficiency in naringenin formation compared to plant CHI<sup>8</sup>. Deletion of the lid structure cause a solvent exposed active site and resulted in an active phenotype with about 5-fold decreased efficiency,

though still 4-fold higher than for plant CHI. The specific hydrogen bonding and  $\pi$ -stacking of the substrate with R125 is not mandatory for efficient catalysis, because the R125A variant shows almost unchanged catalytic efficiency (Table 1). Thus, the lid structure enhances efficient catalysis by reducing the dielectric constant of the active site, but additional sequence specific interactions are ruled out. The lack of specific strong interactions between the lid structure and the protein core is in agreement with the SAXS findings that the substrate-free CHI data can be best modelled with an ensemble of structures having an equilibrium state of open and closed lids. Shielding the active site against the solvent to prevent side reactions is obviously not the case with respect to the solvent exposed active site of plant CHI<sup>2</sup>. The lid closing provokes some allosteric changes of side-chains causing an in plane shift of the imidazole center of H33 to the substrate by about 2 Å (Fig. 3). These conformational changes derive only from lid closing, but are not induced by substrate binding, because the same events are observed in the closed conformation without ligand as well.

## CONCLUSION

The tertiary structure of bacterial CHI with its ferredoxin-like fold is completely different to the plant CHI. The bacterial CHI is only related to plant CHI with respect to the products of the catalyzed oxa-Michael-addition. The general acid-base mechanism of both enzymes relies in bacterial CHI on the H33 side-chain in contrast to the catalytic water molecule of plant CHI. The substrate-binding site of plant CHI is solvent exposed during the catalytic process, but the bacterial CHI has a lid structure closing the active site upon substrate binding. The shielding lid is unambiguously the main reason for the 19-fold higher efficiency of the bacterial CHI.

## EXPERIMENTAL SECTION

**Protein expression and purification.** The CHI gene from *Eubacterium ramulus* was cloned, expressed and purified as described previously<sup>12</sup>. The production of selenomethionine substituted protein was performed in *E. coli* Rosetta (DE3) by metabolic inhibition<sup>23</sup>. The purification protocol was the same as that for the native protein.

**Crystallization and data collection.** Crystals were obtained using the hanging-drop method using an equal volume of protein to reservoir solution followed by one round of streak seeding with cat whiskers<sup>24</sup>. The protein concentration used was 7.5 mg ml<sup>-1</sup> and determined spectrophotometrically at 280 nm. First little crystals were obtained with a reservoir solution containing 2.0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.2 M NaCl and 0.1 M HEPES pH 7.5. To improve the quality of the crystals they were seeded with a cat whisker to drops with a reservoir of 1.7 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.2 M NaCl and 0.1 M HEPES pH 7.5. (2S)-naringenin bound crystals were obtained using 2.0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.2 M NaCl, 0.1 M HEPES pH 7.5, 6% (v/v) ethanol and 370 µM of naringenin chalcone as crystallization conditions. Crystallization was performed at 293 K. X-ray dif-

fraction data were collected at 100 K on beamline 14.1 at the BESSY II Synchrotron (Berlin, Germany). For all crystals a cryo-protectant solution containing 22% (v/v) glycerol, 2.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.2 M NaCl and 0.1 M HEPES pH 7.5 was used. The crystals of the native CHI and the SeMet-derivative were isomorphous, and belonged to the orthorhombic space group *I*<sub>2</sub>,*2*,*2*, or *I*<sub>2</sub>*2*<sub>2</sub>. All diffraction images were processed with XDS<sup>25</sup> using the graphical user interface XDSapp<sup>26</sup>. In-house data of a substrate-free crystal were collected with cryo-protectant at 100 K with a Saturn 92 CCD detector mounted on a Micromax 007 rotating anode source (Rigaku MSC). These data were processed with MOSFLM<sup>27</sup> and scaled with SCALA<sup>28</sup>. The structure factor amplitudes were calculated using TRUNCATE<sup>29</sup>. Data collection and processing statistics are summarized in Table S1.

**Structure determination and refinement.** The phases for the selenomethionine data were calculated for both possible space groups with SHELXC/D/E<sup>30</sup> using the GUI HKL2map<sup>31</sup>. Only the space group *I*<sub>2</sub>,*2*,*2*, resulted in interpretable electron density maps. We identified 54 selenium positions (Fig. S8) for nine selenium-methionine residues found in the sequence of bacterial CHI assuming posttranslational cleavage of the N-terminal methionine. The model was built manually into the electron density using COOT<sup>32</sup>. Six monomers were assigned to the asymmetric unit and refined with REFMAC5, including TLS segments<sup>33</sup>. The substrate complex and the in-house measured data set could only be solved by molecular replacement using one monomer of the initially solved structure as model, due to the differences in the cell parameters. Molecular replacement was carried out with the program PHASER<sup>34</sup> and subsequent refinement with REFMAC5 as described above. The quality of the refined protein models was validated using MOLPROBITY<sup>35</sup>. Molecular graphics were prepared with PYMOL<sup>36</sup>. Refinement statistics are listed in Table S2.

**Small angle X-ray scattering.** The synchrotron radiation X-ray scattering data of CHI with and without of naringenin chalcone (substrate concentration was varied between 2 µM and 100 mM) as well as the CHI\_Alid construct were collected on the P12 beamline of the EMBL Hamburg on the storage ring PETRA III (DESY, Hamburg, Germany) using a PILATUS 2M pixel detector, a sample-detector distance of 3.1 m, and a wavelength of  $\lambda = 1.24$  Å covering the range of momentum transfer of  $0.008 \text{ \AA}^{-1} < s < 0.47 \text{ \AA}^{-1}$ . For each construct, several solute concentrations in the range of 0.8 to 8.3 mg ml<sup>-1</sup> in 50 mM sodium phosphate buffer pH 6.8 were measured. To monitor for the radiation damage, 20 successive 50 ms exposures of protein solutions were compared and no significant changes were observed. The data were normalized to the intensity of the transmitted beam and radially averaged. The scattering of the buffer without protein solute was subtracted and the difference curves were scaled for protein concentration. The low angle data measured at lower protein concentrations were extrapolated to infinite dilution and merged with the higher concentration data to yield the final composite scattering curves. The data pro-

cessing steps were performed using the program PRIMUS<sup>37</sup> of the atsas software suite<sup>38</sup>.

Radius of gyration  $R_g$  and forward scattering intensity  $I(o)$ , were independently determined using Guinier analysis<sup>39</sup> and the indirect Fourier transformation approach of the program GNOM<sup>40</sup>. Additionally, the maximum particle dimension  $D_{max}$  was obtained from the latter approach. Molecular masses of protein constructs ( $MM_{SAXS}$ ) were calculated by comparing the extrapolated forward scattering intensities with that of a reference sample of bovine serum albumin (BSA), ( $MM_{ref} = 66$  kDa) together with concentration information. The excluded volume of the hydrated protein  $V_p$  was obtained with DATPOROD<sup>38</sup> and used to extract an independent estimate of the molecular mass ( $MM_{POROD}$ ). For globular proteins, hydrated protein volumes in  $\text{\AA}^3$  are about 1.7 times the molecular masses in Da. The statistics of data collection are summarized in Table S5.

The scattering patterns from our high-resolution structural models were calculated using the program CRYSTAL<sup>41</sup> and used to determine the fit of these models to the experimental scattering data. Given the atomic coordinates of a structural model, the program minimizes the discrepancy  $\chi^2$  between the experimental and theoretical scattering intensities by adjusting the excluded volume of the particle and the contrast of the hydration layer.

**SAXS modeling.** Flexibility analysis of the lid regions of CHI in solution was conducted using the available crystallographic models as input for the ensemble optimization method (EOM). This approach seeks to best fit the experimental scattering profile with an ensemble of conformations<sup>42</sup>. The protein was modeled as a hexamer and the lid regions of each subunit (residues 107 – 130) were defined as flexible and their possible conformations modelled with RANCH<sup>42</sup> producing 10,000 random configurations, while the rest of the hexameric protein was kept fixed. A genetic algorithm was employed to find the set of conformations best fitting the SAXS data. The structures selected from the random pool of structures were analyzed with respect to the  $R_g$  distribution. SAXS data and interpretation are summarized in Figures S10 – S13.

**Mutant generation.** The mutations for determining the catalytic residues were introduced by PCR with one primer carrying the desired mutation and a primer binding at the T7 terminator sequence of the pET-vector (named pET\_RP). All primers used are listed in the Table S3. After initial denaturation at 95°C for 5 min, the cycling program was followed for 30 cycles: 45 s, 95°C denaturation, 45 s, respective melting temperature for primer annealing, 90 s, and 72°C elongation. A final elongation step was performed over 5 minutes at 72°C.

With the resulting PCR-product a MegaWhoP-PCR<sup>43</sup> was performed, after the following procedure: for 5 min at 68°C, initial denaturation for 1 min at 95°C, the first cycling program was followed for 10 cycles: 30 s, 95°C denaturation, 45 s, 55°C for primer annealing, 7 min, 62°C elongation. Afterwards a second cycling program followed with 14 cycles: 30 s, 95°C denaturation, 45 s, 55°C for pri-

mer annealing, and 11 min, 68°C elongation. The non-methylated DNA was digested by *Dpn*I and transformed into *E. coli* BL21 (DE3).

The deletion of the lid (amino-acid residues 109 – 130) was introduced using the FastCloning method. After initial denaturation for 5 min at 95°C, the cycling program was followed for 25 cycles: 30 s, 95°C denaturation, 30 s, 50°C melting temperature for primer annealing, and 6.5 min, 72°C elongation. The non-methylated DNA was digested by *Dpn*I and transformed into *E. coli* BL21 (DE3).

**Enzyme assay.** The chalcone isomerase activity was measured spectrophotometrically as result of the conversion from naringenin chalcone to naringenin at 368 nm ( $\epsilon = 29,068 \text{ mM}^{-1} \text{ cm}^{-1}$ ) (Fig. S9). The reaction mixture (0.5 ml, 2 mm cuvette) contained 50 mM sodium phosphate-buffer pH 6.4, 1 mg BSA, varying naringenin chalcone concentration (dissolved in DMSO stock solution), and an appropriate amount of enzyme. The activity is not affected by DMSO concentrations up to 5%.

**Melting points (Tm)** were determined by circular dichroism spectroscopy (CD). The purified enzymes were subjected to CD measured at 222 nm using a Jasco V-650 with a heat rate of 0.5°C min<sup>-1</sup> in the range of 20 – 80°C (Table S5).

**<sup>1</sup>H NMR experiments.** For mechanistic studies, the enzyme was pre-incubated in <sup>2</sup>H<sub>2</sub>O for 1 h. The enzyme assay mentioned above was performed in the deuterated solvent and the product extracted twice with ethyl acetate. The organic solvent was evaporated with a SpeedVac and the residual product measured in DMSO-d<sub>6</sub>. <sup>1</sup>H NMR spectra (Fig. S7) were collected with a Bruker Avance 600 MHz spectrometer equipped with an inverse <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P quadrupole resonance cryoprobehead and z-field gradients. The data were processed and analyzed using the Topspin 3.1 software (Bruker).

## ASSOCIATED CONTENT

### Supporting Information

X-ray crystallographic data, SAXS data, NMR spectra, and supportive figures and tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

### Data deposition

Atomic coordinates and structure factors have been deposited at the Protein Data Bank for CHI, SeMet-CHI, CHI (low resolution), and the (2S)-naringenin complex under accessions codes 4c9s, 4c9t, 3zph, and 4d06, respectively. SAXS data and models of CHI, (2S)-naringenin complex, and CHI<sub>Δ</sub>lid have been deposited at the small-angle scattering biological database under accession codes SASDAL6, SASDAM6, SASDAN6).

## AUTHOR INFORMATION

### Corresponding Author

winfried.hinrichs@uni-greifswald.de

### Author Contributions

M.T. cloned, expressed, purified and crystallized the proteins, and performed kinetics assays. H.K. prepared the CHI $\Delta$ C-term variant and determined T<sub>m</sub>. M.T. and G.J.P. collected and processed X-ray diffraction data, performed crystallographic phasing calculations and refined the models. M.T., G.J.P. and W.H. analyzed the protein structures. A.T. and D.S. contributed the SAXS experiments and data analyses. J.D. and K.W. carried out and analyzed the NMR experiments. U.T.B. and W.H. initiated the joint project, organized and designed the scope of the study. All authors were involved in discussing data and preparing the manuscript.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENT

Diffraction data for this study were collected at beamline 14.1 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron-storage ring (Berlin-Adlershof, Germany). The SAXS data was collected at beamline P12 operated by EMBL Hamburg at the PETRA-III ring (DESY, Hamburg, Germany). M. Backes (Symrise AG, Holzminden, Germany) synthesized the naringenin chalcone. This work was supported by the "Bundesministerium für Bildung und Forschung" within the "Biokatalyse 2021" cluster (FKZ: 0315365 and 031A109). M.T. thanks the "Landesgraduierten-Kolleg of Mecklenburg-Vorpommern" for a Ph.D. grant. A.T. was supported by the EMBL Interdisciplinary Postdoc Programme (EIPOD) under Marie Curie COFUND actions and by the BMBF research grant BioSCAT, contract No 05K12YE1.

## ABBREVIATIONS

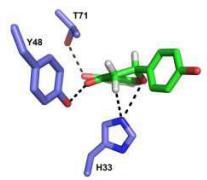
BSA bovine serum albumin, CHI chalcone isomerase, CHS chalcone synthase, ERED enoate reductase.

## REFERENCES

- (1) Firmin, J. L.; Wilson, K. E.; Rossen, L.; Johnston, A. W. B. *Nature* **1986**, *324*, 90.
- (2) Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nat. Struct. Mol. Biol.* **2000**, *7*, 786.
- (3) Ngaki, M. N.; Louie, G. V.; Philippe, R. N.; Manning, G.; Pojer, F.; Bowman, M. E.; Li, L.; Larsen, E.; Wurtele, E. S.; Noel, J. P. *Nature* **2012**, *485*, 530.
- (4) Herles, C.; Braune, A.; Blaut, M. *Arch. Microbiol.* **2004**, *181*, 428.
- (5) Schneider, H.; Schwartz, A.; Collins, M. D.; Blaut, M. *Arch. Microbiol.* **1999**, *171*, 81.
- (6) Winkel-Shirley, B. *Plant Physiol.* **2001**, *126*, 485.
- (7) Nising, C. F.; Bräse, S. *Chem. Soc. Rev.* **2012**, *41*, 988.
- (8) Jez, J. M.; Bowman, M. E.; Noel, J. P. *Biochemistry* **2002**, *41*, 5168.
- (9) Jez, J. M.; Noel, J. P. *J. Biol. Chem.* **2002**, *277*, 1361.
- (10) Braune, A.; Gütschow, M.; Engst, W.; Blaut, M. *Appl. Environ. Microbiol.* **2001**, *67*, 5558.
- (11) Schneider, H.; Blaut, M. *Arch. Microbiol.* **2000**, *173*, 71.
- (12) Gall, M.; Thomsen, M.; Peters, C.; Pavlidis, I. V.; Jonczyk, P.; Grünert, P. P.; Beutel, S.; Scheper, T.; Gross, E.; Backes, M.; Geißler, T.; Ley, J. P.; Hilmer, J.-M.; Krammer, G.; Palm, G. J.; Hinrichs, W.; Bornscheuer, U. T. *Angew. Chem. Int. Ed.* **2014**, *53*, 1439.
- (13) Schoefer, L.; Braune, A.; Blaut, M. *Appl. Environ. Microbiol.* **2004**, *70*, 6131.
- (14) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389.
- (15) Gensheimer, M.; Mushegian, A. *Protein Sci.* **2004**, *13*, 540.
- (16) Wilson, D.; Pethica, R.; Zhou, Y. D.; Talbot, C.; Vogel, C.; Madera, M.; Chothia, C.; Gough, J. *Nucleic Acids Res.* **2009**, *37*, D380.
- (17) Abendroth, J.; Rice, A. E.; McLuskey, K.; Bagdasarian, M.; Hol, W. G. *J. J. Mol. Biol.* **2004**, *338*, 585.
- (18) Krissinel, E.; Henrick, K. *J. Mol. Biol.* **2007**, *372*, 774.
- (19) Krissinel, E.; Henrick, K. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2004**, *60*, 2256.
- (20) Dietmann, S.; Park, J.; Notre Dame, C.; Heger, A.; Lappe, M.; Holm, L. *Nucleic Acids Res.* **2001**, *29*, 55.
- (21) de Geus, D. C.; Thomassen, E. A. J.; Hagedoorn, P.-L.; Pannu, N. S.; van Duijn, E.; Abrahams, J. P. *J. Mol. Biol.* **2009**, *387*, 192.
- (22) Dgany, O.; Gonzalez, A.; Sofer, O.; Wang, W.; Zolotnitsky, G.; Wolf, A.; Shoham, Y.; Altman, A.; Wolf, S. G.; Shoseyov, O.; Almog, O. *J. Biol. Chem.* **2004**, *279*, 51516.
- (23) Doublie, S. *Methods Enzymol.* **1997**, *276*, 523.
- (24) Bergfors, T. *J. Struct. Biol.* **2003**, *142*, 66.
- (25) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125.
- (26) Krug, M.; Weiss, M. S.; Heinemann, U.; Mueller, U. *J. Appl. Crystallogr.* **2012**, *45*, 568.
- (27) Leslie, A. *Daresbury, Warrington, WA 4AD, UK* **1991**.
- (28) Evans, P. R. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 72.
- (29) French, S.; Wilson, K. *Acta Crystallogr., Sect. A* **1978**, *34*, 517.
- (30) Sheldrick, G. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 479.
- (31) Pape, T.; Schneider, T. R. *J. Appl. Crystallogr.* **2004**, *37*, 843.
- (32) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126.
- (33) Collaborative Computational Project Number 4 *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760.
- (34) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.
- (35) Chen, V. B.; Arendall, W. B., III; 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: Biol. Crystallogr.* **2010**, *66*, 12.
- (36) DeLano, W. L. *Abstr. Pap. Am. Chem. Soc.* **2009**, 238.
- (37) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *J. Appl. Crystallogr.* **2003**, *36*, 1277.
- (38) Petoukhov, M. V.; Franke, D.; Shkumatov, A. V.; Tria, G.; Kikhney, A. G.; Gajda, M.; Gorba, C.; Mertens, H. D. T.; Konarev, P. V.; Svergun, D. I. *J. Appl. Crystallogr.* **2012**, *45*, 342.
- (39) Guinier, A. *Ann. Phys. Paris* **1939**, *12*, 161.
- (40) Svergun, D. I. *J. Appl. Crystallogr.* **1992**, *25*, 495.
- (41) Svergun, D.; Barberato, C.; Koch, M. H. *J. Appl. Crystallogr.* **1995**, *28*, 768.
- (42) Bernado, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007**, *129*, 5656.
- (43) Miyazaki, K. In *Synthetic Biology, Pt B: Computer Aided Design and DNA Assembly*; Voigt, C., Ed.; Elsevier Academic Press Inc: San Diego, 2011; Vol. 498, p 399.

Insert Table of Contents artwork here

---



# **Structure and catalytic mechanism of the evolutionarily unique bacterial chalcone isomerase**

Maren Thomsen<sup>†</sup>, Anne Tuukkanen<sup>‡</sup>, Jonathan Dickerhoff<sup>†</sup>, Gottfried J. Palm<sup>†</sup>, Hanna Kratzat<sup>†</sup>, Dmitri I. Svergun<sup>‡</sup>, Klaus Weisz<sup>†</sup>, Uwe T. Bornscheuer<sup>†</sup> & Winfried Hinrichs<sup>†</sup>

<sup>†</sup>Institute of Biochemistry, University of Greifswald, Felix-Hausdorff-Str. 4, 17489 Greifswald, Germany.

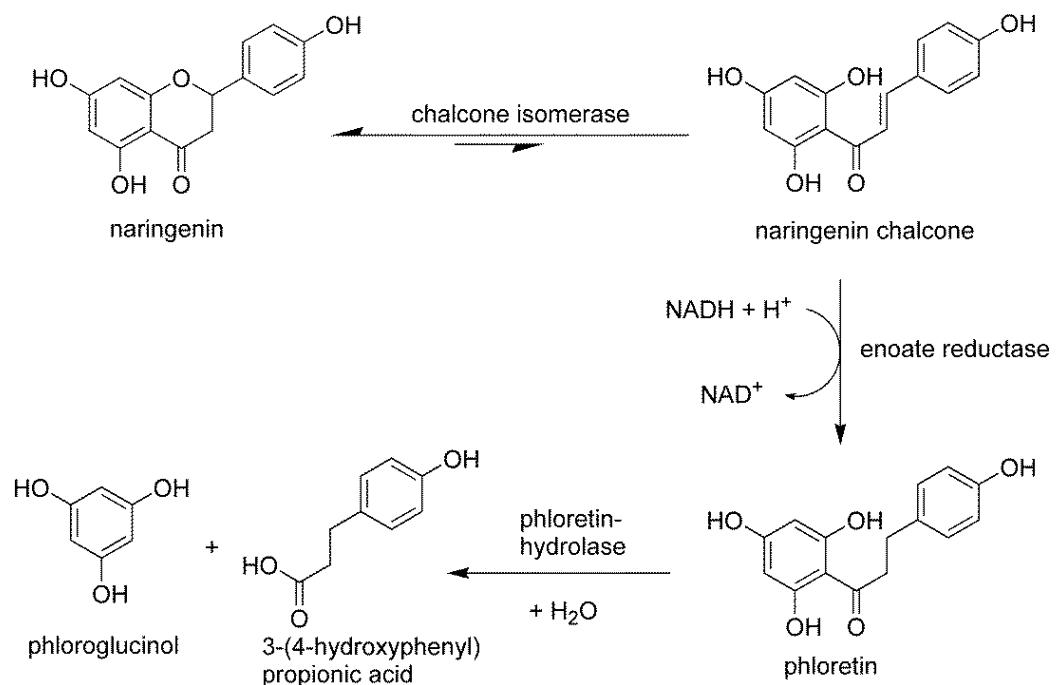
<sup>‡</sup>European Molecular Biology Laboratory, Hamburg Outstation, EMBL c/o DESY, Notkestr. 85, 22603 Hamburg, Germany.

---

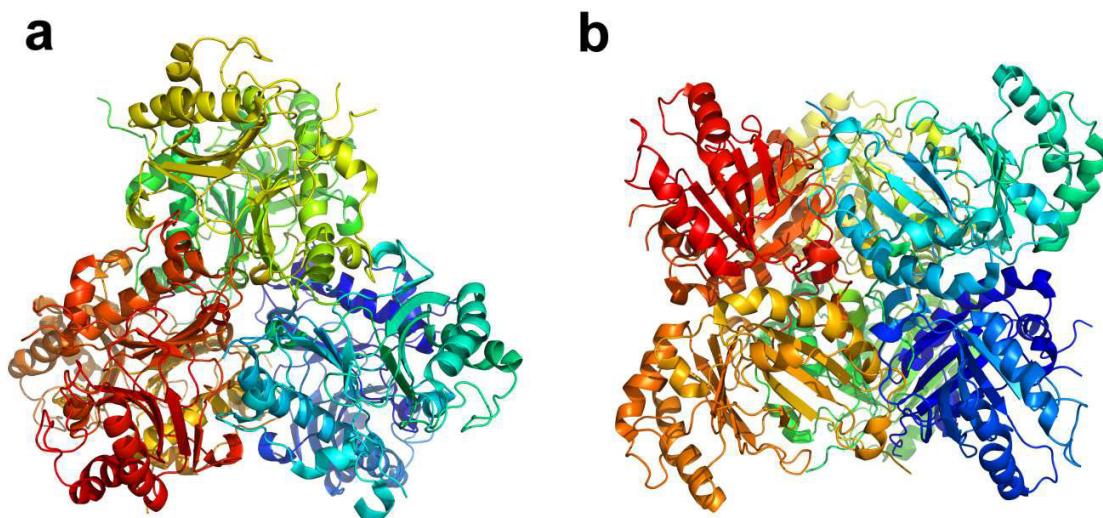
## **Supporting Information**

### **Table of Contents**

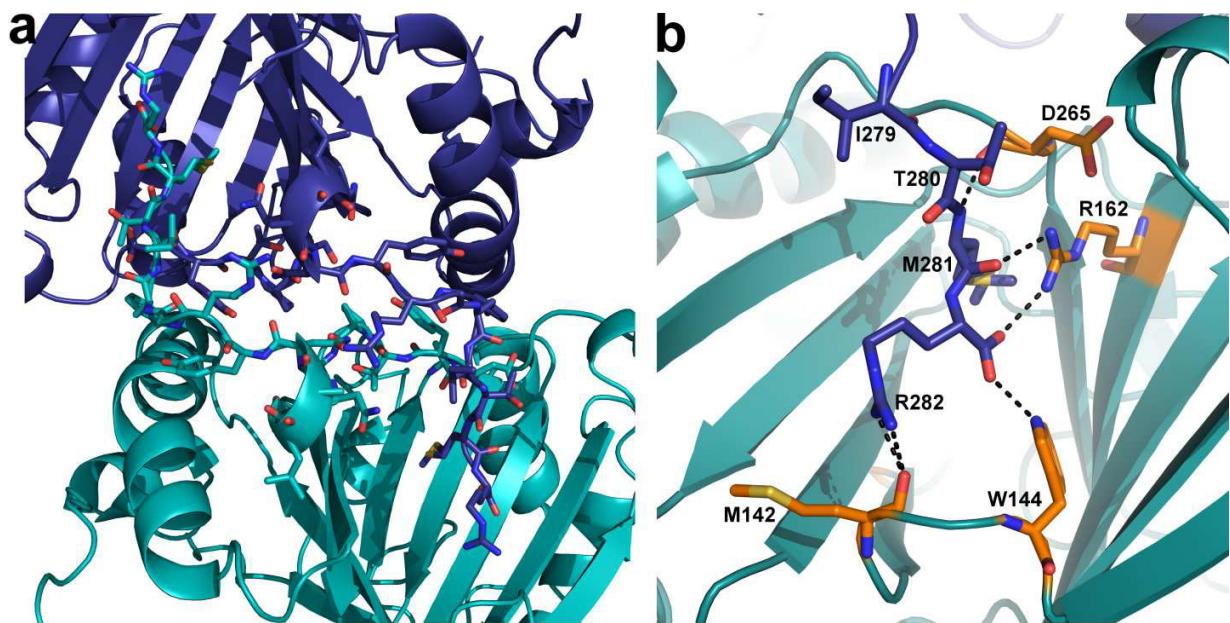
<b>Figure S1.....</b>	<b>S2</b>
<b>Figure S2.....</b>	<b>S2</b>
<b>Figure S3.....</b>	<b>S3</b>
<b>Figure S4.....</b>	<b>S4</b>
<b>Figure S5.....</b>	<b>S5</b>
<b>Figure S6.....</b>	<b>S8</b>
<b>Figure S7.....</b>	<b>S9</b>
<b>Figure S8.....</b>	<b>S11</b>
<b>Figure S9.....</b>	<b>S12</b>
<b>Table S1.....</b>	<b>S14</b>
<b>Table S2.....</b>	<b>S15</b>
<b>Table S3.....</b>	<b>S16</b>
<b>Table S4.....</b>	<b>S16</b>
<b>Table S5.....</b>	<b>S16</b>
<b>Table S6.....</b>	<b>S17</b>
<b>Figure S10.....</b>	<b>S18</b>
<b>Figure S11.....</b>	<b>S19</b>
<b>Figure S12.....</b>	<b>S20</b>
<b>Figure S13.....</b>	<b>S21</b>
<b>References.....</b>	<b>S22</b>



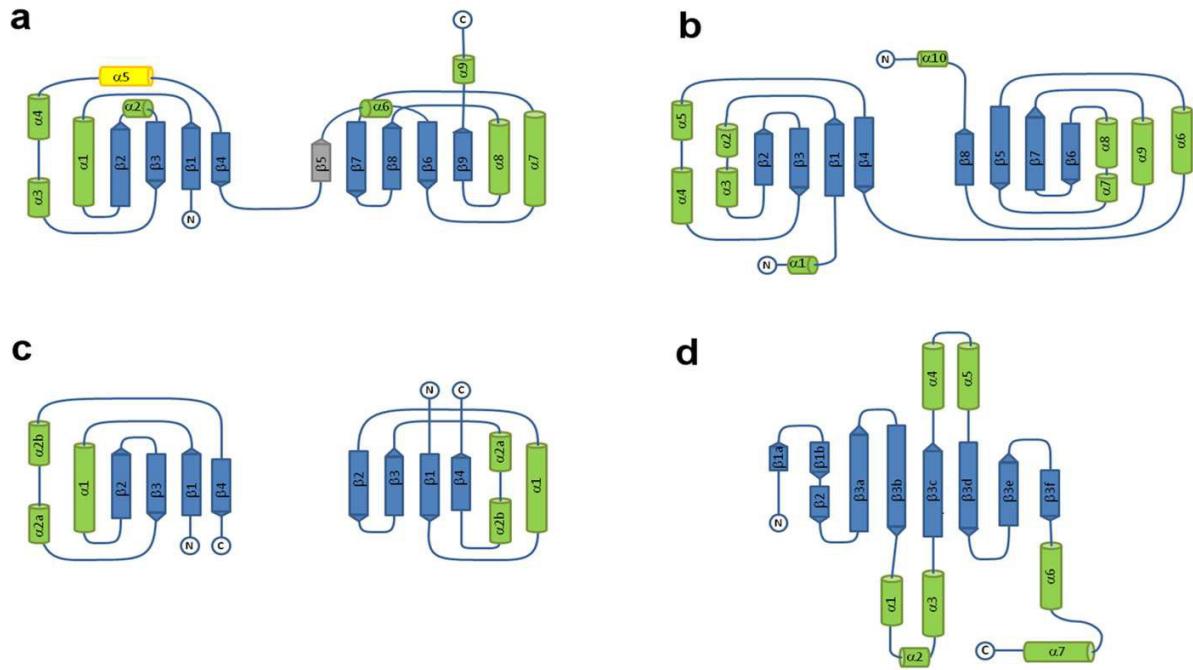
**Figure S1 | Naringenin degradation pathway of *Eubacterium ramulus*.** The equilibrium is on the side of (2S)-naringenin, but chalcone degradation of *Eubacterium ramulus* continues due to the steady state equilibrium.



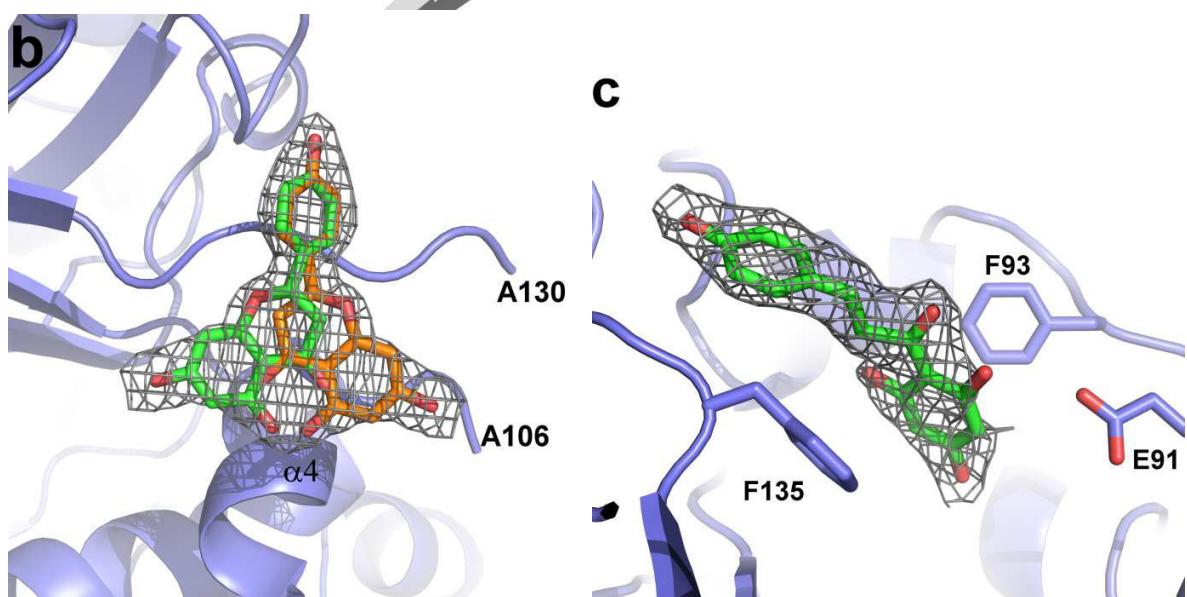
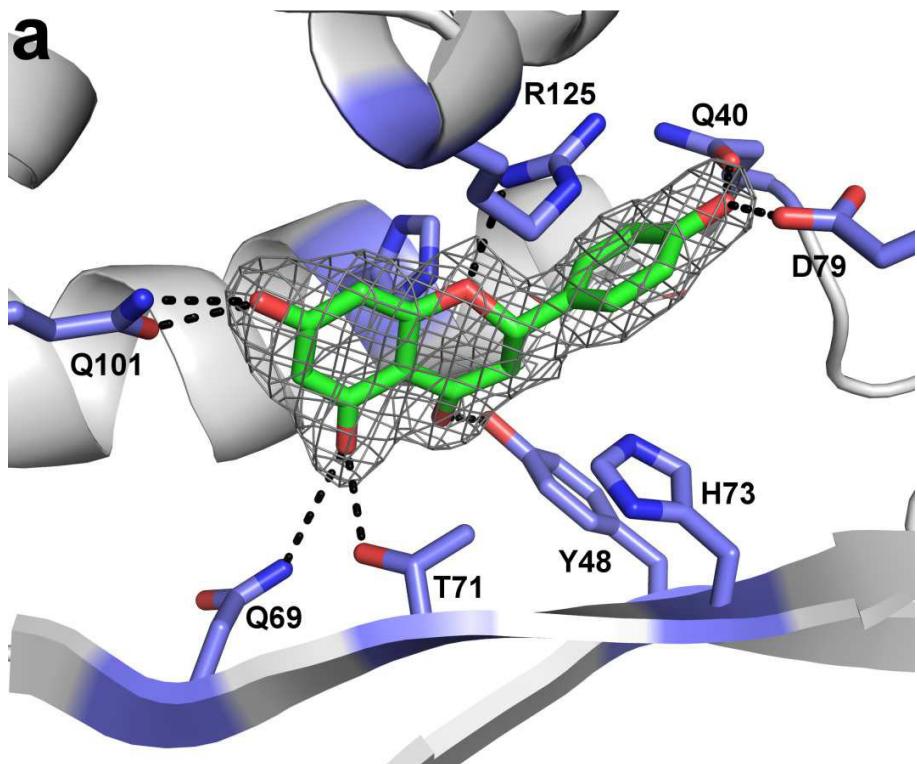
**Figure S2 | D<sub>3</sub> symmetry of the quaternary structure of CHI.** **a**, View down the local 3-fold axis, **b**, view rotated horizontally by 90°.

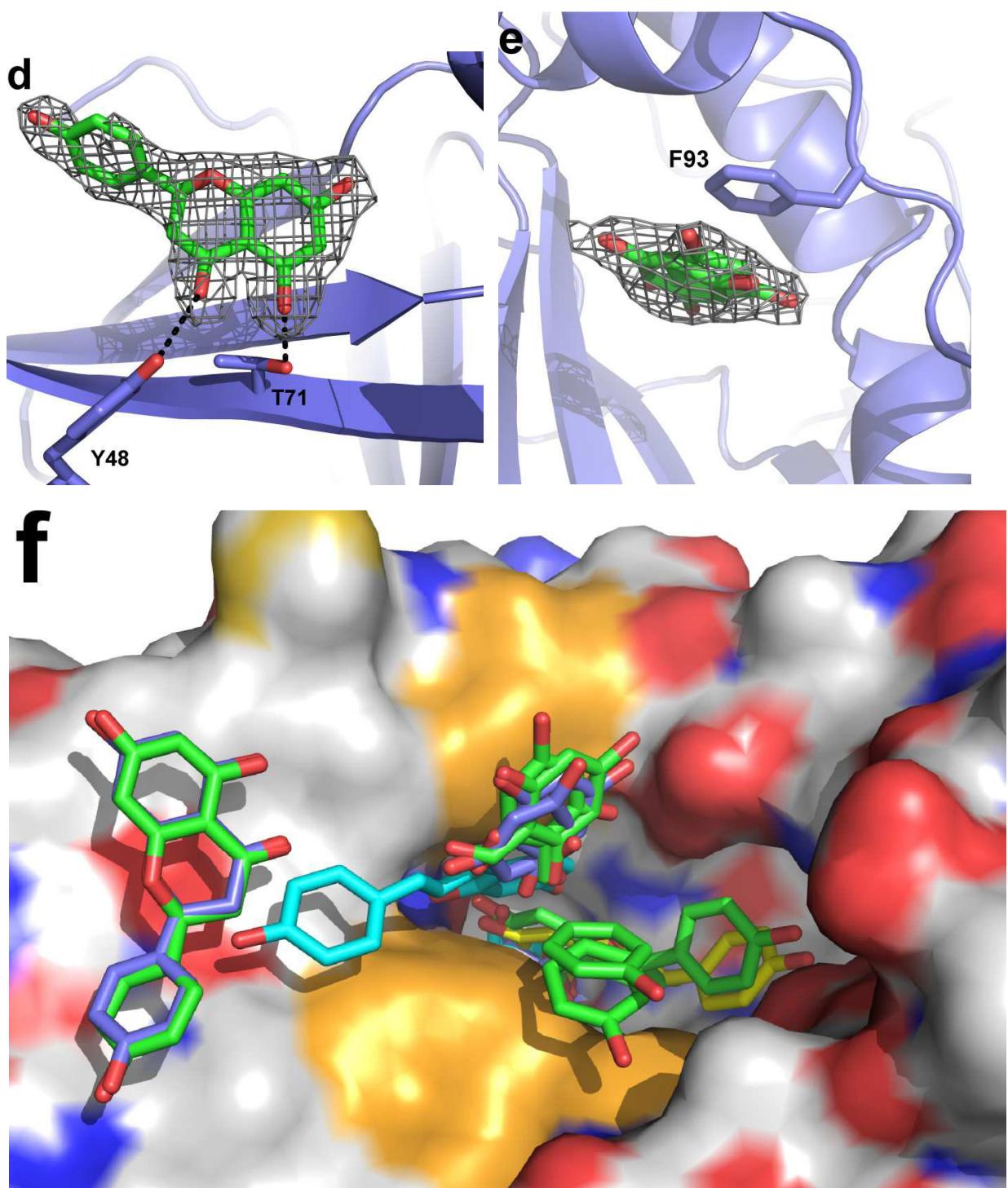


**Figure S3 | C-terminal amino-acid residues 279 – 282 penetrate into the core of a neighbored monomer. a,** Each C-terminus of CHI (residues 264 – 282) contributes to dimer formation. **b,** The hydrogen-bonding pattern of the C-terminal amino-acid residues indicate a remarkable enthalpy term for dimer stabilization.



**Extended Data Figures 4a-d** | Topology diagrams with  $\alpha$ -helices are shown as green cylinders and the  $\beta$  strands as blue arrows (right). The N- and C-termini are labelled. **a**, Bacterial CHI, **b**, chlorite dismutase from *Dechloromonas aromaticata* (PDB entry 3Q08), and **c**, SP1 from *Populus tremula* (PDB entry 1TR0) dimer orientation for comparison with CHI. **d**, In contrast to the ferredoxin-like domains (ordered  $\beta_2$ ,  $\beta_3$ ,  $\beta_1$ ,  $\beta_4$ ) all  $\beta$ -strands of plant CHI from *Medicago sativa* (PDB entry 1EYP) are consecutively arranged  $\beta_{1a}$  to  $\beta_{3a-f}$ .

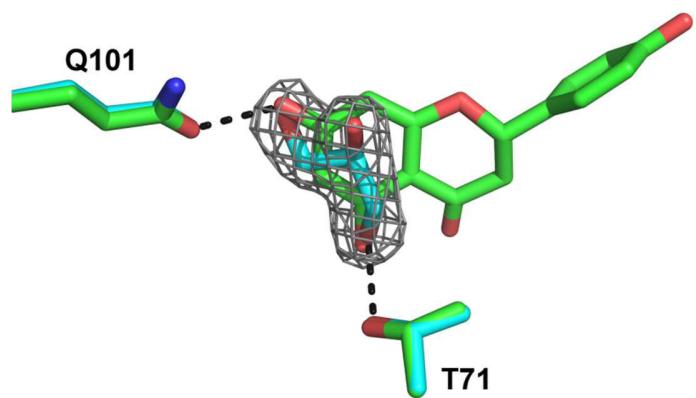




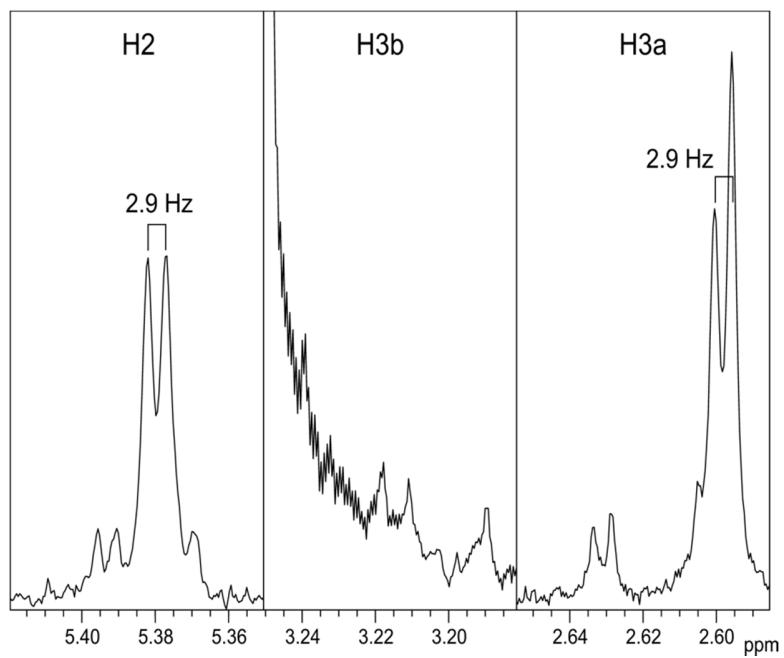
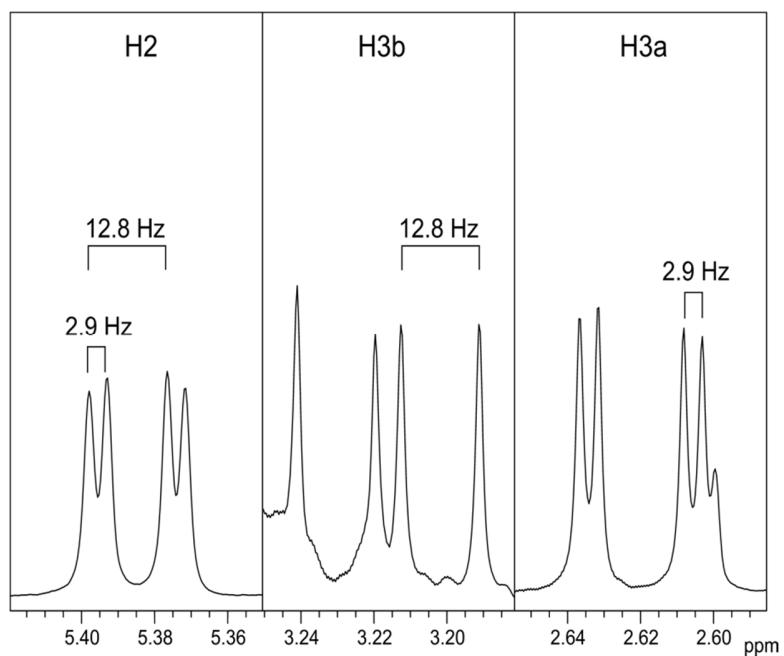
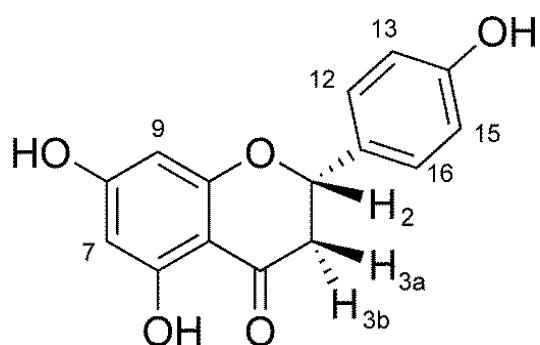
**Figure S5a-f | Various substrate and product binding motifs.** In all CHI monomers with open or closed lid several binding motifs of substrate (naringenin chalcone) and product ((2S)-naringenin) are observed in the active site and in the entry tunnel. This is due to the artificial high naringenin chalcone concentration for co-crystallization. The trimer with open conformation

(monomers B, D, F) shows several additional binding motifs. The  $2F_{\text{obs}} - F_{\text{calc}}$  electron density maps are contoured at  $1\sigma$  level.

- a**, (2S)-naringenin in the active site as observed for monomers with closed lid (A, C, and E).
- b**, (2S)-naringenin on a crystallographic two-fold axis between the monomer D and its symmetry mate ( $-x, -y + \frac{1}{2}, z$ ). This is also detected (not shown) between monomer F and its symmetry mate ( $-x + \frac{1}{2}, y, -z$ ).
- c**, Naringenin chalcone bound in the entry region of the active site of monomer B (open lid).
- d**, (2S)-naringenin in the active site of monomer D that has an open lid conformation.
- e**, Alternative positions of the di-phenolic fragment in monomer F.
- f**, Superimposition of all CHI monomers with observed substrate or product sites. The surface of monomers with open lid conformations (here D) is show as template. Ligands of monomer B (cyan), of monomer D (green), and of monomer F (blue) are shown as stick models. The active site bound naringenin (yellow sticks) observed in monomers A, C, and E (all in closed conformations) is shown for comparison. The surfaces of the interacting hydrophobic residues Phe93, Phe135 and Phe137 are highlighted in orange.

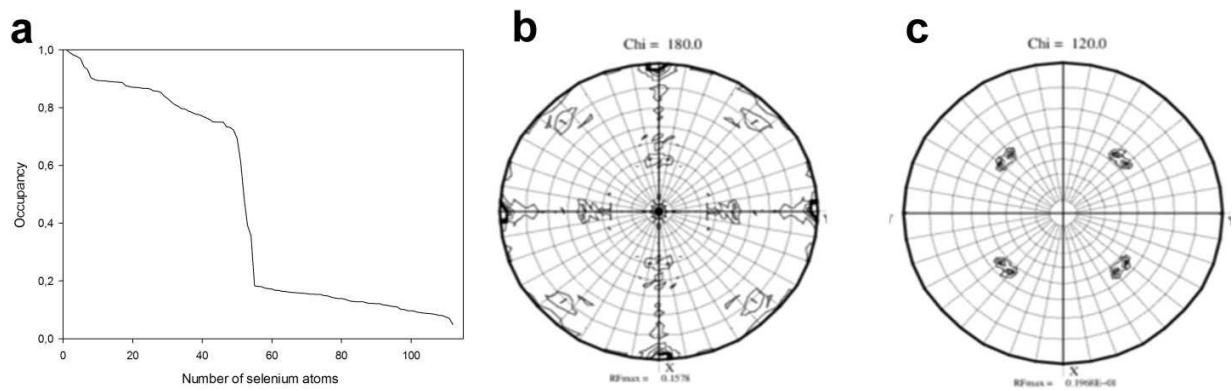


**Figure S6 | Glycerol replacing the hydrogen bonds of the phenolic 4',6'-dihydroxy-fragment of the flavanone.** Superposition of monomer A of the (2S)-naringenin complex and substrate-free CHI (side chains with yellow and light blue C-atoms, respectively). The local  $2F_{\text{obs}} - F_{\text{calc}}$  electron density map of glycerol contoured at  $1\sigma$  is shown.

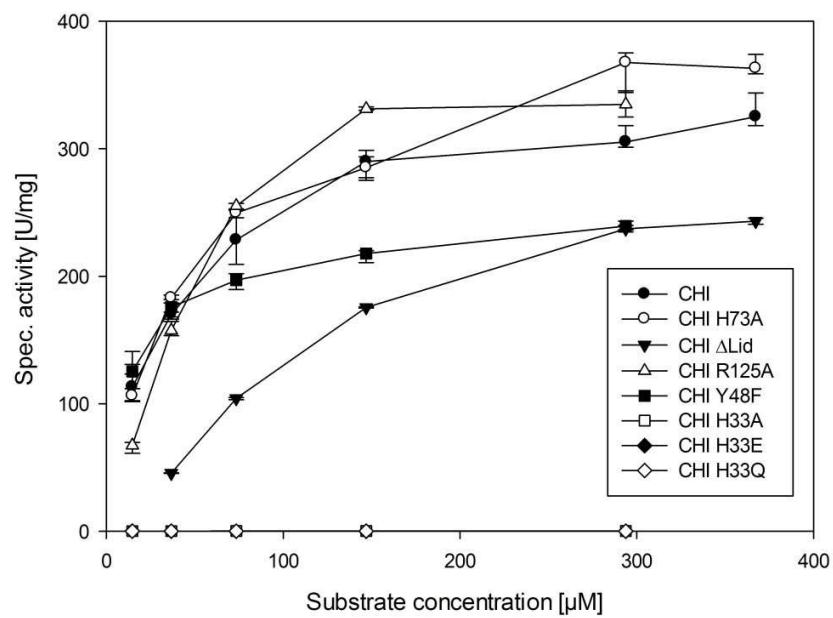
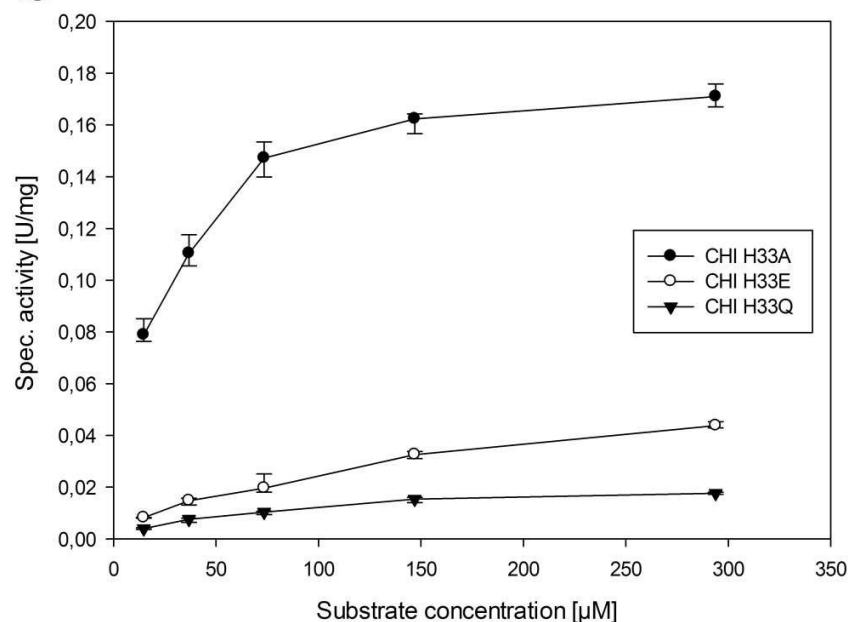


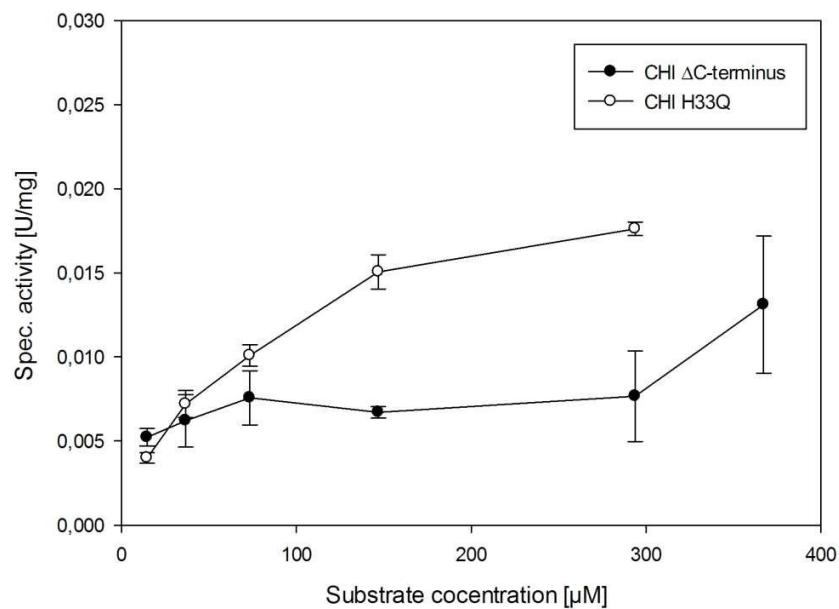
**Figure S7 | Mechanistic studies with  $^1\text{H}$ -NMR measurements.**  $^1\text{H}$  NMR spectral regions of (2S)-naringenin (top) enzymatically produced in  $\text{H}_2\text{O}$  (middle) and in  $\text{D}_2\text{O}$  (bottom). The signal of the proton H3b at the *pro-S* position of C3 is mostly lost through H/D exchange in  $\text{D}_2\text{O}$  solution, confirming its participation in the catalytic cycle.

**(2S)-Naringenin:**  $^1\text{H}$  NMR (600 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) = 2.62 (1H, d/d,  $J$  = 2.9 Hz, 17.1 Hz, H3a), 3.22 (1H, d/d,  $J$  = 12.8 Hz, 17.1 Hz, H3b), 5.38 (1H, d/d,  $J$  = 2.9 Hz, 12.8 Hz, H2), 5.82 (2H, s, H7, H9), 6.73 (2H, d,  $J$  = 8.5 Hz, H13, H15), 7.26 (2H, d,  $J$  = 8.5 Hz, H12, H16), 9.55 (1H, s, OH), 10.75 (1H, s, OH), 12.10 (1H, s, OH).



**Figure S8 | Phasing and packing information.** **a**, Occupancy of the Se positions of selenomethionine labelled CHI found with SHELXD<sup>1,2</sup>. Self-rotation function of native CHI diffraction data calculated with Molrep<sup>3</sup> in space group  $I2_12_12_1$  at, **b**, chi = 180°, and, **c**, chi= 120° revealing three independent twofold and two threefold non-crystallographic axes.

**a****b**

**C**

**Figure S9 | Enzymatic activity of CHI variants.** **a**, The contribution of supposed catalytic residues of CHI for the conversion of naringenin chalcone to (2S)-naringenin is determined. **b**, The CHI variants His33Ala, His33Glu, and His33Gln showed significantly reduced activity. **c**, The C-terminal truncated variant CHI<sub>ΔC-terminus</sub> (missing residues 278 – 282) is less active than CHI\_H33Q and still behaves as hexamer in gel-filtration, but is much less stable than the wild-type CHI.

**Table S1 | Data collection and processing statistics.**

Values in parentheses are for the last resolution shell.

data set	native	Se-met data set	In-house dataset	naringenin
X-ray source	BESSY II/BL 14.1	BESSY II/BL 14.1	Rigaku MicroMax007	BESSY II/BL 14.1
wavelength (Å)	0.97935	0.97935	1.5418	0.97857
temperature (K)	100	100	110	100
space group	$I2_12_12_1$	$I2_12_12_1$	$I2_12_12_1$	$I2_12_12_1$
a /b/c (Å)	171.2/192.2/204.6	172.3/192.6/203.8	177.1/203.1/206.1	181.2/188.0/196.2
max. resolution (Å)	1.8 (1.91 – 1.8)	1.98 (2.10 – 1.98)	2.8 (2.95 – 2.8)	2.0 (2.12 – 2.0 )
unique reflections	309268	460230*	91147	436945*
redundancy	4.15 (4.08)	7.72 (7.3)	4.1 (4.0)	3.49(3.41)
$R_{\text{merge}}$	9.9% (70.6%)	20.1% (125.9%)	19.2% (98.7%)	12.1% (61.7%)
$I/\sigma(I)$	12.25 (2.14)	11.41 (1.75)	5.4 (1.2)	8.94 (1.93)
completeness (%)	99.6 (97.6)	99.7 (98.1)	99.6 (99.9)	99.5 (98.4)
Wilson $B$ factor (Å <sup>2</sup> )	28.5	29.7	41.5	27.2

\* Friedel pairs not merged.

**Table S2 | Refinement statistics.**

Data set	native	Se-Met data set	In-house dataset	naringenin
Resolution (Å)	48.1 – 1.8	48.5 – 2.0	20.1 – 2.8	49.2 – 2.0
R/R <sub>free</sub> (%)†	13.2/15.2	13.4/15.6	24.8/29.6	14.0/16.2
Protein residues/water molecules	1538/2500	1506/2086	1564/1061	1623/1893
No. of Cl <sup>-</sup> /SO <sub>4</sub> <sup>2-</sup> /glycerol/substrate	18/1/6/-	15/1/6/-	4/0/5/-	17/0/2/8
R.m.s.d. from ideality				
Bond lengths (Å)	0.0157	0.0185	0.0119	0.0192
Bond angles (°)	1.707	1.823	1.551	1.938
Average <i>B</i> factor (Å <sup>2</sup> )	22.5	20.5	36.3	24.5
Ramachandran statistics‡ (%)				
Most favored regions	98.2	98.2	97.1	98.3
Outliers	none	0.2	0.1	0.1
PDB entry code	4c9s	4c9t	3zph	4d06

† R<sub>free</sub> = analogous R-factor for 5% randomly chosen reflections excluded from refinement.

‡Categories were defined by MOLPROBITY<sup>4</sup>.

**Table S3 | Sequences of primers used to introduce desired mutations.**

Primer name	Sequence
pET28b_CHI_H33A	5'-TAAAGTAGCGATTCCGGACAGC-3'
pET28b_CHI_H33E	5'-TAAAGTAGAAATTCCGGACAGC-3'
pET28b_CHI_H33Q	5'-TAAAGTACAGATTCCGGACAGC-3'
pET28b_CHI_H73A	5'-GCAGCTGACAGAGGCGCACTGGTTAG-3'
pET28b_CHI_Y48F	5'-CCAAATTGCATTTATCCGTCC-3'
pET_RP	5'-CTAGTTATTGCTCAGCGG-3'
pET28b_CHI_Dlid_fw	5'-GATCCCGGCAGCAGCAGAAGGAAATCC -3'
pET28b_CHI_Dlid_rv	5'-GGATTTCTCTGCTGCTGCCGGGATC -3'
pET28b_CHI_DeltaC-terminus_fw	5'-GCGGCCGCTTATCTCATGGTGATTATCC-3'
pET28b_CHI_DeltaC-terminus_rv	5'-GGATAAACCATGAGATAAGCGGCCGC-3'

**Table S4 | Overview of the secondary structure alignments performed with PDBeFold<sup>5</sup>.**Sequence alignment with SCOPE<sup>6</sup> identified no related proteins.

CHI domain	superposed on (PDB entry)	RMSD on C $\alpha$ -atoms (Å)	Aligned with residues (of target protein)	Q-score
CHI, solvent exposed domain	SP1 (1TR0)	2.37	89 (106)	0.30
CHI, catalytic domain	SP1 (1TR0)	2.94	75 (106)	0.25
CHI, catalytic domain	CHI, solvent exposed domain	2.54	79 (152)	0.22
CHI, catalytic domain	chlorite dismutase (3Q08)	3.22	70 (241)	0.20
CHI, monomer	monooxygenase ACTVA-ORF6 (1N5V)	2.71	85 (258)	0.14
CHI, monomer	chlorite dismutase (3Q08)	4.68	170 (241)	0.13
CHI, monomer	plant CHI from <i>Medicago sativa</i> (1EYP)	2.72	74 (212)	0.055
CHI, monomer	fatty acid binding protein from <i>Arabidopsis thaliana</i> (4DOI)	2.86	73 (221)	0.049

**Table S5 | Melting points (T<sub>m</sub>) of CHI.**

protein	T <sub>m</sub> [°C]
CHI	47.9 ± 0.3
CHI $\Delta$ C-term	39.0 ± 1.2

## SAXS experiments

**Table S6 | SAXS Data Collection and Model-free Parameters.**

<i>Data collection parameters</i>	CHI	naringenin-CHI	CHI_Alid
Instrument	P12 at EMBL/DESY, storage ring PETRA III, Germany		
Beam geometry		0.2 × 0.12 mm <sup>2</sup>	
Wavelength (Å)		1.24	
<i>q</i> -range (Å <sup>-1</sup> )		0.008 – 0.47	
Exposure time (ms)		20 × 50	
Concentration range (mg ml <sup>-1</sup> )	0.1 – 18.2	0.8 – 2.6	1.3 – 9.8
Temperature (K)	283	283	283
<i>Structural parameters*</i>			
<i>I</i> (0) (arbitrary units) (from <i>P</i> ( <i>r</i> ))	25450 ± 10	24640 ± 10	26710 ± 10
<i>R</i> <sub>g</sub> (from <i>P</i> ( <i>r</i> )) (Å)	37 ± 2	36 ± 2	35 ± 2
<i>I</i> (0) (arbitrary units) (from Guinier)	25450 ± 30	25100 ± 20	27220 ± 20
<i>R</i> <sub>g</sub> (Å) (from Guinier)	36 ± 3	37 ± 3	36 ± 3
<i>D</i> <sub>max</sub> (Å)	130 ± 5	110 ± 5	110 ± 5
Porod volume (10 <sup>3</sup> Å <sup>3</sup> )	320 ± 20	320 ± 20	270 ± 20
<i>Molecular mass determination*</i>			
MM <sub>POROD</sub> (from Porod volume) (kDa)	190 ± 10	190 ± 10	160 ± 10
Contrast ( $\Delta\rho \times 10^{10}$ cm <sup>-2</sup> )	3.047	3.047	3.047
MM <sub>saxs</sub> (from <i>I</i> (0), (kDa))	190 ± 20	190 ± 20	150 ± 10
Calculated hexameric MM from sequence (kDa)	194.4	194.4	181.2
SASBDB entry code	SASDAL6	SASDAM6	SASDAN6
<i>Software employed</i>			
Primary data reduction	Automated radial averaging <sup>7</sup>		
Data processing	PRIMUS <sup>8</sup>		
<i>Ab initio</i> analysis	DAMMIN <sup>9</sup>		
Validation and averaging	DAMAVER <sup>10</sup>		
Rigid body modelling	CORAL <sup>11</sup>		
Computation of model intensities	CRYSTAL <sup>12</sup>		

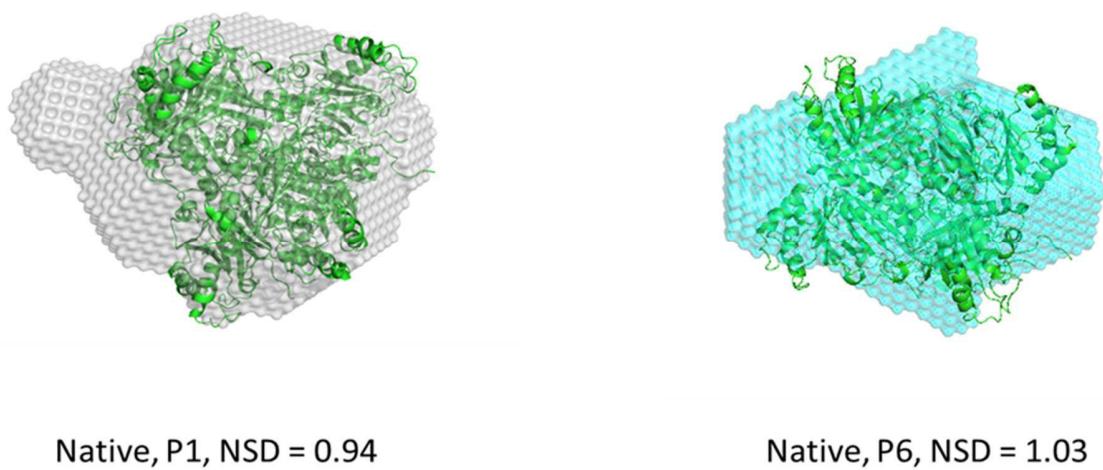
\*Reported for infinite dilution of concentration series measurements

## SAXS results

Small-angle x-ray scattering experiments (SAXS) of the wild-type CHI, its naringenin complex and the CHI<sub>Δ</sub>lid variant clearly show differences caused by lid deletion or lid closing upon naringenin binding.

### 1. Native CHI without naringenin ligand, theoretical MW = 194 kDa

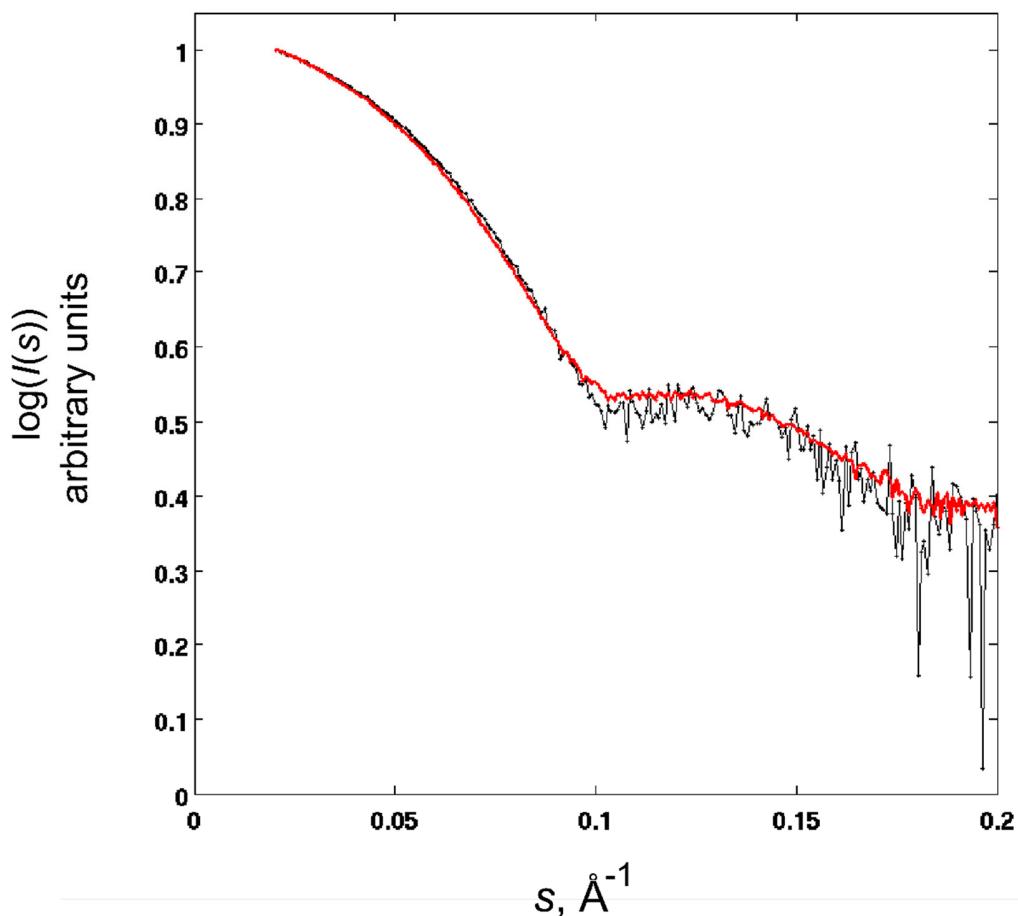
SAXS of the native protein without ligands indicate no ionic strength dependent oligomerization (range of 0 – 1.0 M NaCl). The SAXS-based molecular mass estimates confirm hexameric assembly observed in the crystallographic structures. *Ab initio* models of CHI are in excellent agreement with the overall shape of the high-resolution structure. The normalized spatial discrepancy (NSD) values for P1 and P6 reconstructions are 0.94 and 1.03, respectively<sup>13</sup>. However, the theoretical scattering based on the crystallographic hexamer structure and addition of the missing lids using a single conformation does not fit the experimental SAXS data ( $\chi = 3.4$ ).



**Figure S10 | The *ab initio* models in P1 and P6 symmetries.** *Ab initio* models based on SAXS data have an excellent agreement with the crystallographic structure. The normalized spatial discrepancy (NSD) values for P1 and P6 reconstructions are given.

## 2. CHI with naringenin ligand, theoretical MW = 194 kDa

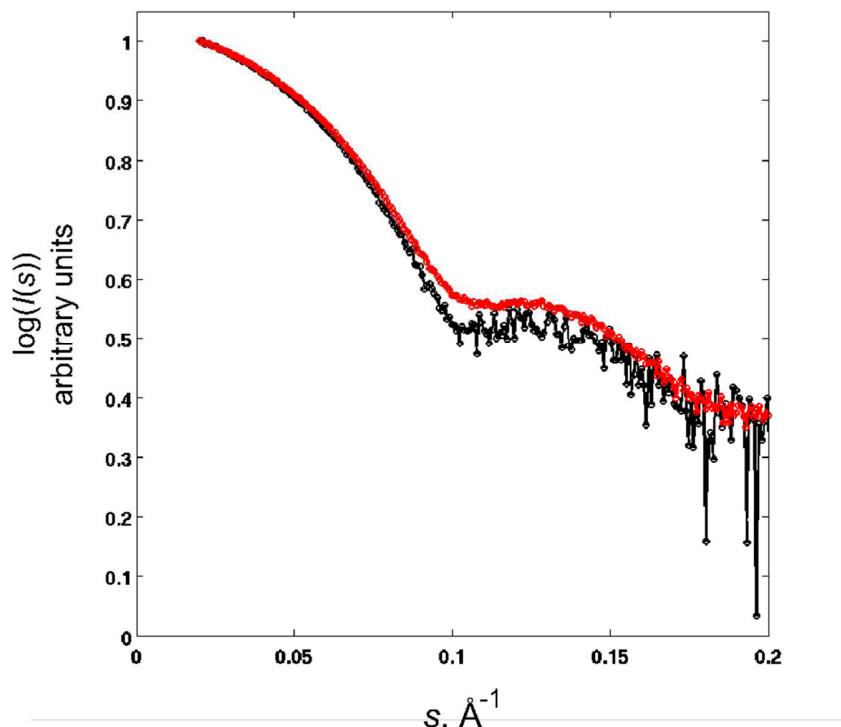
The overall structure of the ligand complex of CHI is more compact than that of the native protein that was observed as differences in  $R_g$  and  $D_{\max}$  values. In addition, there is a change of the minimum of the SAXS profile of around  $\sim 0.1 \text{ \AA}^{-1}$  indicating further a more globular structure than the ligand-free protein has. The structure with all loops closed (constructed using the crystallographic trimer model with closed loops in the crystallographic structure and symmetry operations) yielded a goodness-of-the-fit values of  $\chi = 1.0$  with the experimental data.



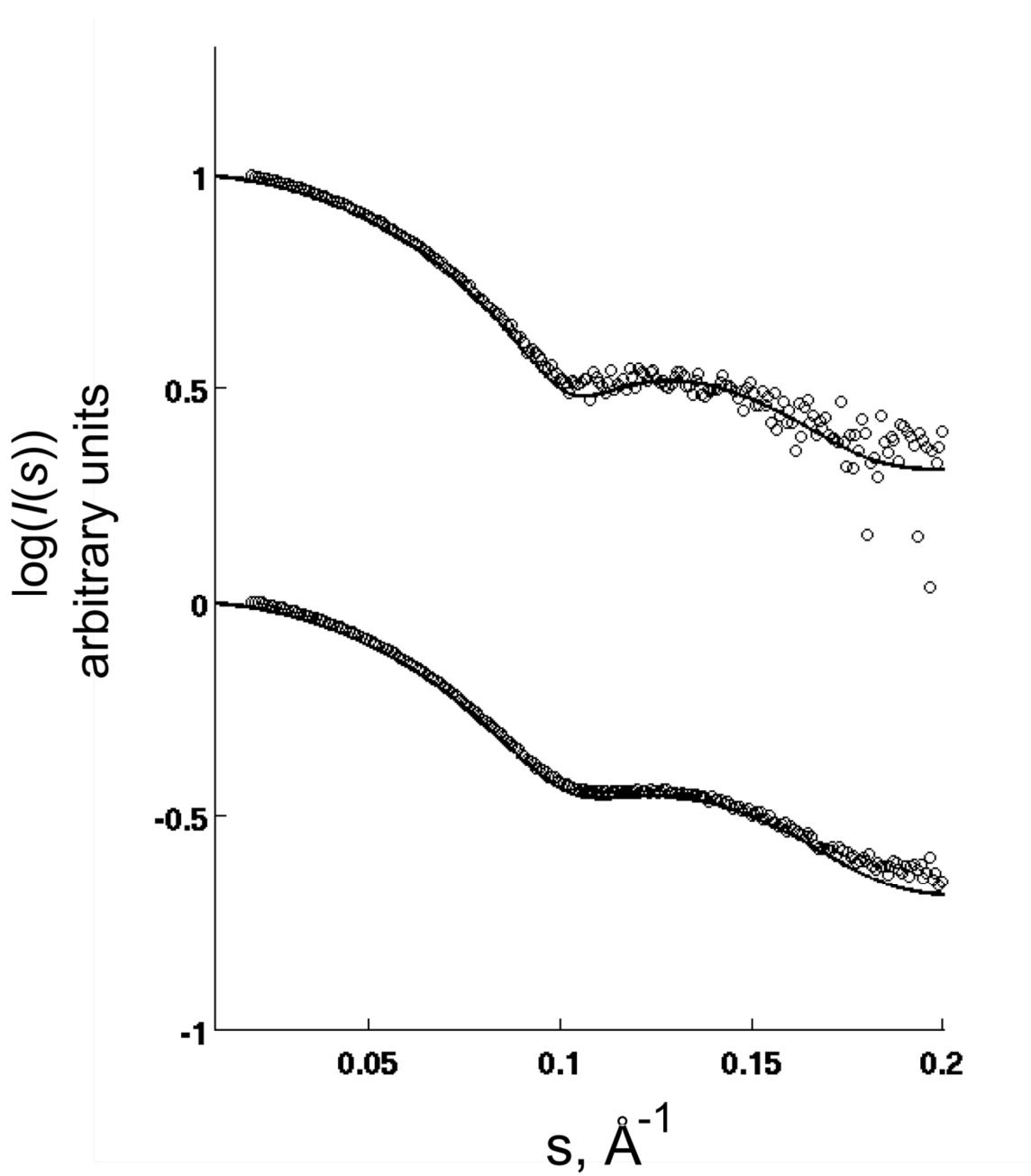
**Figure S11 | SAXS data comparison of CHI in the absence (red) and presence (black) of naringenin.** Concentrations used are as following 13.5  $\mu\text{M}$  CHI, 1 mM naringenin.

### 3. CHI<sub>Δ</sub>lid variant, theoretical MW = 181 kDa

The observed  $R_g$  of the deletion mutant CHI<sub>Δ</sub>lid (missing amino-acid residues His109 – Ala130) is in accordance with the expectations. The best CRYSTAL fit was obtained with the hexameric structure which is missing the lids regions ( $\chi = 1.4$ ).



**Figure S12 | SAXS data comparison of the naringenin complex of CHI (red line) with CHI<sub>Δ</sub>lid (black line).** Concentrations used are as following 13.5 μM CHI, 1 mM naringenin.



**Figure S13 | CRYSTAL fits of the naringenin-bound and unbound CHI structures (PDB codes 4d06 and 4c9s, resp.) to the SAXS data measured in the presence of 1mM naringenin (upper curve) and on the CHI<sub>Δ</sub>lid variant (lower curve). The goodness-of-the-fit values  $\chi$  are 1.0 and 1.4 for the ligand-bound and unbound structures, respectively.**

## References

- (1) Sheldrick, G. *Acta Crystallogr. D* **66**, 479-485 (2010).
- (2) Pape, T. & Schneider, T. R. *J. Appl. Crystallogr.* **37**, 843-844 (2004).
- (3) Collaborative Computational Project Number 4 *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994, 50**, 760.
- (4) Chen, V. B.; Arendall, W. B., III; 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: Biol. Crystallogr.* **2010, 66**, 12.
- (5) Krissinel, E.; Henrick, K. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2004, 60**, 2256.
- (6) Fox, N. K.; Brenner, S. E.; Chandonia, J. M. *Nucleic Acids Res.* **2014, 42**, D304.
- (7) Petoukhov, M. V.; Franke, D.; Shkumatov, A. V.; Tria, G.; Kikhney, A. G.; Gajda, M.; Gorba, C.; Mertens, H. D. T.; Konarev, P. V.; Svergun, D. I. *J. Appl. Crystallogr.* **2012, 45**, 342.
- (8) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *J. Appl. Crystallogr.* **2003, 36**, 1277.
- (9) Svergun, D. I. *Biophys. J.* **1999, 76**, 2879.
- (10) Volkov, V. V.; Svergun, D. I. *J. Appl. Crystallogr.* **2003, 36**, 860.
- (11) Petoukhov, M. V.; Konarev, P. V.; Kikhney, A. G.; Svergun, D. I. *J. Appl. Crystallogr.* **2007, 40**, S223.
- (12) Svergun, D. I.; Barberato, C.; Koch, M. H.J. *J. Appl. Crystallogr.* **1995, 28**, 768.
- (13) Kozin, M. & Svergun, D.I. *J. Appl. Crystallogr.* **2001, 34**, 33.

# **Publikation III**

# Crystallization and preliminary X-ray diffraction studies of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*

**Maren Thomsen,<sup>‡</sup> Lilly Skalden,<sup>‡</sup> Gottfried J. Palm, Matthias Höhne, Uwe T. Bornscheuer and Winfried Hinrichs\***

Institut für Biochemie, Universität Greifswald,  
 Felix-Hausdorff-Strasse 4, D-17489 Greifswald,  
 Germany

<sup>‡</sup> These authors contributed equally to this work.

Correspondence e-mail:  
 winfried.hinrichs@uni-greifswald.de

Received 4 October 2013  
 Accepted 11 November 2013

The (*R*)-selective amine transaminase from *Aspergillus fumigatus* was expressed in *Escherichia coli* and purified to homogeneity. Bright yellow crystals appeared while storing the concentrated solution in the refrigerator and belonged to space group *C222*<sub>1</sub>. X-ray diffraction data were collected to 1.27 Å resolution, as well as an anomalous data set to 1.84 Å resolution that was suitable for S-SAD phasing.

## 1. Introduction

Transaminases belong to the pyridoxal-5'-phosphate (PLP)-dependent enzymes and catalyze the reversible transfer of an amino group to an  $\alpha$ -keto acid, ketone or aldehyde (Hayashi, 1995). The PLP and the catalytic lysine side chain are the key elements in this reaction (Eliot & Kirsch, 2004).

Transaminases are of biotechnological significance because of their ability to produce enantiopure amines from prochiral precursors. These amines are applied as ingredients or synthons in medicine, agrochemistry, pharmacy and chemistry (Merck, 2001; Deng *et al.*, 1995; Martens *et al.*, 1986; Höhne & Bornscheuer, 2009).

Based on their substrate range, transaminases can be divided into  $\alpha$ -transaminases,  $\omega$ -transaminases and amine transaminases. Whereas the substrates of  $\alpha$ -transaminases require a carboxylate in the  $\alpha$  position, the substrates of  $\omega$ -transaminases have up to five extra C atoms between the terminal amino function and the carboxylate. The substrates of amine transaminases can lack the carboxyl group completely (Höhne & Bornscheuer, 2012; Mani Tripathi & Rama-chandran, 2006). Amine transaminases often show excellent enantio-selectivity and can be grouped into two classes. (*R*)-Amines are generated by (*R*)-selective amine transaminases when the quinoid intermediate of the reaction is protonated from the catalytic lysine at the *si*-site (Hanson, 1966). Alternatively, the (*S*)-amine is produced by an (*S*)-amine transaminase when the protonation occurs at the *re*-site.

*Aspergillus fumigatus* is a mildew which can cause respiratory allergy. It is a thermophilic saprophytic fungus with a worldwide distribution (Latgé, 1999). The sequence of an (*R*)-selective amine transaminase from *A. fumigatus* was identified by an *in silico* search (Höhne *et al.*, 2010) and is available online at NCBI (NCBI Reference Sequence XP\_748821.1).

Several structures of  $\alpha$ -transaminases have been described and these enzymes have been studied in detail (Schwarzenbacher *et al.*, 2004; Han *et al.*, 2006). Recently, a few crystal structures of non-homologous (*S*)-selective amine transaminases have been published (Steffen-Munsberg *et al.*, 2013; Humble *et al.*, 2012). In contrast, only a homology model of an (*R*)-selective amine transaminase from an *Arthrobacter* species based on a  $\beta$ -amino-acid aminotransferase (PDB entry 3daa) has been published (Savile *et al.*, 2010). Here, we describe expression, purification, crystallization and initial crystallographic results to elucidate the structure of the (*R*)-selective amine transaminase (AspFum) from *A. fumigatus*.



© 2013 International Union of Crystallography  
 All rights reserved

## 2. Materials and methods

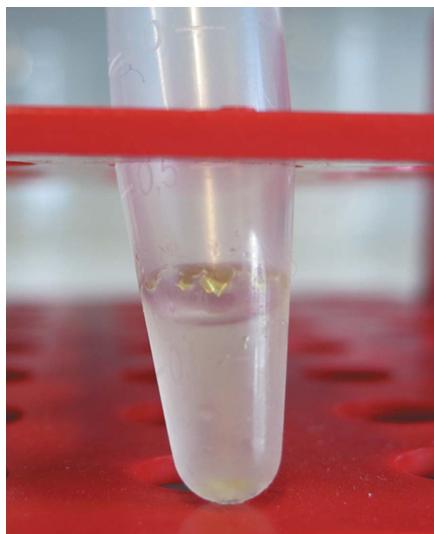
### 2.1. Protein expression and purification

The gene for the amine transaminase was expressed in *Escherichia coli* BL21 (DE3) cells containing the expression vector pET-22b, which encodes the sequence of the amine transaminase including an additional C-terminal His<sub>6</sub> tag (SGSHHHHH; Höhne *et al.*, 2010). The recombinant protein consists of 332 amino-acid residues with a molecular weight of 37.16 kDa. The cells were grown at 310 K in 400 ml LB medium containing 0.1 mg ml<sup>-1</sup> ampicillin until an OD<sub>600</sub> of 0.4 was reached. The temperature was then reduced to the expression temperature of 293 K and the cells were further incubated until they reached an OD<sub>600</sub> of 0.7. Expression of the protein was induced by the addition of 1 mM IPTG. The cells were harvested 20 h after induction (Höhne *et al.*, 2010).

The cell pellet was resuspended in 50 mM sodium phosphate buffer pH 7.5, 300 mM sodium chloride (buffer *A*) containing an additional 0.1 mM PLP and 30 mM imidazole. Cell disruption was performed by two passages through a French press at 10.3 MPa. The resulting suspension was centrifuged for 45 min at 10 000g. The filtrated supernatant was applied onto a nickel-NTA column (GE Healthcare). After washing with three column volumes of buffer *A* containing 60 mM imidazole at a flow rate of 5 ml min<sup>-1</sup>, the protein was eluted with buffer *A* containing 300 mM imidazole. The amine transaminase-containing fractions were identified using an acetylphenone assay (Schätzle *et al.*, 2009), collected and pooled. The pooled protein was then desalting by gel chromatography against 20 mM tricine buffer pH 7.5, 10 µM PLP at a flow rate of 2 ml min<sup>-1</sup> (Höhne *et al.*, 2010). The desired concentration of AspFum was achieved by ultrafiltration with Vivaspin 6 columns (molecular-weight cutoff 10 kDa; Sartorius Stedim).

### 2.2. Crystallization

Initial crystallization hits were obtained with a variety of PEG-based conditions (JBScreen Classic 1–10, Jena Bioscience) within 4 d. However, all diffraction images of these crystals were not indexable. Suitable crystals of AspFum appeared after six months in an Eppendorf reaction tube containing concentrated protein (10.7 mg ml<sup>-1</sup>) and 20 mM tricine pH 7.5 with 10 µM PLP at 277 K.



**Figure 1**

Crystals of the (*R*)-selective amine transaminase from *A. fumigatus* grown in an Eppendorf reaction tube.

**Table 1**

Data-collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

Data set	Native	Anomalous
Beamline	14.1, BESSY II	14.1, BESSY II
Detector	Pilatus 6M	Pilatus 6M
Wavelength (Å)	0.91841	1.77122
Temperature (K)	100	100
Orthorhombic space group	C222 <sub>1</sub>	C222 <sub>1</sub>
Unit-cell parameters (Å)	$a = 102.2, b = 120.9,$ $c = 135.4$	$a = 102.2, b = 120.9,$ $c = 135.4$
Resolution range (Å)	50.0–1.27 (1.35–1.27)	50.0–1.84 (1.95–1.84)
No. of unique reflections	426722 (68273)	135117 (17260)
Multiplicity	3.38 (3.3)	5.6 (2.6)
$R_{\text{merge}}$ (%)	6.3 (60.3)	3.9 (9.5)
Mean $I/\sigma(I)$	13.2 (2.0)	30.29 (8.34)
$CC_{1/2}^{\dagger}$ (%)	99.9 (73.0)	99.9 (98.8)
Completeness (%)	99.1 (97.9)	95.9 (75.8)
Overall $B$ factor from Wilson plot (Å <sup>2</sup> )	17.4	18.8
Total rotation, increment (°)	180, 0.1	360, 0.1

<sup>†</sup> CC<sub>1/2</sub> is the percentage correlation between intensities from random half data sets (Karplus & Diederichs, 2012).

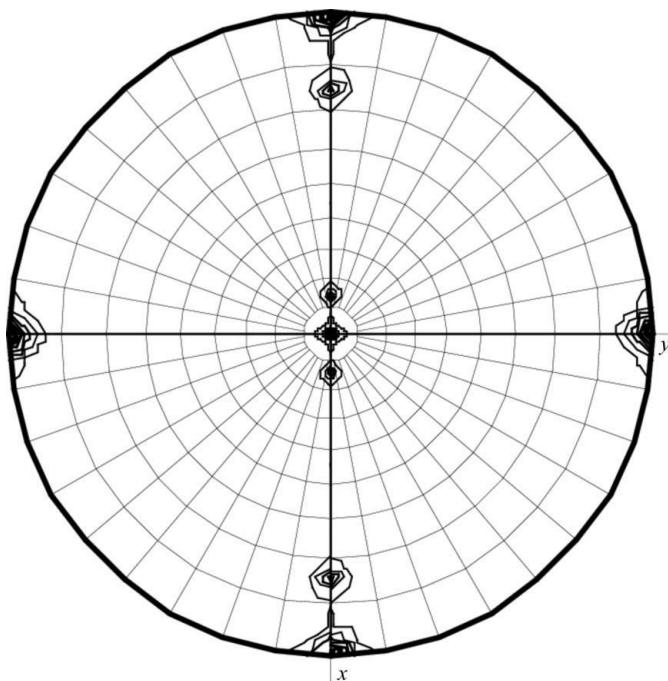
The crystals could easily be seen by eye (>1 mm) and had a very bright yellow colour, suggesting bound PLP (Fig. 1). Only the small crystals (<0.4 mm) present at the bottom of the tube diffracted to high resolution. The mechanical stress on the large crystals attached to the wall of the reaction tube while fishing and upon cooling led to loss of diffraction quality.

### 2.3. Data collection and X-ray crystallographic analysis

For cryoprotection, a solution consisting of 35% (v/v) glycerol, 20 mM tricine pH 7.5, 10 µM PLP was used. X-ray diffraction data were collected at 100 K on beamline 14.1 at the BESSY II synchrotron source, Berlin, Germany (Mueller *et al.*, 2012). Two data sets were collected from one crystal. The first was collected at a wavelength of 0.9184 Å using the highest intensity and the second was collected at 1.77 Å to obtain a large anomalous signal from the S atoms present in the protein. The resolution range of the anomalous data set was limited by the detector geometry. All diffraction images were processed with *XDS* (Kabsch, 2010) using the graphical user interface *XDSapp* (Krug *et al.*, 2012). The rotation function was calculated using *MOLREP* (Vagin & Teplyakov, 2010; Winn *et al.*, 2011) with a resolution range of 30–3 Å and a radius of integration of 30 Å. Data-collection and processing statistics are given in Table 1.

## 3. Results and discussion

The (*R*)-selective amine transaminase from *A. fumigatus* was successfully expressed, purified and crystallized and X-ray diffraction data collection was performed. The calculation of the Matthews coefficient  $V_M$  (Matthews, 1968) as 2.9 Å<sup>3</sup> Da<sup>-1</sup> with a corresponding solvent content of 58% for two monomers offers the most probable solution. The self-rotation function (Fig. 2) shows an independent noncrystallographic twofold axis. Based on the self-rotation function and the Matthews coefficient, we deduced the presence of a dimer in the asymmetric unit. The structure could be solved directly at the beamline using the SAS protocol of the automated crystal structure-determination platform *Auto-Rickshaw* (Panjikar *et al.*, 2005), which incorporates *SHELXC* (Sheldrick, 2001), *SHELXD* (Schneider & Sheldrick, 2002), *ABS* (Hao, 2004), *SHELXE* (Sheldrick, 2002) and *DM* (Cowtan, 1994). Automatic tracing using *ARP/wARP* (Perrakis *et al.*, 1999) yielded 97% of the polypeptide model and indeed shows

**Figure 2**

The self-rotation function at  $\chi = 180^\circ$  for the diffraction data of (*R*)-selective amine transaminase from *A. fumigatus* in space group  $C222_1$  reveals one independent twofold axis with noncrystallographic symmetry. In the orthorhombic space group the dyad-related monomers and their rotational symmetry mates display 16 noncrystallographic relationships including eight twofold axes. In the packing arrangement these axes coincide pairwise, causing four peaks (60% of the origin) in the self-rotation function at  $\chi = 180^\circ$  with  $\omega = 15$  or  $75^\circ$  and  $\varphi = 0$  or  $180^\circ$ .

a dimer in the asymmetric unit. Currently, manual completion of the model and refinement against the high-resolution data is in progress.

Similarly to this amine transaminase, we have crystallized another (*R*)-selective amine transaminase from *Neosartorya fischeri* (96% sequence identity) from a concentrated protein solution without adding a specific precipitant.

MT thanks the Landesgraduiertenkolleg Mecklenburg-Vorpommern for financial support. We thank the European Union (KBBE-2011-5, grant No. 289350) for financial support within the European Union Seventh Framework Programme. Diffraction data were collected on BL14.1 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron-storage ring (Berlin-Adlershof, Germany).

## References

- Cowtan, K. (1994). *Jnt CCP4/ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34–38.
- Deng, L., Mikusová, K., Robuck, K. G., Scherman, M., Brennan, P. J. & McNeil, M. R. (1995). *Antimicrob. Agents Chemother.* **39**, 694–701.
- Eliot, A. C. & Kirsch, J. F. (2004). *Annu. Rev. Biochem.* **73**, 383–415.
- Han, Q., Robinson, H., Gao, Y. G., Vogelaar, N., Wilson, S. R., Rizzi, M. & Li, J. (2006). *J. Biol. Chem.* **281**, 37175–37182.
- Hanson, K. R. (1966). *J. Am. Chem. Soc.* **88**, 2731–2742.
- Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.
- Hayashi, H. (1995). *J. Biochem.* **118**, 463–473.
- Höhne, M. & Bornscheuer, U. T. (2009). *ChemCatChem*, **1**, 42–51.
- Höhne, M. & Bornscheuer, U. T. (2012). *Enzymes in Organic Synthesis*, edited by W. Drauz, H. Gröger & O. May, pp. 779–820. Weinheim: Wiley-VCH.
- Höhne, M., Schätzle, S., Jochens, H., Robins, K. & Bornscheuer, U. T. (2010). *Nature Chem. Biol.* **6**, 807–813.
- Humble, M. S., Cassimjee, K. E., Håkansson, M., Kimbung, Y. R., Walse, B., Abedi, V., Federsel, H. J., Berglund, P. & Logan, D. T. (2012). *FEBS J.* **279**, 779–792.
- Kabsch, W. (2010). *Acta Cryst. D66*, 125–132.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Krug, M., Weiss, M. S., Heinemann, U. & Mueller, U. (2012). *J. Appl. Cryst.* **45**, 568–572.
- Latgé, J.-P. (1999). *Clin. Microbiol. Rev.* **12**, 310–350.
- Mani Tripathi, S. & Ramachandran, R. (2006). *J. Mol. Biol.* **362**, 877–886.
- Martens, J., Günther, K. & Schickedanz, M. (1986). *Arch. Pharm.* **319**, 461–465.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Merck (2001). *The Merck Index: An Encyclopedia of Chemicals, Drugs and Biologicals*. Whitehouse Station: Merck Manuals.
- Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). *J. Synchrotron Rad.* **19**, 442–449.
- Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst. D61*, 449–457.
- Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
- Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W. & Hughes, G. J. (2010). *Science*, **329**, 305–309.
- Schätzle, S., Höhne, M., Redestad, E., Robins, K. & Bornscheuer, U. T. (2009). *Anal. Chem.* **81**, 8244–8248.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst. D58*, 1772–1779.
- Schwarzenbacher, R. et al. (2004). *Proteins*, **55**, 759–763.
- Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
- Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Usón, I. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 333–351. Dordrecht: Kluwer Academic Publishers.
- Steffen-Munsberg, F., Vickers, C., Thontowi, A., Schätzle, S., Tumlirsch, T., Humble, M. S., Land, H., Berglund, P., Bornscheuer, U. T. & Höhne, M. (2013). *ChemCatChem*, **5**, 150–153.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D66*, 22–25.
- Winn, M. D. et al. (2011). *Acta Cryst. D67*, 235–242.

## **Publikation IV**

# Crystallographic characterization of the (R)-selective amine transaminase from *Aspergillus fumigatus*

Maren Thomsen,<sup>‡</sup> Lilly Skalden,<sup>‡</sup>  
Gottfried J. Palm, Matthias  
Höhne, Uwe T. Bornscheuer and  
Winfried Hinrichs\*

Institute of Biochemistry, University of  
Greifswald, Felix-Hausdorff-Strasse 4,  
17489 Greifswald, Germany

<sup>‡</sup> These authors contributed equally to this  
work.

Correspondence e-mail:  
winfried.hinrichs@uni-greifswald.de

The importance of amine transaminases for producing optically pure chiral precursors for pharmaceuticals and chemicals has substantially increased in recent years. The X-ray crystal structure of the (R)-selective amine transaminase from the fungus *Aspergillus fumigatus* was solved by S-SAD phasing to 1.84 Å resolution. The refined structure at 1.27 Å resolution provides detailed knowledge about the molecular basis of substrate recognition and conversion to facilitate protein-engineering approaches. The protein forms a homodimer and belongs to fold class IV of the pyridoxal-5'-phosphate-dependent enzymes. Both subunits contribute residues to form two active sites. The structure of the holoenzyme shows the catalytically important cofactor pyridoxal-5'-phosphate bound as an internal aldimine with the catalytically responsible amino-acid residue Lys179, as well as in its free form. A long N-terminal helix is an important feature for the stability of this fungal (R)-selective amine transaminase, but is missing in branched-chain amino-acid aminotransferases and D-amino-acid aminotransferases.

Received 19 December 2013  
Accepted 15 January 2014

PDB reference: (R)-selective  
amine transaminase, 4chi

## 1. Introduction

During the last decade, interest in transaminases has increased strongly (Koszelewski *et al.*, 2010; Kroutil *et al.*, 2013; Malik *et al.*, 2012; Mathew & Yun, 2012; Rudat *et al.*, 2012; Tufvesson *et al.*, 2011). Many new transaminases have been discovered and applied in organic syntheses to obtain optically pure amines and non-natural amino acids for chemical and pharmaceutical applications (Höhne & Bornscheuer, 2012). This includes oxazolone derivatives used for the treatment of diabetes (Sutin *et al.*, 2007), rivastigmine serving in the treatment of Alzheimer's disease (Fuchs *et al.*, 2010; Rösler *et al.*, 1999), a protected kedarcidine aglycon useful as an antitumour antibiotic (Ogawa *et al.*, 2009), mexiletine for the treatment of cardiac arrhythmia (Koszelewski, Clay *et al.*, 2009; Koszelewski, Pressnitz *et al.*, 2009) and imagabalin, which has been suggested for the treatment of generalized anxiety disorder (Midelfort *et al.*, 2013).

Transaminases belong to the pyridoxal-5'-phosphate (PLP)-dependent enzymes. Besides transamination, the cofactor PLP facilitates a broad variety of other enzymatic reactions such as racemization, decarboxylation and elimination, where it serves as an electron sink to stabilize carbanion intermediates (Christen & Mehta, 2001). The reaction catalyzed by transaminases is the reversible conversion of  $\alpha$ -keto acids, ketones and aldehydes to the corresponding amino acids or amines (Hayashi, 1995). The catalysis itself is divided into two half-reactions. During the first half-reaction the amino group of a

suitable amino donor is transferred to PLP to yield pyridoxamine-5'-phosphate (PMP) with the simultaneous release of the co-product, the deaminated donor. In the second half-reaction the amino acceptor is converted to the corresponding amine and PLP is thus regenerated (Eliot & Kirsch, 2004; Jansonius, 1998). Transaminases can be used in the kinetic resolution of racemic amines and amino acids with a maximum yield of one enantiomer or in asymmetric synthesis starting from prochiral ketones to yield the corresponding optically pure amine at up to 100% yield, if a suitable method to shift the equilibrium to amine formation is employed. In particular, the latter method makes them very useful in the production of building blocks for pharmaceuticals (Martens & Schickedanz, 1986; Blaser, 2002).

Transaminases can be divided into  $\alpha$ -transaminases,  $\omega$ -transaminases and amine transaminases based on their substrate scope. Whereas the substrates of  $\alpha$ -transaminases require a carboxylate group in the  $\alpha$ -position to the carbonyl function,  $\omega$ -transaminases also accept substrates with several C atoms (Schrewe *et al.*, 2013) between the carbonyl and the carboxylic acid function and, typically, the ketone or aldehyde function is at the (sub-)terminal C atom of the substrate. Amine transaminases convert ketones to amines and do not require a carboxylate group in the substrate (Höhne & Bornscheuer, 2012).

Seven fold classes of PLP-dependent enzymes are currently known, and transaminases have been identified in classes I and IV (Eliot & Kirsch, 2004; Jansonius, 1998). All of the members of these fold classes share the characteristic that the smallest catalytic unit is a homodimer (Eliot & Kirsch, 2004). The monomer can be divided into a large and a small domain. The two active sites lie at the interface between the domains, and amino-acid residues of each monomer contribute to the catalytic centre. The active sites of fold classes I and IV can be regarded as mirror images. Whereas (*S*)-selective amine transaminases occur in fold class I, (*R*)-selective amine transaminases belong to fold class IV (Jansonius, 1998; Eliot & Kirsch, 2004). This assignment also matches observations during protonation in the catalytic mechanism. In (*R*)-selective amine transaminases the *si*-site (Hanson, 1966) of the generated quinoid intermediate is solvent-facing, whereas in the (*S*)-selective amine transaminases it is the *re*-site.

To enable the production of enantiopure compounds, amine transaminases with both enantiopreferences are required. In 2010, Höhne and coworkers discovered 17 (*R*)-selective amine transaminases using an *in silico* search (Höhne *et al.*, 2010). To find these putative (*R*)-selective amine transaminases sequences, the *in silico* search was based on the determination of specific sequence motifs which characterize either D-amino-acid aminotransferases (D-ATAs) or branched-chain amino-acid aminotransferases (BCATs) to filter out motifs for (*R*)-selective amine transaminases. Based on these criteria, the sequences of BCATs and D-ATAs could be excluded and the remaining sequences (approximately 0.4% of all investigated sequences) were experimentally confirmed to be (*R*)-selective amine transaminases (Höhne *et al.*, 2010). Structures of  $\alpha$ -transaminases (Schwarzenbacher *et al.*, 2004; Han *et al.*,

**Table 1**

Data-collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

Data set	Native	Anomalous
Beamline	BL14.1, BESSY II	BL14.1, BESSY II
Detector	Pilatus 6M	Pilatus 6M
Wavelength (Å)	0.91841	1.77122
Temperature (K)	100	100
Space group	C222 <sub>1</sub>	C222 <sub>1</sub>
Unit-cell parameters (Å)	$a = 102.2, b = 120.9, c = 135.4$	$a = 102.2, b = 120.9, c = 135.4$
Resolution range (Å)	50.0–1.27 (1.35–1.27)	50.0–1.84 (1.95–1.84)
No. of unique reflections	426722 (68273)	135117 (17260)
Multiplicity	3.38 (3.3)	5.6 (2.6)
$R_{\text{merge}}^{\dagger}$ (%)	6.3 (60.3)	3.9 (9.5)
Mean $I/\sigma(I)$	13.2 (2.0)	30.29 (8.34)
$CC_{1/2}^{\ddagger}$ (%)	99.9 (73.0)	99.9 (98.8)
Completeness (%)	99.1 (97.9)	95.9 (75.8)
Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> )	17.4	18.8
Total rotation/increment (°)	180/0.1	360/0.1

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the observed intensity and  $\langle I(hkl) \rangle$  is the average intensity of multiple measurements.  $\ddagger CC_{1/2}$  is the percentage correlation between intensities from random half data sets (Karpus & Diederichs, 2012).

2006) and also of a few (*S*)-selective amine transaminases have been published and investigated (Humble *et al.*, 2012; Sayer *et al.*, 2013; Steffen-Munsberg *et al.*, 2013), but a structural analysis of an (*R*)-selective amine transaminase has not been published to date. Presently, a homology model of an amine transaminase from *Arthrobacter* sp. is the only existing toehold (Savile *et al.*, 2010).

In this paper, we present the crystal structure analysis of the (*R*)-selective amine transaminase from the fungus *Aspergillus fumigatus*.

## 2. Materials and methods

### 2.1. Expression and purification

The expression, purification and crystallization of the (*R*)-selective amine transaminase from *A. fumigatus* were performed as reported previously (Thomsen *et al.*, 2013).

### 2.2. Crystallization and diffraction data collection

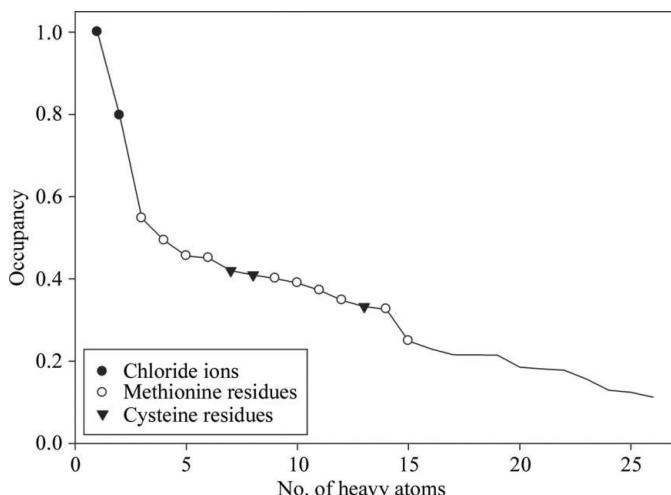
For cryoprotection, a solution consisting of 35% glycerol, 20 mM tricine pH 7.5, 10 µM PLP was used. X-ray diffraction data were collected at 100 K on beamline 14.1 at the BESSY II synchrotron, Berlin, Germany. Two data sets were collected from one crystal. The first data set at a wavelength of 0.918 Å was obtained using the highest intensity of the storage ring and the second was collected at a wavelength of 1.771 Å to obtain the highest anomalous signal of the S atoms present in the protein. The resolution of the anomalous data set was limited by the detector size. All diffraction images were processed with *XDS* (Kabsch, 2010) using the graphical user interface *XDSapp* (Krug *et al.*, 2012). Data-collection and processing statistics are summarized in Table 1.

### 2.3. Structure determination of the holoenzyme by S-SAD phasing

The crystal structure of the (*R*)-selective amine transaminase from *A. fumigatus* was determined by single-wavelength anomalous dispersion sulfur (S-SAD) phasing using the ‘native crystals SAS’ protocol of the automated crystal structure determination platform *Auto-Rickshaw* (Panjikar *et al.*, 2005). SAD was preferred over MAD because higher CC (all/weak) parameters were obtained in *SHELXD*. The automated SAS protocol incorporates *SHELXC* (Sheldrick *et al.*, 2001) for data preparation as well as *SHELXD* (Schneider & Sheldrick, 2002) to find heavy-atom positions. With two cysteine and ten methionine residues per monomer, we searched for 24 S-atom positions per asymmetric unit. The resolution limit for substructure determination and initial phasing was set to 2.5 Å. The best solution obtained resulted in CC (all/weak) of 27.11/16.93 and a PATFOM of 3.68. The program *ABS* (Hao, 2004) determined the correct hand of the substructure, which was subsequently used by *SHELXE* (Sheldrick, 2002) for initial phasing. SAD phasing statistics are listed in Table 2. Density modification was performed with *DM* (Cowtan, 1994). Automatic tracing using *ARP/wARP* (Perrakis *et al.*, 1999) yielded 97% of the polypeptide model at 1.84 Å resolution. Manual completion of the model was carried out with *Coot* (Emsley & Cowtan, 2004). Final refinement with anisotropic *B* factors was carried out with data extending to 1.27 Å resolution using *REFMAC5* (Murshudov *et al.*, 2011). The quality of the refined protein model was validated using *MolProbity* (Chen *et al.*, 2010). Refinement statistics are listed in Table 3. All molecular graphics were prepared using *PyMOL* (Delano, 2002).

### 2.4. Docking studies

The docking studies were performed with *YASARA* (Krieger *et al.*, 2002) with default parameters using the polypeptide chains of the homodimer of our crystallographic



**Figure 1**

Occupancies of the heavy-atom sites found by *SHELXD* (Schneider & Sheldrick, 2002). Meaningful heavy atoms are labelled by locating their positions in the refined model.

**Table 2**  
Summary of SAD phasing.

<i>SHELXD</i>	
CC (all)	27.11
CC (weak)	16.93
PATFOM	3.68
<i>SHELXE</i>	
CC between $E_{\text{obs}}$ and $E_{\text{calc}}$	20.47
CC for partial structure against native data	47.86
FOM	0.725
MapCC	0.901
No. of residues built by <i>ARP/wARP</i>	626

**Table 3**  
Refinement statistics.

Resolution (Å)	50.0–1.27
Working/test reflections	208810/10991
$R/R_{\text{free}}^{\dagger}$ (%)	10.3/12.7
No. of protein residues	639
No. of water/glycerol molecules	994/2
No. of ions ( $\text{Cl}^-/\text{K}^+/\text{Na}^+$ )	4/4/2
R.m.s.d. from ideality	
Bond lengths (Å)	0.014
Bond angles (°)	1.842
Average <i>B</i> factors (Å <sup>2</sup> )	
Protein (5910 atoms)	14.7
Water (1031 atoms)	32.1
Others (96 atoms)	15.8
Ramachandran statistics‡ (%)	
Most favored regions	97.64
Outliers	0
PDB code	4chi

†  $R = \sum_{hkl} |||F_{\text{obs}}| - |F_{\text{calc}}||| / \sum_{hkl} |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factors, respectively.  $R_{\text{free}}$  is analogous to the *R* factor for 5% of the diffraction data excluded from refinement. ‡ Categories were defined by *MolProbity*.

model. The enantiomeric PLP adducts of (*R*)- and (*S*)- $\alpha$ -methylbenzylamine [*(R*)- $\alpha$ - and (*S*)- $\alpha$ -MBA] were generated in *YASARA* and energy minimization was performed to the lowest energy conformation. The completely flexible ligands were then alternatively docked into the active site. The chosen simulation cell was defined to be 18 × 17 × 18 Å around the catalytic residue Lys179. All residues of the active site and the active-site loop were included. H atoms were added in riding positions. The correct solution of the docking analysis was distinguished by the orientation of the cofactor PLP. The comparison of the docked enantiomeric PLP adducts with the PLP in the solved crystal structure led to the final assignment of the correct enantiomer.

## 3. Results

### 3.1. Structure analysis

The phasing contributions of the chloride ions and the S atoms of Cys and Met are shown in Fig. 1. Interestingly, the highest occupancy is observed for two chloride ions, but not for the other possible elements (S, P or K), even taking into account that these atoms show clear signals in the final anomalous electron-density map. The initial phasing based on the anomalous diffraction at 1.84 Å resolution was sufficient for automatic tracing. Refinement using the high-resolution

data converged to an  $R$  and  $R_{\text{free}}$  of 10.3 and 12.7%, respectively.

The final model contains 639 amino-acid residues of two polypeptide chains (*A* and *B*), two PLP molecules, four potassium ions, four chloride ions, two glycerol molecules and 994 water molecules. Ions are assigned according to electron density and meaningful chemical terms and refinement conditions. Both polypeptide chains are well defined by the electron-density maps (Supplementary Fig. S2<sup>1</sup>) and the final model is consistent with the anomalous map (Supplementary Fig. S3). Some residues with poor electron density at the N- and C-termini (monomer *A*, Met1 and Ser322–His332; monomer *B*, Met1–Ala2 and Ser322–His332) were excluded from the structural model. The cofactor PLP was modelled into the active site using  $F_{\text{obs}} - F_{\text{calc}}$  difference maps and refined with a summed occupancy of 0.8. The occupancies of the two PLP states were assigned so that the *B* factors were consistent with those of neighbouring residues. The remaining occupancy of 0.2 was filled with a single phosphate in the same position as the phosphate group of the cofactor.

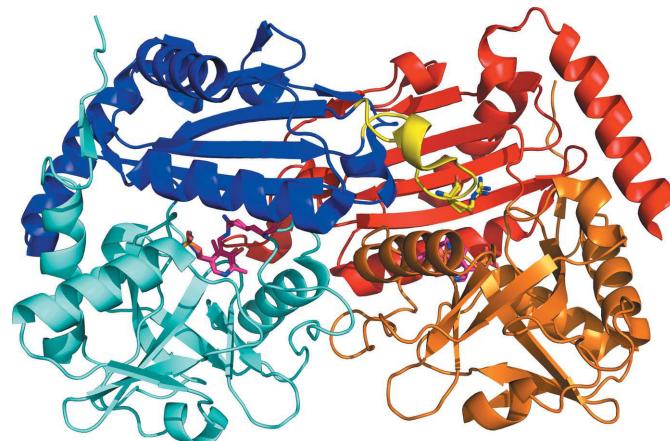
Additional positive difference electron density was observed in the substrate-binding site within covalent bond distance of the cofactor PLP in each monomer (Supplementary Fig. S4). Owing to the low occupancy of the ligand, we could not conclusively model this density. All compounds used in purification and crystallization and common metabolites were ruled out; also, GC-MS-analysis of acid-denatured and heat-denatured enzyme did not uncover the identity of this ligand. Besides the tricine molecule in the buffer, no amine or carbonyl compounds were added after cell disruption. Nevertheless, D-amino acids were also tested as possible ligands. In this case, the  $\alpha$ -carboxyl group could not be modelled into the small binding pocket.

Alternative conformations were modelled for 138 amino-acid side chains out of 639 residues (~20%). Some peptide backbone O atoms could also be modelled in alternative conformations. Differences in the main-chain conformation could be detected for residues Thr204–Gly206. The final refinement and validation statistics are shown in Table 3.

### 3.2. Overall fold

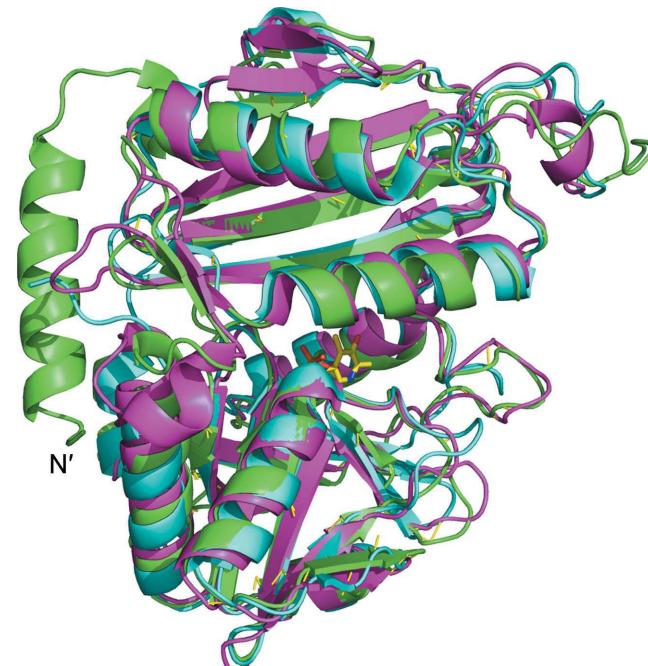
The (*R*)-selective amine transaminase crystallized in space group  $C222_1$  with two monomers in the asymmetric unit forming a homodimer (Fig. 2). Each polypeptide chain is constituted of 332 amino-acid residues with a molecular weight of 37.1 kDa. The tertiary structure of one subunit consists of the typical fold of enzymes belonging to the fold class IV of PLP-dependent enzymes, as first described for D-ATA from *Bacillus* sp. (Sugio *et al.*, 1995). The subunit divides into a small domain (N-terminus to Pro144) with an  $\alpha/\beta$ -structure, an inter-domain loop (Tyr145–Met149) and a large domain (Ala150 to the C-terminus) with a pseudo-barrel structure (Fig. 2). The enzyme belongs to fold class IV of PLP-dependent enzymes, and the overall structure is very similar to

those of BCATs and D-ATAs, with the best fit to the BCAT from *Thermus thermophilus* [PDB entry 1wrv; root-mean-square difference on  $C^\alpha$  atoms (r.m.s.d.) of 1.8 Å, fitting 297 residues; RIKEN Structural Genomics/Proteomics Initiative, unpublished work] and the D-ATA from *Bacillus* sp. YM-1



**Figure 2**

Overall structure of the (*R*)-selective amine transaminase from *A. fumigatus* viewed normal to the molecular dyad. The subunits of the homodimer are shown in blue and red, respectively. The monomer is divided into colour-coded domains: the small domain (blue, red) with the active-site loop (yellow) and the large domain (cyan, orange). The active-site loop derived from the left subunit is shown in yellow (with Arg126 as a stick model). The cofactor PLP is bound to Lys179 at the domain interface (shown as a stick model in pink).



**Figure 3**

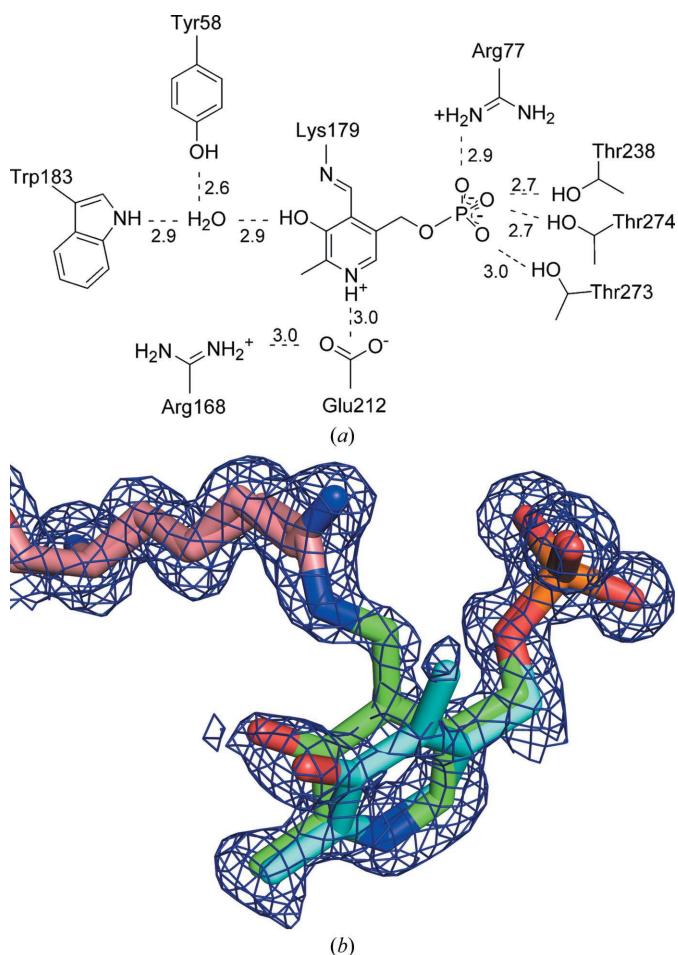
Comparison of the overall monomer fold between the (*R*)-selective amine transaminase from *A. fumigatus* (green), D-amino-acid aminotransferase (PDB entry 3lqs; r.m.s.d. 2.0 Å, cyan) and branched-chain amino-acid aminotransferase (PDB entry 1wrv; r.m.s.d. 1.8 Å, violet) distinctly shows the unique long N-terminal helix found in the (*R*)-selective amine transaminase. The cofactor PLP is shown as a stick model in yellow.

<sup>1</sup> Supporting information has been deposited in the IUCr electronic archive (Reference: DZ5319).

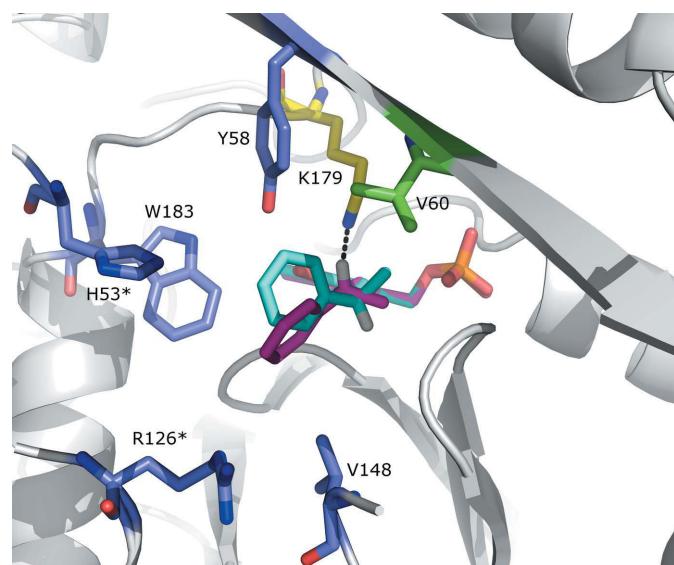
(PDB entry 3lqs; r.m.s.d. of 2.0 Å, fitting 280 residues; Lepore *et al.*, 2010). However, the (*R*)-selective amine transaminase from *A. fumigatus* has an additional long N-terminal  $\alpha$ -helix (Met4–Arg20) which has a significant effect on protein stability, as discussed below (Fig. 3). Other differences in the backbone folding compared with BCATs and d-ATAs are only visible for loop regions on the surface. Residues from each domain as well as residues from the other subunit of the dimer participate in forming the active site. The active-site loop (Gly121\*–Asn135\*; residues labelled with an asterisk belong to the other subunit) limits access to the active site and is contributed by the other subunit.

### 3.3. Cofactor binding

Well defined electron density is observed for the cofactor PLP, which is located at the bottom of the active site between the small and the large domains (Fig. 2). There are two distinct states observed for the cofactor. One state (occupancies of 0.5 and 0.4 in monomers *A* and *B*, respectively) is covalently bound to the active-site residue Lys179 of the large domain



(Fig. 4*a*), whereas the other state (occupancies of 0.3 and 0.4 in monomers *A* and *B*, respectively) represents an adduct with an unidentified ligand (see §3.1) and a free lysine (Fig. 4*b*). The covalently bound PLP shows the typical distorted aldimine bond of PLP-dependent enzymes. The bond angles deviate from the ideal 120° and the internal aldimine bond is out of the plane of the pyridoxyl ring by about 11.5 and 14° in monomers *A* and *B*, respectively. This typical geometry has been found in various crystal structures of other PLP-dependent enzymes and it is supposed that the release of strain on breaking the internal aldimine bond enhances the catalytic ability (Dubnovitsky *et al.*, 2005; Hayashi *et al.*, 1998). In the free state Lys179 has a distinct alternative conformation and is involved in a hydrogen-bonding network with Arg77 and the phosphate group of PLP via a water molecule. The pyridoxyl ring shows two distinct orientations, and the phosphate group is tightly bound and is involved in several hydrogen-bond interactions (with His74, Arg77, Thr273, Thr274, Ile237 and Thr238) as an anchor for the cofactor. Residues Ile237 and Thr238 are located at the N-terminus of helix  $\alpha$ 7 (according to Sugio *et al.*, 1995) such that the dipole moment of the helix additionally facilitates the coordination of the phosphate group. The pyridoxyl ring is sandwiched between Leu234 and the peptide bond of Gly215 to Phe216. The ion pair formed by the highly conserved Glu212 and the N atom of the pyridoxyl ring (N1) provides an electron sink during the reaction mechanism. This glutamate is in turn coordinated by the conserved Arg168. Anchoring of PLP by the phosphate and N1 coordination



limits movement of the pyridoxyl ring to a rotation around the C5–C5A bond by about 19° (Fig. 4b).

### 3.4. Deletion of the N-terminal helix

Previously, we performed crystallization studies on the (*R*)-amine transaminase from *Neosartorya fischeri* (96% identity to the amine transaminase from *A. fumigatus*). Presently, the diffraction images of the obtained crystals are not indexable. Based on a homology model built from D-amino-acid amino-

transferase (PDB entry 3daa; Peisach *et al.*, 1998) with a flexible N-terminus, an N-terminal deletion of 22 amino-acid residues was introduced to improve the crystallization quality. Whereas the wild-type amine transaminase could be over-expressed as soluble protein, this N-terminal deletion resulted in insoluble protein. The insolubility of this amine transaminase variant could not be prevented by changing the expression temperature, varying the inducer concentration, altering the induction time or the co-expression of chaperones.

In the solved X-ray structure of (*R*)-amine transaminase from *A. fumigatus*, the N-terminus forms a long helix corresponding to the sequence that was deleted in the (*R*)-amine transaminase mutant from *N. fischeri*. It is obvious that the N-terminal helix is important for the soluble expression of fungal amine transaminases, but remarkable hydrophobic patches on the surface of the modelled truncated enzyme are not observed.

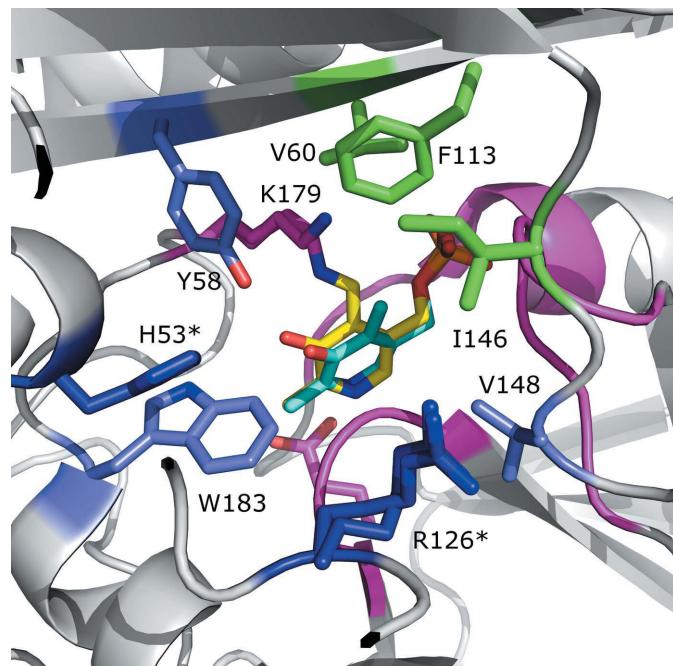
### 3.5. Structural design of the active site

Via docking studies performed with the program YASARA, it could be demonstrated that the active site of the (*R*)-selective amine transaminase from *A. fumigatus* is divided into a small and a large binding pocket (Fig. 5). Docking was performed with the substrate adducts of (*R*)- $\alpha$ -methylbenzylamine [(*R*)- $\alpha$ -MBA] and (*S*)- $\alpha$ -MBA to PLP, which starts the first deamination cycle. The pyridoxyl rings and the phosphate group of the modelled substrate adducts superposed very well with the free PLP state (r.m.s.d. of 0.14 Å) of the X-ray structure, indicating good quality of the docking results. In every docking run the methyl group was bound in the small binding pocket formed by the residues Val60, Phe113 and Ile146 (Fig. 6). The aromatic ring was coordinated in the large binding pocket which is built by the residues His53\*, Tyr58, Arg126\*, Val148 and Trp183. Although (*S*)- $\alpha$ -MBA-PLP could be docked without clashes, the enantioselectivity can be explained by the orientation of (*R*)- $\alpha$ -MBA-PLP and (*S*)- $\alpha$ -MBA-PLP to the catalytically active lysine residue. This

residue initiates the deamination reaction by deprotonation and is only at a reasonable distance (2.8 Å) for abstraction of the proton from (*R*)- $\alpha$ -MBA-PLP which points directly towards the Lys179 N<sup>c</sup> atom. In contrast, the proton of (*S*)- $\alpha$ -MBA-PLP points in the opposite direction and cannot be abstracted by Lys179.

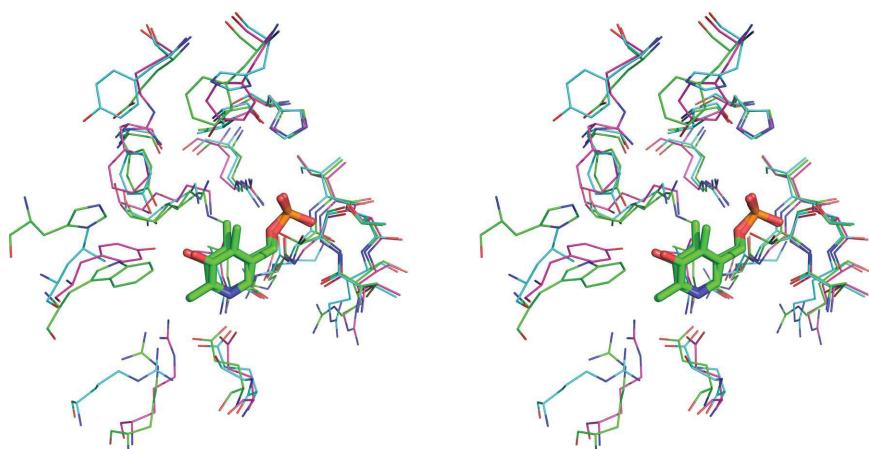
### 3.6. Active-site comparison

Whereas the binding of the cofactor and the backbone are conserved, a comparison of the active-site residues responsible for substrate recognition of the (*R*)-selective amine transaminases (*R*-ATAs) with the active sites of BCATs and D-ATAs shows clearly that no amino-acid residues other than the catalytic Lys179 and Glu212 are conserved (Fig. 7). When considering



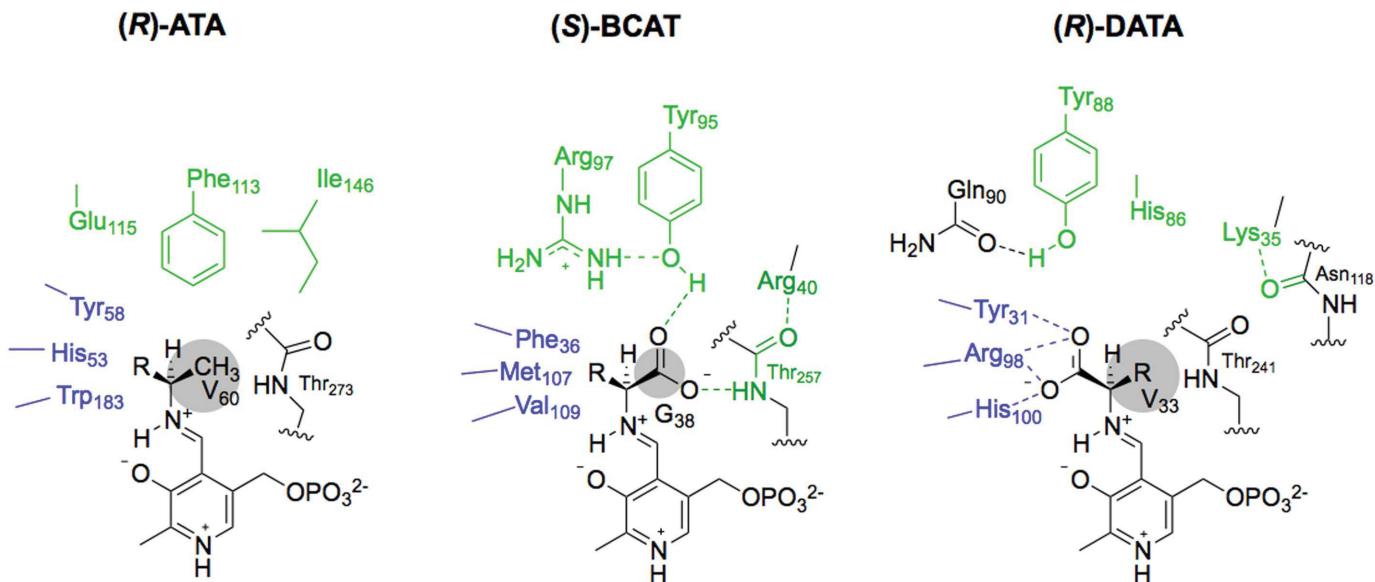
**Figure 6**

Active-site architecture. The residues forming the small binding pocket are shown in green and the amino-acid residues responsible for forming the large binding pocket are shown in blue (residues which originate from the other subunit are shown in dark blue and are marked with asterisks; PLP-binding residues are coloured violet).



**Figure 7**

Stereo representation of the active-site comparison between the (*R*)-selective amine transaminase from *A. fumigatus* (green), D-amino acid aminotransferase (PDB entry 3lqs; cyan) and branched-chain amino-acid aminotransferase (PDB entry 1wrv; violet). The inter-domain loop is omitted for clarity.

**Figure 8**

Schematic drawing of the active-site arrangement of enzymes of fold class IV in comparison to the active site of the (*R*)-selective amine transaminase from *A. fumigatus* [(*R*)-ATA], the (*S*)-branched-chain amino-acid aminotransferase [(*S*)-BCAT; PDB entry 1iye] and D-amino-acid aminotransferase [(*R*)-DATA; PDB entry 3daa]. Grey spheres indicate the space-filling requirements of residues Val60, Gly38 and Val33, respectively.

substrates to be converted by (*R*)-selective amine transaminases, the substitution of the carboxyl group by a methyl group inverts the priority according to the Cahn–Ingold–Prelog rule. Hence, the active-site architecture, based on the definition of the small and large pockets, was postulated to be more similar to the BCATs than to the D-ATAs. This can now be verified by the crystal structure with the docking analysis (see §3.5) as well as the sequence motif Y/(F)VZ (with Z preferably being glutamate) postulated to be an important feature in the structural design of the active site of (*R*)-selective amine transaminases (Höhne *et al.*, 2010). Whereas the large pocket of the D-ATAs is mostly built by small hydrophobic residues, the pocket volume is reduced in BCATs and in the (*R*)-selective amine transaminase by bulky amino-acid residues (Fig. 8). A search of the DALI secondary-structure database (Holm & Rosenström, 2010) revealed the inter-domain loop as an active-site limiting feature. The loop in question is two amino-acid residues longer than the equivalent loop found in D-ATAs and therefore contributes to the small binding pocket of the active site. Whereas the small pocket of the D-ATAs harbours predominantly positively charged residues to coordinate the carboxylate group, in BCATs and *R*-ATAs aromatic side chains form a hydrophobic environment for the mostly hydrophobic substituents that are accepted. As mentioned above, the entrance of the active site of the *R*-ATAs is limited by the active-site loop. Interestingly, similar to dual substrate recognition by the (*S*)-amine transaminase (Steffen-Munsberg *et al.*, 2013), the active-site loop of the *R*-ATA also has a highly flexible Arg126 (slightly increased *B* factors and two alternative conformations of residues Arg126, Gly127 and Ser128). By flipping in and out of the active site, Arg126 could facilitate the coordination of the negatively charged carboxylate of the amino acceptor pyruvate as well as the binding of uncharged

substrate in the same pocket of the active site. This assumption needs to be investigated further *via* mutagenesis.

#### 4. Conclusion

In the era of rational protein design, crystal structure or NMR analyses at atomic resolution are the most valuable tools for protein-engineering experiments. The crystal structure of the (*R*)-selective amine transaminase elucidated here for the enzyme from *A. fumigatus* provides essential information and insights into understanding how substrate recognition occurs in (*R*)-selective amine transaminases and distinguishes them from other enzymes of fold class IV. For further understanding of substrate binding and enantioselectivity, soaking and co-crystallization experiments are in progress.

*Note added in proof.* A crystal structure analysis of the (*R*)-selective  $\omega$ -transaminase from *Aspergillus terreus* has recently been published (Lyskowski *et al.*, 2014).

MT thanks the ‘Landesgraduiertenkolleg Mecklenburg-Vorpommern’ for financial support. We thank the European Union (KBBE-2011-5, grant No. 289350) for financial support within the European Union Seventh Framework Programme. Diffraction data were collected on BL14.1 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron-storage ring (Berlin-Adlershof, Germany). We also thank Professor M. Lalk (University Greifswald, Germany) for the GC-MS analysis.

#### References

- Blaser, H.-U. (2002). *Adv. Synth. Catal.* **344**, 17.

- 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. (2010). *Acta Cryst. D* **66**, 12–21.
- Christen, P. & Mehta, P. K. (2001). *Chem. Rec.* **1**, 436–447.
- Cowtan, K. (1994). *Jnt CCP4/ESF-EACBM Newslet. Protein Crystallogr.* **31**, 34–38.
- Delano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Dubnovitsky, A. P., Ravelli, R. B. G., Popov, A. N. & Papageorgiou, A. C. (2005). *Protein Sci.* **14**, 1498–1507.
- Eliot, A. C. & Kirsch, J. F. (2004). *Annu. Rev. Biochem.* **73**, 383–415.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst. D* **60**, 2126–2132.
- Fuchs, M., Koszelewski, D., Tauber, K., Kroutil, W. & Faber, K. (2010). *Chem. Commun.* **46**, 5500.
- Han, Q., Robinson, H., Gao, Y. G., Vogelaar, N., Wilson, S. R., Rizzi, M. & Li, J. (2006). *J. Biol. Chem.* **281**, 37175–37182.
- Hanson, K. R. (1966). *J. Am. Chem. Soc.* **88**, 2731–2742.
- Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.
- Hayashi, H. (1995). *J. Biochem.* **118**, 463–473.
- Hayashi, H., Mizuguchi, H. & Kagamiyama, H. (1998). *Biochemistry*, **37**, 15076–15085.
- Höhne, M. & Bornscheuer, U. T. (2012). *Enzymes in Organic Synthesis*, edited by O. May, H. Gröger & W. Drauz, pp. 779–820. Weinheim: Wiley-VCH.
- Höhne, M., Schätzle, S., Jochens, H., Robins, K. & Bornscheuer, U. T. (2010). *Nature Chem. Biol.* **6**, 807–813.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Humble, M. S., Cassimjee, K. E., Håkansson, M., Kimbung, Y. R., Walse, B., Abedi, V., Federsel, H. J., Berglund, P. & Logan, D. T. (2012). *FEBS J.* **279**, 779–792.
- Jansonius, J. N. (1998). *Curr. Opin. Struct. Biol.* **8**, 759–769.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Koszelewski, D., Clay, D., Rozzell, D. & Kroutil, W. (2009). *Eur. J. Org. Chem.* **2009**, 2289–2292.
- Koszelewski, D., Pressnitz, D., Clay, D. & Kroutil, W. (2009). *Org. Lett.* **11**, 4810–4812.
- Koszelewski, D., Tauber, K., Faber, K. & Kroutil, W. (2010). *Trends Biotechnol.* **28**, 324–332.
- Krieger, E., Koraimann, G. & Vriend, G. (2002). *Proteins*, **47**, 393–402.
- Kroutil, W., Fischereder, E. M., Fuchs, C. S., Lechner, H., Mutti, F. G., Pressnitz, D., Rajagopalan, A., Sattler, J. H., Simon, R. C. & Sirola, E. (2013). *Org. Process Res. Dev.* **17**, 751–759.
- Krug, M., Weiss, M. S., Heinemann, U. & Mueller, U. (2012). *J. Appl. Cryst.* **45**, 568–572.
- Lepore, B. W., Liu, D., Peng, Y., Fu, M., Yasuda, C., Manning, J. M., Silverman, R. B. & Ringe, D. (2010). *Biochemistry*, **49**, 3138–3147.
- Lyskowski, A., Gruber, C., Steinkellner, G., Schurmann, M., Schwab, H., Gruber, K. & Steiner, K. (2014). *PLoS One*, **9**.
- Malik, M. S., Park, E.-S. & Shin, J.-S. (2012). *Appl. Microbiol. Biotechnol.* **94**, 1163–1171.
- Martens, J. G. K. & Schickedanz, M. (1986). *Arch. Pharm.* **319**, 461–465.
- Mathew, S. & Yun, H. (2012). *ACS Catal.* **2**, 993–1001.
- Midelfort, K. S., Kumar, R., Han, S., Karmilowicz, M. J., McConnell, K., Gehlhaar, D. K., Mistry, A., Chang, J. S., Anderson, M., Villalobos, A., Minshull, J., Govindarajan, S. & Wong, J. W. (2013). *Protein Eng. Des. Sel.* **26**, 25–33.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Ogawa, K., Koyama, Y., Ohashi, I., Sato, I. & Hirama, M. (2009). *Angew. Chem. Int. Ed.* **48**, 1110–1113.
- Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst. D* **61**, 449–457.
- Peisach, D., Chipman, D. M., Van Ophem, P. W., Manning, J. M. & Ringe, D. (1998). *Biochemistry*, **37**, 4958–4967.
- Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Mol. Biol.* **6**, 5.
- Rösler, M., Anand, R., Cicin-Sain, A., Gauthier, S., Agid, Y., Dal-Bianco, P., Stähelin, H. B., Hartman, R. & Gharabawi, M. (1999). *BMJ*, **318**, 633–638.
- Rudat, J., Brucher, B. R. & Syldatk, C. (2012). *AMB Express*, **2**, 11.
- Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W. & Hughes, G. J. (2010). *Science*, **329**, 305–309.
- Sayer, C., Isupov, M. N., Westlake, A. & Littlechild, J. A. (2013). *Acta Cryst. D* **69**, 564–576.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst. D* **58**, 1772–1779.
- Schrewe, M., Ladkau, N., Bühler, B. & Schmid, A. (2013). *Adv. Synth. Catal.* **355**, 1693–1697.
- Schwarzenbacher, R. et al. (2004). *Proteins*, **55**, 759–763.
- Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
- Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Uson, I. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 333–351. Dordrecht: Kluwer Academic Publishers.
- Steffen-Munsberg, F., Vickers, C., Thontowi, A., Schätzle, S., Tumliorsch, T., Svedendahl Humble, M., Land, H., Berglund, P., Bornscheuer, U. T. & Höhne, M. (2013). *ChemCatChem*, **5**, 150–153.
- Sugio, S., Petsko, G. A., Manning, J. M., Soda, K. & Ringe, D. (1995). *Biochemistry*, **34**, 9661–9669.
- Sutin, L., Andersson, S., Bergquist, L., Castro, V. M., Danielsson, E., James, S., Henriksson, M., Johansson, L., Kaiser, C., Flyrén, K. & Williams, M. (2007). *Bioorg. Med. Chem. Lett.* **17**, 4837–4840.
- Thomsen, M., Skalden, L., Palm, G. J., Höhne, M., Bornscheuer, U. T. & Hinrichs, W. (2013). *Acta Cryst. F* **69**, 1415–1417.
- Tufvesson, P., Lima-Ramos, J., Jensen, J. S., Al-Haque, N., Neto, W. & Woodley, J. M. (2011). *Biotechnol. Bioeng.* **108**, 1479–1493.

## Supplement

# Crystallographic characterisation of the (R)-selective amine transaminase from *Aspergillus fumigatus*

Maren Thomsen‡, Lilly Skalden‡, Gottfried J. Palm, Matthias Höhne, Uwe T. Bornscheuer and Winfried Hinrichs

‡ These authors contributed equally to this work.

Institute of Biochemistry, University of Greifswald, Felix-Hausdorff-Str. 4, Greifswald, Germany

Correspondence e-mail: [winfried.hinrichs@uni-greifswald.de](mailto:winfried.hinrichs@uni-greifswald.de)

### Cloning of the amine transaminase mutant of *Neosartorya fischeri*

To delete 22 amino acids at the N-terminus an additional *Nde*I restriction site was inserted into the sequence of the amine transaminase of *Neosartorya fischeri* after the second amino acid alanine. The insertion was made by a QuikChange PCR with the Quik Lightning Multi Enzyme kit (Agilent Technologies). The QuikChange was done as described in the manual with the forward primer CTGCTGGAACGTAGCCATATGGCGTTCTCTAAAGGTATTG. The following PCR program was used: Hold 95°C for 2 min before 30 cycles with 95°C for 20 sec, 55°C for 30 sec and 65°C for 5 min started. Finally 65°C was hold additional for 5 min. After the QuikChange PCR a *Dpn*I digestion from the same kit followed. 1 µl *Dpn*I was mixed with 25 µl PCR reagent and incubated for 5 min at 37°C before the enzyme was inactivated by 80°C for 20 min. The resulting plasmid was verified through sequencing. The final deletion of the 22 amino acids was done through a restriction digestion with *Nde*I enzyme (New England Biolabs). 20 µl plasmid was mixed with 7 µl NEB-buffer 4, 3 µl *Nde*I enzyme and 40 µl dist. water. The digestion was made at 37°C for 2 h before the enzyme was inactivated for 10 min at 80°C. A PCR purification with the High-Pure PCR Cleanup Micro kit (Roche) followed. The purification was done as described in the manual. Finally the ligation of the plasmid was carried out with the T4-ligase (Thermo Scientific). 5.8 µl purified PCR product was mixed with 0.7 µl T4-ligase buffer, 0.5 µl T4-ligase and 1 µl dist. water. The following program was used: Hold 20°C for 2 h, hold 16°C for 4 h, hold 14°C for 3 h, hold 12°C for 3 h, hold 10°C for 2 h and hold 72°C for 10 min. The resulting plasmid was verified by sequencing.

### Expression of the wild type and the mutant of *Neosartorya fischeri*

The expression of the amine transaminase mutant and wild type from *Neosartorya fischeri* was carried out like the expression of the amine transaminase from *Aspergillus fumigatus* as described (Höhne *et al.*, 2010). To optimise the expression of the mutant, the same expression protocol was used with exception of the following varied parameter. The expressions at different temperatures were additionally carried out at 15°C, 25°C and 30°C. The expression at different induction times were also tested with an induction OD<sub>600</sub> of 1.5 and 5. Also different inducer concentrations of 0.1 mM and 0.5 mM IPTG were investigated. The expression optimisation was carried out with chaperones present in the Takara chaperone kit (Takara Bio INC.). The plasmids carrying the chaperones were co-transformed with the amine transaminase gene carrying plasmid. During all cultivations and expressions 7/OD<sub>600</sub> samples were taken.

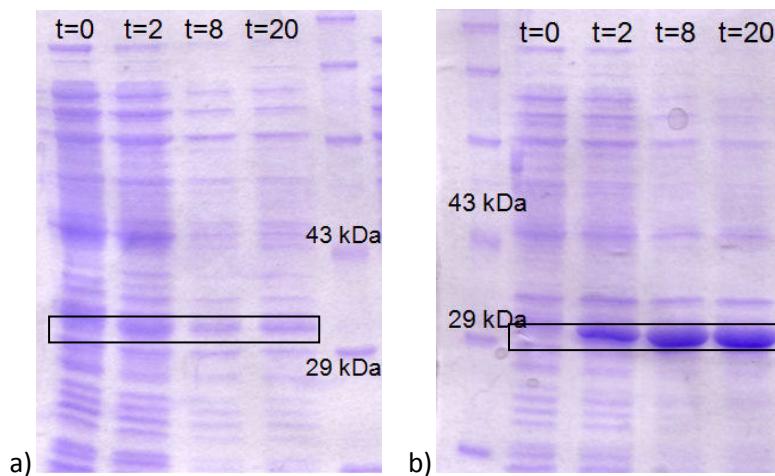
### Transformation and co-transformation of chaperones

Competent *E. coli* BL21 cells were incubated for 30 min on ice with 1  $\mu$ l of the plasmid. Afterwards a heat shock at 42°C for 45 sec followed. After 5 min incubation on ice, 250  $\mu$ l LB-Soc was added and incubated at 37°C for 1 h. Finally, 150  $\mu$ l culture was plated on agar plates with specific antibiotic resistance (ampicillin for WT and mutant of the amine transaminase of *Neosartorya fischeri*). The co-transformation was done with in each case 1  $\mu$ l of the plasmid carrying the chaperones and the amine transaminase encoding genes. The resulting culture was plated on agar plates with ampicillin (100  $\mu$ g/ml) and chloramphenicol (50  $\mu$ g/ml).

### SDS-PAGE

The 7/OD<sub>600</sub> samples were disrupted by Fast Prep with 4  $\text{ms}^{-1}$  for 20 sec two times. After centrifugation at 13000g the supernatant was separated and the insoluble fraction was washed with sodium phosphate buffer (50 mM, pH 7.5). Finally, the insoluble fraction was re-suspended in sodium phosphate buffer (50 mM, pH 7.5).

SDS-PAGEs with a 10% separation gel and 4% collecting gel was used. 15  $\mu$ l normalized cultivation sample was mixed with 15  $\mu$ l sample buffer before a denaturation at 95°C for 10 min happened. 20  $\mu$ l prepared sample was loaded to the SDS-PAGE. As protein marker the commercial available protein marker from Carl Roth (200 kDa to 14 kDa) was used. With 200V and 50 mM the SDS-PAGE run for 45 min before it was stained with coomassie blue. After destaining the gels, the protein bands were visible.



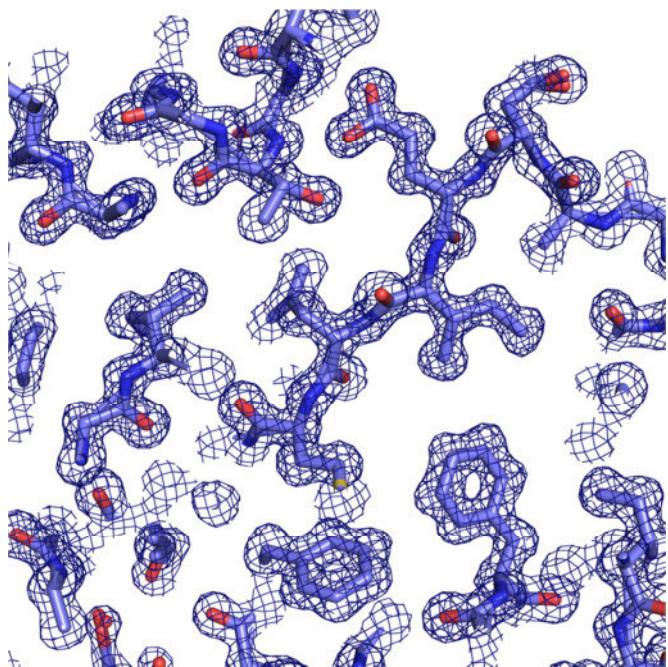
**SI-Figure 1:** SDS-PAGE of the amine transaminases of *Neosartorya fischeri*, wild type and mutant.  
a) Soluble fraction b) insoluble fraction

**SI-Table 1:** Data-collection and processing statistics.

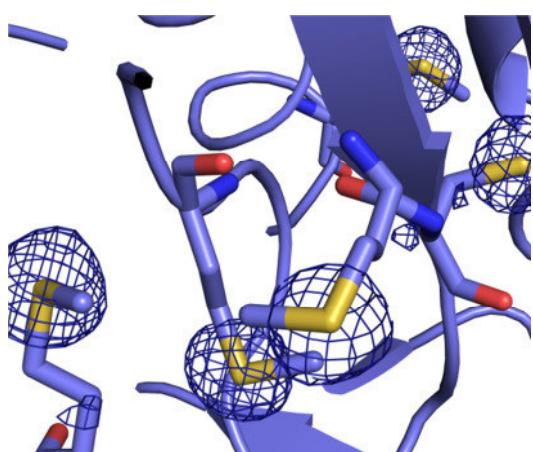
Values in parentheses are for the outermost resolution shell.

	Native	Anomalous
Beamline	BESSY II/14.1	BESSY II/14.1
Detector	PILATUS 6M	PILATUS 6M
Wavelength (Å)	0.91841	1.77122
Temperature (K)	100	100
Orthorhombic space group	$C222_1$	$C222_1$
Unit-cell parameters (Å)		
$a/b/c$	102.2/120.9/135.4	102.2/120.9/135.4
Resolution range (Å)	50.0 – 1.27 (1.35 – 1.27)	50.0 – 1.84 (1.95 – 1.84)
No. of unique reflections	426722 (68273)	135117 (17260)
Multiplicity	3.38 (3.3)	5.6 (2.6)
$R_{\text{merge}}$ (%)	6.3 (60.3)	3.9 (9.5)
Mean $I/\sigma(I)$	13.2 (2.0)	30.29 (8.34)
$\text{CC}_{1/2}$ (%) †	99.9 (73.0)	99.9 (98.8)
Completeness (%)	99.1 (97.9)	95.9 (75.8)
Overall $B$ factor from Wilson plot (Å <sup>2</sup> )	17.4	18.8
Total rotation, increment (°)	180, 0.1	360, 0.1

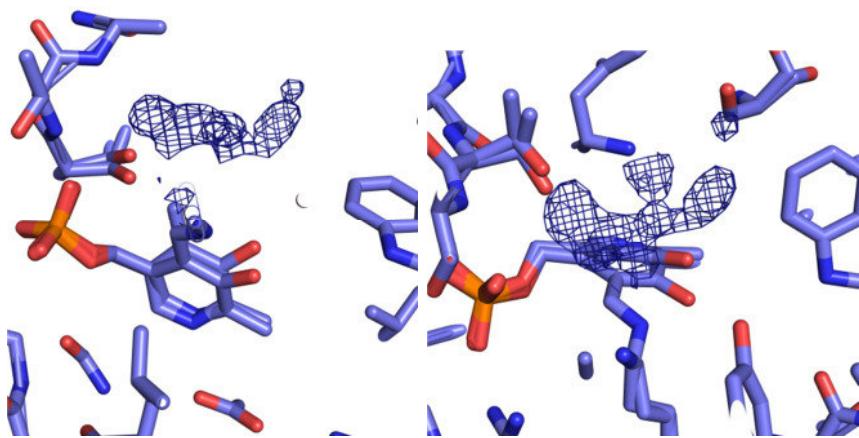
†  $\text{CC}_{1/2}$  is the percentage correlation between intensities from random half data sets (Karplus & Diederichs, 2012).



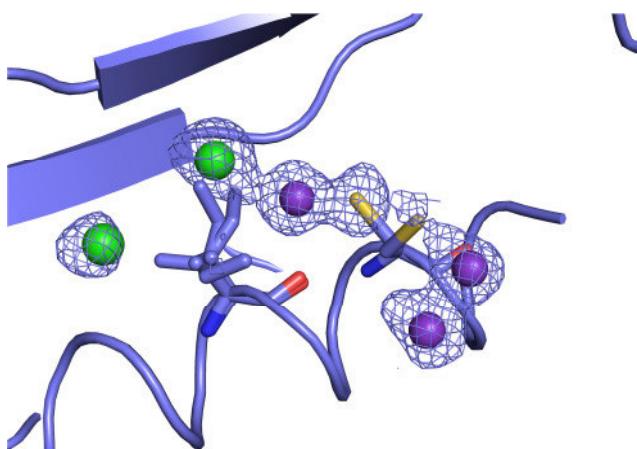
**SI-Figure 2:**  $2F_{\text{obs}} - F_{\text{calc}}$  difference map of a representative region of the enzyme calculated at  $1.27 \text{ \AA}$  resolution contoured at  $1\sigma$  level.



**SI-Figure 3:** Anomalous map contoured at  $3\sigma$ .



**SI-Figure 4:**  $F_{\text{obs}} - F_{\text{calc}}$  difference maps of additional positive difference electron density at  $3\sigma$  level.



**SI-Figure 5:** Anomalous map contoured at  $3\sigma$  around Ile78 and Cys82, ions are shown as spheres (potassium: violet; chloride: green).

## References

- Höhne, M., Schätzle, S., Jochens, H., Robins, K., Bornscheuer, U. T. (2010). *Nat. Chem. Biol.* **6**, 807-813.  
Karplus, P. A., Diederichs, K. (2012). *Science* **336**, 1030-1033.

# **Publikation V**

# Structural and biochemical characterization of the dual substrate recognition of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*

Lilly Skalden, Maren Thomsen, Matthias Höhne, Uwe T. Bornscheuer and Winfried Hinrichs

Institut für Biochemie, Universität Greifswald, Germany

## Keywords

amine transaminase; dual substrate recognition; gabaculine; pyridoxal-5'-phosphate; X-ray structure

## Correspondence

W. Hinrichs, Institut für Biochemie, Universität Greifswald, Felix-Hausdorff-Straße 4, 17489 Greifswald, Germany  
Fax: +49 0 3834 864373  
Tel.: +49 0 3834 864356  
E-mail: [winfried.hinrichs@uni-greifswald.de](mailto:winfried.hinrichs@uni-greifswald.de)  
U. T. Bornscheuer, Institut für Biochemie, Universität Greifswald, Felix-Hausdorff-Straße 4, 17489 Greifswald, Germany  
Fax: +49 0 3834 864373  
Tel.: +49 0 3834 864367  
E-mail: [uwe.bornscheuer@uni-greifswald.de](mailto:uwe.bornscheuer@uni-greifswald.de)

Lilly Skalden and Maren Thomsen contributed equally to this work

(Received 5 September 2014, revised 23 October 2014, accepted 13 November 2014)

doi:10.1111/febs.13149

## Introduction

Amines are important building blocks for a range of pharmaceuticals, fine-chemicals and agrochemicals [1]. Although several chemical methods to produce chiral amines have been developed, biocatalytic routes have become increasingly important. In the past few years, many amine transaminases (ATA) with complementary enantio-preferences have been discovered [1–3] and protein engineering has enabled substantial alteration of the substrate scope, as

Chiral amines are important precursors for the pharmaceutical and fine-chemical industries. Because of this, the demand for enantiopure amines is currently increasing. Amine transaminases can produce a large spectrum of chiral amines in the (*R*)- or (*S*)-configuration, depending on their substrate scope and stereo-preference, by converting a prochiral ketone into the chiral amine while using alanine as the amine donor producing pyruvate as an  $\alpha$ -keto acid product. In order to guide the protein engineering of transaminases to improve substrate specificity and enantioselectivity, we carried out a crystal structure analysis at 1.6 Å resolution of the (*R*)-amine transaminase from *Aspergillus fumigatus* with the bound inhibitor gabaculine. This revealed that Arg126 has an important role in the dual substrate recognition of this enzyme because mutating this residue to alanine reduced substantially the ability of the enzyme to use pyruvate as an amino acceptor.

## Database

Coordinates and structure factors have been deposited with the Protein Data Bank under accession code [4UUG](#).

shown for the large-scale synthesis of Sitagliptin [4]. These enzymes are the most suitable for making chiral amines because they can form optically pure amines from a racemic mixture at kinetic resolutions or, in the asymmetric synthesis mode, from ketones usually using alanine/pyruvate as the donor/acceptor pair. Various improvements in asymmetric synthesis have been made to increase the yield of amine formation [5,6].

## Abbreviations

ATA, amine transaminase; *m*CPP, *m*-carboxyphenylpyridoxamine phosphate; PDB, Protein Data Bank; PLP, pyridoxal-5'-phosphate.

Depending upon their substrate scope, transaminases can be divided into  $\alpha$ -,  $\omega$ - and amine transaminases [7].  $\alpha$ -Transaminases convert substrates with the amino group in an alpha position to the carboxylate.  $\omega$ -Transaminases may have several carbon atoms between the carboxylate and the terminal amino group [8]. ATAs are able to convert ketones directly, so a carboxylate is not required. All transaminases are pyridoxal-5'-phosphate (PLP)-dependent enzymes that transfer an amino group from an amino donor to an amino receptor. In the seven different PLP-depending fold classes, transaminases are located in fold classes I and IV [9–12].

Transamination takes place in two half reactions. In the first half reaction, the amino group is transferred from an amino donor to the cofactor PLP via several intermediates. The corresponding ketone or  $\alpha$ -keto acid of the amino donor is released and pyridoxamine-5'-phosphate is formed. In the second half-reaction, the amino group is transferred from pyridoxamine-5'-phosphate via the same types of intermediates in the reverse order to the amino acceptor. This is converted into the amine or amino acid and PLP is regenerated [9,13].

$\alpha$ -Transaminases commonly use amino acids like aspartate or alanine as an amino donor [14]. Even if alanine is not the preferred amino donor,  $\omega$ -transaminases, as well as ATAs, can use alanine as the donor [15,16]. Dual recognition is the ability of the transaminase to use the same substrate-binding pocket by accepting hydrophobic (e.g. phenyl group) as well as hydrophilic (e.g. carboxylic group) substituents [17]. This phenomenon has been described for different types of transaminases and has been reported for branched-chain amino acid transaminases and (S)-selective ATAs. Branched-chain amino acid transaminases, which belong to fold class IV, use hydrophilic residues to coordinate the carboxylate at the border of the large binding pocket, but the main hydrophobic character of this pocket is maintained [2,17]. Dual substrate recognitions have also been reported in fold class I, where (S)-ATAs and aromatic amino acid aminotransferases are grouped. For the aromatic amino acid aminotransferase from *Paracoccus denitrificans*, loop rearrangements with conformational changes of side chains in the active site caused by the type of substrate are known [17]. In the (S)-ATA from *Silicibacter pomeroyi* [Protein Data Bank (PDB) code: [3HMU](#)] movement of the amino acid residue Arg417 has been described [18]. Depending on the substrate, either L-alanine or  $\alpha$ -methylbenzylamine ( $\alpha$ -MBA) is bound, and Arg417 flips in or out of the active site to constitute hydrogen bonding with the carboxylate of L-alanine.

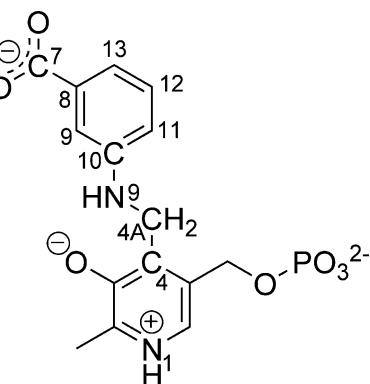
(R)-ATAs, which belong to fold class IV, also show dual substrate recognition, and can use  $\alpha$ -MBA as well

as D-alanine as an amino donor, similarly to (S)-selective ATAs [15,19].

To investigate this dual substrate recognition in detail for (R)-ATA, a thorough knowledge of the protein structure is necessary. Whereas structures of  $\alpha$ -transaminases and (S)-ATA are well documented in the literature and the protein structure database [18,20,21], until recently, the structures of (R)-ATA were missing. Homology models [4] and substrate scope studies [15] provided a hint of the overall fold and the active site, but not a reliable explanation of dual substrate recognition.

Recently, we reported the crystal structure of the (R)-ATA from *Aspergillus fumigatus* (PDB code: [4CHI](#)) [22]. Two additional structures have been published in the meantime for the (R)-ATAs from *Aspergillus terreus* (PDB code: [4CE5](#)) [23] and *Nectria haematococca* (PDB codes: [4CMD](#) and [4CMF](#)) [24]. Detailed analysis of the protein structures – especially to understand dual substrate recognition – requires inhibitors such as gabaculine to be bound to the active site. Gabaculine (5-amino-1,3-cyclohexadienylcarboxylic acid) is a neurotoxic natural product from *Streptomyces toyocaensis* [25] and is known to be a covalent inhibitor for transaminases. During the transamination of gabaculine, a cyclohexatrienyl system, which is bound to the cofactor PLP, is formed. After spontaneous aromatization, *m*-carboxyphenylpyridoxamine phosphate (*m*CPP, Scheme 1) is formed and this finally inhibits the enzyme irreversibly [26].

To date, the substrate recognition and selectivity of (R)-ATAs have been discussed based on the structural determination of a gabaculine complex (PDB code: [4CMF](#)) [24], but only in comparison with structural and mutational data for a related enzyme (PDB code: [4CE5](#)) [23]. These studies resulted in different models of substrate recognition.



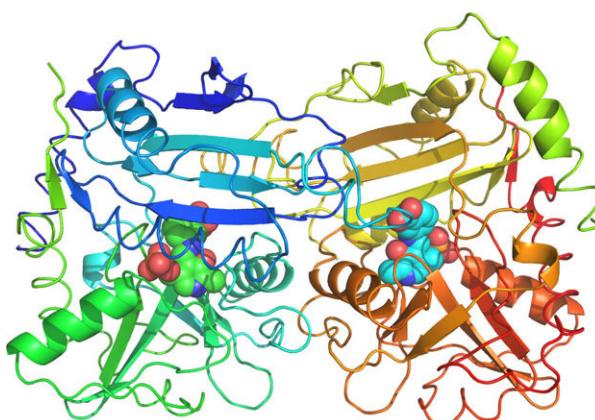
**Scheme 1.** Structure of *m*-carboxyphenylpyridoxamine phosphate (*m*CPP).

Here, we report the crystal structure of the (*R*)-ATA from *A. fumigatus* with the bound inhibitor gabaculine, as well as mutational studies of the same enzyme to clarify the dual substrate recognition phenomenon for (*R*)-ATAs.

## Results

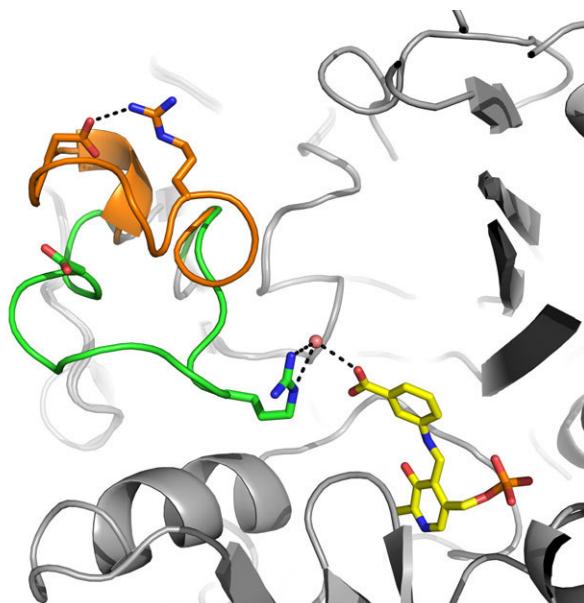
### Inhibitor binding to the active site

Crystal structure analysis of the ATA from *A. fumigatus* with bound gabaculine was carried out to obtain detailed insight into substrate recognition. Crystals of the native enzyme crystallized in the trigonal space group  $P\bar{3}_121$  and were soaked with gabaculine to form the covalent inhibitor-adduct *mCPP*. The resulting crystal structure could be solved and refined to a resolution of 1.6 Å. The asymmetric unit contains two monomers forming a homodimer (Fig. 1) displaying a fold typical of class IV PLP-dependent enzymes. Each monomer contributes to the active site of the other by providing an active-site loop (Gly121–Asn135). Monomer B with the active-site loop of monomer A forms a closed active-site loop conformation, whereas monomer A with the loop contributed by monomer B is predominantly in an open conformation (occupancy of 0.7). However, some residues (besides Gly127–Asp132) of the active-site loop of monomer B can be traced in the electron-density maps with a low occupancy of 0.3 in the closed conformation. Superposition of these different loop conformations of the monomers showed that the main difference in the backbone folding is restricted to the active-site loop only (rmsd on C $\alpha$ : 1.5 Å; Fig. 2).

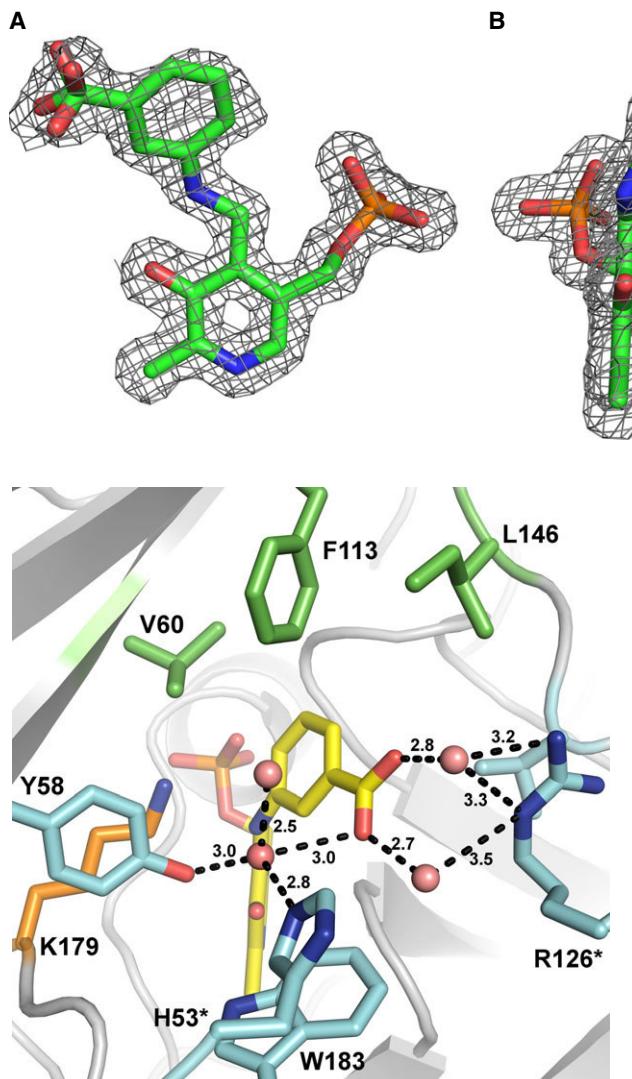


**Fig. 1.** Overall presentation of the (*R*)-ATA from *Aspergillus fumigatus* with *mCPP* atoms shown as van der Waal's spheres (color code: carbon, yellow; oxygen, red; nitrogen, blue).

The inhibitor is clearly bound in both monomers with full occupancy in monomer B and an occupancy of 0.8 for the benzoic acid moiety in monomer A, owing to minor conformational disorder. However, they display the same overall orientation in the active site (rmsd on all atoms: 0.09 Å). These two states differ only in the orientation of the carboxylate (C7), which rotates in the open form (Fig. 3). Both states are in agreement with the previously reported classification of the small and the large binding pockets [22]. The C-atoms C11–C12 of the benzoic acid moiety are surrounded by the small binding pocket which is formed by the hydrophobic residues Val60, Phe113 and Leu146, whereas the C-atoms C8, C9, C13 and the steric-demanding carboxylate are coordinated by His53\*, Arg126\* (\*both of the adjacent monomer) and Tyr58; Val148 and Trp183 compose the large binding pocket. The carboxylate of *mCPP* is coordinated by residues His53\* and Tyr58 of the large binding pocket via a water molecule (Fig. 4). The proposed responsible residue Arg126\* for the dual substrate recognition has a direct distance of 3.8 Å to the carboxylate of *mCPP*. However, here, coordination is also facilitated via a water molecule (Fig. 4).



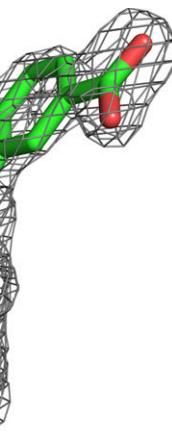
**Fig. 2.** Comparison of the open and closed active-site loop conformation by superposition of monomers A and B. In the closed loop conformation (green) R126 forms a water-mediated salt bridge with the carboxylate of *mCPP* (shown as stick model in yellow). The open loop conformation (orange) is stabilized by an alternative salt bridge of Arg126 with the carboxylate of Asp132 (R126-N $\eta$ 1–D132-O $\delta$  2.1 Å). The C $\alpha$  position of Arg126 is shifted by ~ 11.2 Å. Hydrogen bonds are indicated by stippled lines.



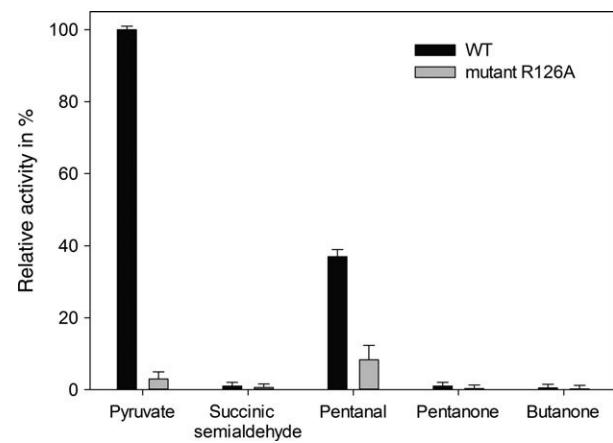
**Fig. 4.** Presentation of the active site with bound *mCPP* (yellow) showing the coordination of the carboxylate mediated via water molecules to the residues Arg126\*, His53\* and Y58. Distances are displayed in Å. Residues defining the small binding pocket are shown in green and those defining the large binding pocket in light blue. The catalytic Lys179 is shown in orange.

#### Influence of the Arg126Ala mutation on the substrate recognition

To obtain deeper insight into the dual substrate recognition of the (*R*)-ATA from *A. fumigatus*, the Arg126Ala variant was generated to investigate the influence of this residue on the conversion of different compounds because this arginine might have a similar function as Arg417 reported previously for the (*S*)-ATA from *Silicibacter pomeroyi* [18]. The mutant was generated by a MegaWhop-PCR, over-expressed in *Escherichia coli*, purified and analyzed



**Fig. 3.** The  $2F_o - F_c$  electron-density maps contoured at  $1\sigma$  of the *mCPP*. (A) *mCPP* in the open active-site loop conformation. Note alternative conformations of the carboxylate on top. (B) *mCPP* in the closed active-site loop conformation with the view rotated about  $90^\circ$ .



**Fig. 5.** Comparison of the relative activity of the (*R*)-ATA from *Aspergillus fumigatus* with different amino acceptors. All data are given relative to the wild-type activity with pyruvate as amino acceptor (100% =  $1105 \text{ mU} \cdot \text{mg}^{-1}$  protein).

using the acetophenone assay [27] for its ability to accept pyruvate, succinic semialdehyde, 2-butanone, 2-pentanone or pentanal. A comparison of the specific activities of the wild-type and the Arg126Ala variant for the five amino acceptors revealed that the activity of the Arg126Ala for pyruvate is significantly reduced to only residual activity (3%) compared with the wild-type. Also, the activity against pentanal is reduced; nevertheless, it is higher than the activity for pyruvate. The low activity towards the other substrates remained almost unaffected (Fig. 5).

#### Metal-binding sites

At the surface of monomer B, a metal ion was observed with an octahedral coordination sphere established by three water molecules (2.37–2.40 Å), two carbonyl

O-atoms of Leu133 and Asn136 (2.20–2.33 Å), and the side chain of Asn137 (2.42 Å). Refinement could not discriminate between the isoelectronic ions of  $\text{Na}^+$  or  $\text{Mg}^{2+}$ . The distances of the coordination sphere are suitable for  $\text{Mg}^{2+}$ , but  $\text{Na}^+$  was defined because of the high concentration for crystallization. This coordination site is only possible with the open loop conformation, which is stabilized by crystal contacts. Thus, enzymatic activity is not depending on this putative metal-binding site.

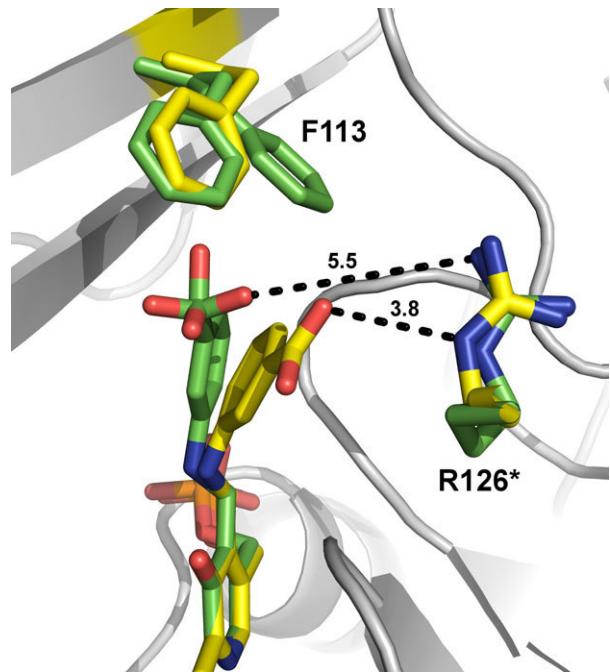
## Discussion

Recently, an inhibitor structure of the (*R*)-ATA from *Nectria haematococca* (PDB code: [4CMF](#)) was published by the Littlechild group [25]. The sequence identity of this transaminase to the here-reported enzyme is 69.1%. In the crystal structure of 4CMF, the Arg126 of the active-site loop – whose guanidinium group is within 5.5 Å of the carboxylate of the *mCPP* – is not important for the binding of the carboxylate [24].

In the here-reported inhibitor complex of (*R*)-ATA from *A. fumigatus* (PDB code: [4UUG](#)), the *mCPP* conformation differs significantly compared with 4CMF. The dihedral angles C4–C4A–N9–C10 (Scheme 1) of these two conformations differ by ~17°, and the carboxylate C-atom C7 is displaced by ~2.7 Å. In the *Nectria haematococca* enzyme the *mCPP* and the pyridoxine are almost in plane to each other, but these planes are at an angle of ~140° in the *mCPP* of the *A. fumigatus* enzyme (Fig. 6). This bend in the inhibitor complex of (*R*)-ATA from *A. fumigatus* results in a relaxed orientation in the active site without clashes with amino acid side chains. By contrast, some distances that are too short are observed in 4CMF, most probably due to some disorder caused by alternative conformations (e.g. Phe113 in Fig. 6).

In addition, compared with 4CMF, the side chain of the Arg126 of 4UUG is closer to the carboxylate of *mCPP* by 1.7 Å and forms a weak salt bridge (Nε–O2 3.8 Å, Fig. 6). This interaction is supported by bridging hydrogen bonds via water molecules to the arginine side chain and the carboxylate of *mCPP* (Fig. 4). The closed active-site loop allows only one distinct conformation of the carboxylate, whereas in the open form at least one additional conformation is observed due to more space and weaker alternative interactions (Fig. 3).

The conformation of the open active-site loop (Fig. 2) is stabilized by contacts to symmetry mates. The closed active-site loop is not influenced by any packing contacts, assuming a distinct active-site func-



**Fig. 6.** Superposition of the inhibitor structures of (*R*)-selective amine transaminase from *Aspergillus fumigatus* and *Nectria haematococca* showing two different conformations of *mCPP*. Yellow C atoms: *mCPP* in chain B of the *Aspergillus fumigatus* enzyme; green C atoms: *mCPP* in chain A of the *Nectria haematococca* enzyme (rmsd on all atoms 0.85 Å). Note the alternative positions of Phe113. N and O atoms in both structures are blue and red, respectively.

tion. In the case of (*S*)-selective ATAs and aromatic amino acid transaminases, which both belong to fold class I, positional changes of the arginine are observed. Arginine flips into the active site to facilitate coordination to carboxylated substrates, but is not involved in binding of hydrophobic substrates [28,29]. Comparison of our results with (*S*)-selective ATAs and the aromatic amino acid transaminase is difficult, because these structures belong to fold class I, whereas the ATA from *A. fumigatus* belongs to fold class IV. Nevertheless, a loop movement is observed in both fold classes [20]. In the ATA from *A. fumigatus* the open active-site loop broadens the entrance to the active site significantly, stabilized by a hydrogen bond between Arg126 and Asp132. This is probably an artificial state induced and stabilized by crystal contacts, but it might be essential for the open conformation of the active-site loop. This assumption needs further investigation to prove the dual substrate recognition of (*R*)-selective ATAs with hydrophobic substrates.

The distinct importance of Arg126 of the active-site loop for binding carboxylated substrates is confirmed

by the Arg126Ala variant. This enzyme variant shows a significant decrease to a residual activity of 3% towards pyruvate compared with the wild-type enzyme. The Arg128Ala variant of the ATA from *A. terreus*, as reported by the Steiner group [23], also supports our observations. Mutations of the flipping arginine of (*S*)-selective ATAs [29] have no significant effect on the activity towards nonpolar amino acceptors, whereas the substrate recognition of nonpolar amino acceptors in (*R*)-selective ATAs requires additional interactions, as indicated by the loss of activity against pentanal.

Interestingly, whereas Arg126 seems to play no important role in the fixed carboxylate-binding position in *Nectria haematococca*, our crystal structure analysis of the inhibitor complex, our kinetic data and the results from the Steiner group [23] support the importance of Arg126 of the active-site loop in determining a carboxylate-binding position in the (*R*)-selective ATA from *A. fumigatus*.

## Conclusion

The crystal structure of the ATA from *A. fumigatus* was solved with the bound inhibitor gabaculine to a resolution of 1.6 Å. The orientation and binding of the carboxylate of pyruvate are facilitated by Arg126 via water molecules. Furthermore, mutagenesis of this arginine present in the active-site loop to alanine results in substantially reduced activity with pyruvate as the amino acceptor. Hence, we conclude that Arg126 is required to finalize substrate recognition and is important for dual substrate recognition in the (*R*)-ATA from *A. fumigatus*.

## Materials and methods

### Chemicals and materials

All chemicals were purchased from Fluka (Buchs, Switzerland), Sigma (Steinheim, Germany), Merck (Darmstadt, Germany), VWR (Hannover, Germany) or Carl Roth (Karlsruhe, Germany) and were used without further purification, unless otherwise specified. Polymerases were obtained from New England Biolabs GmbH (NEB, Frankfurt am Main, Germany) and primers were ordered from Invitrogen (Life Technologies GmbH, Darmstadt, Germany).

### Cloning and mutagenesis

Mutagenesis was performed using the gene encoding the ATA from *A. fumigatus*, which has a C-terminal His tag and was cloned into a pET22b plasmid [2]. The mutant Arg126 Ala was generated by a MegaWhop-PCR. The following

primers were used: specific primer for the mutant Arg126 Ala: CTTCCGGTTAGAACCGCAACACCGGTCAG and the T7 forward primer: TAATACGACTCACTA TAGGG. After digestion of the wild-type plasmid, the mutated plasmid was transformed into competent *E. coli* Top10 cells. The cells were plated on Luria–Bertani–agar plates with 0.1 mg·mL<sup>-1</sup> ampicillin. After overnight culture at 37 °C colonies were picked and the plasmid was isolated using the innuPREP Plasmid Mini Kit from Analytik Jena (Jena, Germany). Sequencing was performed by Eurofins MWG GmbH (Ebersberg, Germany). Finally, the plasmid was transformed into *E. coli* BL21-competent cells.

### Expression, purification and desalting

Expression, purification and desalting of the Arg126Ala variant were performed as described for the wild-type enzyme [30]. The protein content was determined by the BCA assay using a Varian Cary 50 Bio spectrophotometer.

### Activity test

The activity of the wild-type and the Arg126Ala variant was determined by the acetonophenone assay [27]. The ability to process succinic semialdehyde, pentanal, 2-butanone or 2-pentanone as an amino acceptor was investigated by measuring the activity of the variant and the wild-type with α-MBA as the amino donor at 254 nm. Amino acceptors and α-MBA were used at a concentration of 2.5 mM.

### Crystallization and inhibitor soaking

Crystals were obtained by the hanging-drop method and with a reservoir solution containing 0.1 M sodium acetate,

**Table 1.** Data collection and processing statistic. Values in parentheses are for the highest resolution shell.

X-ray source	BL14.1, BESSY II
Detector	Pilatus 6M
Wavelength (Å)	0.91841
Temperature (K)	100
Space group	<i>P</i> 3 <sub>1</sub> 21
<i>a</i> = <i>b</i> / <i>c</i> (Å)	144.4/96.1
Resolution range (Å)	47.28–1.6 (1.63–1.60)
Unique reflections	151504 (7496)
Redundancy	4.5 (4.5)
<i>R</i> <sub>merge</sub> (%)	5.8 (66.8)
<i>R</i> <sub>meas</sub> (%)	6.6 (75.7)
<i>R</i> <sub>pim</sub> (%)	3.1 (35.3)
< <i>I</i> / <i>I</i> <sub>0</sub> >	16.1 (2.3)
CC <sub>1/2</sub> (%)	99.9 (75.6)
Completeness (%)	99.9 (99.8)
Wilson <i>B</i> factor (Å <sup>2</sup> )	26.7
Total rotation/increment (°)	80/0.2

**Table 2.** Refinement statistic.

Data set	(R)-ATA/gabaculine
Resolution range (Å)	47.25–1.6
Working/test reflections	143 859/7626
$R/R_{\text{free}} (\%)^{\text{a}}$	13.4/14.8
Protein residues/water molecules	641/794
No. ligands	10
rmsd from ideality	
Bond lengths (Å)	0.0140
Bond angles (°)	1.6979
Average $B$ factor (Å $^2$ )	21
Ramachandran statistics	
Most favored regions (%)	97.0
Outliers (%)	0.3
PDB entry code	<a href="#">4UUG</a>

<sup>a</sup> $R = \sum |F_o - |F_c|| / \sum |F_o|$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factors, respectively.  $R_{\text{free}}$ . Analogous to R-factor except the summation is over 5% of reflections not included in refinement.

pH 4.6, and 2.0 M sodium formate. The crystallization drop contained an equal volume of reservoir solution and protein solution. The concentration of the protein solution used was 12.2 mg·mL $^{-1}$ . The obtained crystals had only a slight yellow color, and because of this they were first transferred to a solution containing 10% 2-methyl-2,4-pentanediol, 0.1 M sodium acetate pH 4.6, 2.0 M sodium formate and 0.1 mM PLP to ensure full occupation of the cofactor. After 15 min, the crystals obtained a much brighter yellow color and were transferred to a solution containing 10% 2-methyl-2,4-pentanediol, 0.1 M sodium acetate pH 4.6, 2.0 M sodium formate and 0.1 mM gabaculine. This solution had cryo-protectant properties.

### Data collection and structure determination

X-Ray diffraction data were collected at 100 K on beamline 14.1 at the BESSY II synchrotron (Berlin, Germany). All diffraction images were indexed and integrated with XDS [31], scaling and calculation of the structure factors was performed with AIMLESS [32]. Data collection and processing statistics are summarized in Table 1.

The structure was solved with Molecular Replacement using the program PHASER [33]. The structure of the apoenzyme from *A. fumigatus* (PDB code: [4CHI](#)) served as a template [22]. The ligand was built manually into the  $F_o - F_c$  electron-density map with COOT [34]. Occupancies of the mCPP were determined by adapting the  $B$  factors to neighboring amino acid residues. Refinement was carried out with REFMAC5 including TLS segments [35]. The quality of the refined protein model was validated using MOLPROBITY [36]. Refinement statistics are listed in Table 2. All molecular graphics were prepared using PYMOL [37].

### Acknowledgements

MT thanks the Landesgraduiertenkolleg Mecklenburg-Vorpommern for financial support. We thank the European Union (KBBE-2011-5, grant no. 289350) for financial support within the European Union Seventh Framework Programme. Diffraction data were collected on BL14.1 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron-storage ring (Berlin-Adlershof, Germany).

### Author contributions

LS cloned and purified the protein, and performed kinetic assays. MT crystallized, soaked, solved and analyzed the protein structure. UTB and WH initiated the joint project. All authors were involved in discussing data and preparing the manuscript.

### References

- Höhne M & Bornscheuer UT (2009) Biocatalytic routes to optically active amines. *ChemCatChem* **1**, 42–51.
- Höhne M, Schätzle S, Jochens H, Robins K & Bornscheuer UT (2010) Rational assignment of key motifs for function guides in silico enzyme identification. *Nat Chem Biol* **6**, 807–813.
- Kohls H, Steffen-Munsberg F & Höhne M (2014) Recent achievements in developing the biocatalytic toolbox for chiral amine synthesis. *Curr Opin Chem Biol* **19**, 180–192.
- Savile CK, Janey JM, Mundorff EC, Moore JC, Tam S, Jarvis WR, Colbeck JC, Krebber A, Fleitz FJ, Brands J et al. (2010) Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329**, 305–309.
- Höhne M, Kühl S, Robins K & Bornscheuer UT (2008) Efficient asymmetric synthesis of chiral amines by combining transaminase and pyruvate decarboxylase. *ChemBioChem* **9**, 363–365.
- Matcham G, Bhatia M, Lang W, Lewis C, Nelson R, Wang A & Wu W (1999) Enzyme and reaction engineering in biocatalysis: synthesis of *S*-methoxyisopropylamine (= *S*-1-methoxypropan-2-amine). *Chimia* **53**, 584–589.
- Höhne M & Bornscheuer UT (2012) Application of transaminases in organic synthesis. In *Enzymes in Organic Synthesis* (May O, Gröger H & Drauz W, eds), pp. 779–820. Wiley-VCH, Weinheim.
- Schrewe M, Ladkau N, Bühler B & Schmid A (2013) Direct terminal alkylamino-functionalization via multistep biocatalysis in one recombinant whole-cell catalyst. *Adv Synth Catal* **355**, 1693–1697.

- 9 Eliot A & Kirsch J (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu Rev Biochem* **73**, 383–415.
- 10 Percudani R & Peracchi A (2009) The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. *BMC Bioinformatics* **10**, 273–280.
- 11 Berkovitch F, Behshad E, Tang K-H, Enns EA, Frey PA & Drennan CL (2004) A locking mechanism preventing radical damage in the absence of substrate, as revealed by the x-ray structure of lysine 5,6-aminomutase. *Proc Natl Acad Sci USA* **101**, 15870–15875.
- 12 Lepore BW, Ruzicka FJ, Frey PA & Ringe D (2005) The x-ray crystal structure of lysine-2,3-aminomutase from Clostridium subterminale. *Proc Natl Acad Sci USA* **102**, 13819–13824.
- 13 Jansonius JN (1998) Structure, evolution and action of vitamin B6-dependent enzymes. *Curr Opin Struct Biol* **8**, 759–769.
- 14 Cronin CN & Kirsch JF (1988) Role of arginine-292 in the substrate specificity of aspartate aminotransferase as examined by site-directed mutagenesis. *Biochemistry* **27**, 4572–4579.
- 15 Park E-S, Dong J-Y & Shin J-S (2014) Active site model of (*R*)-selective  $\omega$ -transaminase and its application to the production of d-amino acids. *Appl Microbiol Biotechnol* **98**, 651–660.
- 16 Shin J-S & Kim B-G (2001) Comparison of the omega-transaminases from different microorganisms and application to production of chiral amines. *Biosci Biotechnol Biochem* **65**, 1782–1788.
- 17 Hirotsu K, Goto M, Okamoto A & Miyahara I (2005) Dual substrate recognition of aminotransferases. *Chem Rec* **5**, 160–172.
- 18 Steffen-Munsberg F, Vickers C, Thontowi A, Schätzle S, Tumlirsch T, Svedendahl Humble M, Land H, Berglund P, Bornscheuer UT & Höhne M (2013) Connecting unexplored protein crystal structures to enzymatic function. *ChemCatChem* **5**, 150–153.
- 19 Koszelewski D, Göritzer M, Clay D, Seisser B & Kroutil W (2010) Synthesis of optically active amines employing recombinant  $\omega$ -transaminases in *E. coli* cells. *ChemCatChem* **2**, 73–77.
- 20 Humble MS, Cassimjee KE, Håkansson M, Kimbung YR, Walse B, Abedi V, Federsel H-J, Berglund P & Logan DT (2012) Crystal structures of the *Chromobacterium violaceum*  $\omega$ -transaminase reveal major structural rearrangements upon binding of coenzyme PLP. *FEBS J* **279**, 779–792.
- 21 Ford GC, Eichele G & Jansonius JN (1980) Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proc Natl Acad Sci* **77**, 2559–2563.
- 22 Thomsen M, Skalden L, Palm GJ, Höhne M, Bornscheuer UT & Hinrichs W (2014) Crystallographic characterization of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Crystallogr D* **70**, 1086–1093.
- 23 Lyskowski A, Gruber C, Steinikellner G, Schürmann M, Schwab H, Gruber K & Steiner K (2014) Crystal structure of an *R*-selective  $\omega$ -transaminase from *Aspergillus terreus*. *PLoS One* **9**, e87350.
- 24 Sayer C, Martinez-Torres RJ, Richter N, Isupov MN, Hailes HC, Littlechild JA & Ward JM (2014) The substrate specificity, enantioselectivity and structure of the (*R*)-selective amine : pyruvate transaminase from *Nectria haematococca*. *FEBS J* **281**, 2240–2253.
- 25 Kobayashi K, Miyazawa S, Terahara A, Mishima H & Kurihara H (1976) Gabaculine:  $\gamma$ -aminobutyrate aminotransferase inhibitor of microbial origin. *Tetrahedron Lett* **17**, 537–540.
- 26 Rando RR (1977) Mechanism of the irreversible inhibition of  $\gamma$ -aminobutyric acid- $\alpha$ -ketoglutaric acid transaminase by the neurotoxin gabaculine. *Biochemistry* **16**, 4604–4610.
- 27 Schätzle S, Höhne M, Redestad E, Robins K & Bornscheuer UT (2009) Rapid and sensitive kinetic assay for characterization of  $\omega$ -transaminases. *Anal Chem* **81**, 8244–8248.
- 28 Okamoto A, Nakai Y, Hayashi H, Hirotsu K & Kagamiyama H (1998) Crystal structures of *Paracoccus denitrificans* aromatic amino acid aminotransferase: a substrate recognition site constructed by rearrangement of hydrogen bond network. *J Mol Biol* **280**, 443–461.
- 29 Steffen-Munsberg F, Vickers C, Thontowi A, Schätzle S, Meinhardt T, Svedendahl Humble M, Land H, Berglund P, Bornscheuer UT & Höhne M (2013) Revealing the structural basis of promiscuous amine transaminase activity. *ChemCatChem* **5**, 154–157.
- 30 Thomsen M, Skalden L, Palm GJ, Höhne M, Bornscheuer UT & Hinrichs W (2013) Crystallization and preliminary X-ray diffraction studies of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**, 1415–1417.
- 31 Kabsch W (2010) XDS. *Acta Crystallogr D* **66**, 125–132.
- 32 Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* **50**, 760–763.
- 33 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.

- 34 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132.
- 35 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* **67**, 355–367.
- 36 Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12–21.
- 37 Delano WL (2002) PyMOL in <http://www.pymol.org>.

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Ernst-Moritz-Arndt-Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

---

Maren Thomsen  
Unterschrift des Promovenden

# Lebenslauf

## Ausbildung

---

- 07/2007** **Abitur** an der Jürgen-Fuhlendorf-Schule, Gymnasium in Bad Bramstedt, Deutschland
- 10/2007 – 06/2012** Studium der Biochemie an der Ernst-Moritz-Arndt Universität Greifswald, Deutschland
- 10/2011 – 06/2012** **Diplomarbeit** im Studiengang Biochemie an der Ernst-Moritz-Arndt Universität Greifswald, Deutschland  
Betreuer: Prof. W. Hinrichs und Prof. U.T. Bornscheuer  
Thema: Strukturelle und biochemische Charakterisierung einer Chalconisomerase
- 06/2012** Hochschulabschluss mit Auszeichnung als **Diplom-Biochemikerin** an der Ernst-Moritz-Arndt Universität Greifswald, Deutschland
- 10/2012 – 09/2014** Promotions-Stipendiatin des Landesgraduiertenkolleg des Landes Mecklenburg-Vorpommern
- Start 10/2012** **Dissertation** an der Ernst-Moritz-Arndt Universität Greifswald, Deutschland  
Betreuer: Prof. W. Hinrichs  
Thema: Struktur und Funktion der ersten bakteriellen Chalconisomerase und einer (R)-selektiven Amin-Transaminase

## Publikationen in Peer-Reviewed Journals

---

- [1] M. Thomsen, A. Tuukkanen, J. Dickerhoff, G.J. Palm, H. Kratzat, D. Svergun, K. Weisz, U.T. Bornscheuer, W. Hinrichs. Structure and catalytic mechanism of the evolutionarily unique bacterial chalcone isomerase. *submitted*
- [2] L. Skalden\*, M. Thomsen\*, M. Höhne, U.T. Bornscheuer, W. Hinrichs. Structural and biochemical characterization of the dual substrate recognition of the (R)-selective amine transaminase from *Aspergillus fumigatus*. *FEBS J. online*. doi: 10.1111/febs.13149.
- [3] M. Thomsen\*, L. Skalden\*, G.J. Palm, M. Höhne, U.T. Bornscheuer, W. Hinrichs. Crystallographic characterization of the (R)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Cryst. D* **2014**, *70* (4), 1086-1093.
- [4] M. Gall, M. Thomsen, C. Peters, I.V. Pavlidis, P. Jonczyk, P.P. Grünert, S. Beutel, T. Scheper, E. Gross, M. Backes, J.P. Ley, J.M. Hilmer, G. Krammer, G.J. Palm, W. Hinrichs, U.T. Bornscheuer. Enzymatic conversion of flavonoids using bacterial chalcone isomerase and enoate reductase. *Angew. Chem. Int. Ed.* **2014**, *53* (5), 1439-1442. Enzymatische Umsetzung von Flavonoiden

mit einer bakteriellen Chalconisomerase und einer Enoatreduktase, *Angew. Chem.* **2014**, *126* (5), 1463-1466.

- [5] M. Thomsen\*, L. Skalden\*, G.J. Palm, M. Höhne, U.T. Bornscheuer, W. Hinrichs. Crystallization and preliminary X-ray diffraction studies of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Cryst. F* **2013**, *69* (12), 1415-1417.

\*gleichberechtigte Erstautoren

---

## Patentanmeldungen

- [1] J.M. Hilmer, E. Gross, G. Krammer, J.P. Ley, M. Gall, U.T. Bornscheuer, M. Thomsen, C. Peters, P. Jonczyk, S. Beutel, T. Schepers. (2014) Method for the biotechnological manufacture of dihydrochalcones, Patent Anmeldung. EP 2692729, DE 102012213492
- [2] U.T. Bornscheuer, M. Thomsen, W. Hinrichs, E. Gross, T. Geissler, J.P. Ley. Method for the biotechnological manufacture of flavanoneglykoside dihydrochalcones. *submitted* Patent Anmeldung.

---

## Posterpräsentation und Vorträge

- 09/2012** **Vortrag:** Investigation of the flavanone degrading enzymes.  
(15<sup>th</sup> HEC-Meeting in Beilngries, Deutschland)
- 11/2013** **Poster:** The X-ray crystal structure of the first bacterial chalcone isomerase.  
(BESSY II User Meeting, Berlin, Germany)
- 09/2014** **Vortrag:** Bacterial and plant chalcone isomerases have different folds processing the same substrate (17<sup>th</sup> HEC-Meeting in Berlin, Deutschland)

---

## Protein Datenbank-Eintragungen

- 4c9s** Bacterial chalcone isomerase in open conformation from *Eubacterium ramulus* at 1.8 Å resolution
- 4c9t** Bacterial chalcone isomerase in open conformation from *Eubacterium ramulus* at 2.0 Å resolution, SelenoMet derivative
- 3zph** Bacterial chalcone isomerase in closed conformation from *Eubacterium*

*ramulus* at 2.8 Å resolution

- 4d06** Bacterial chalcone isomerase complexed with naringenin
- 4d4f** Mutant P250A of the bacterial chalcone isomerase from *Eubacterium ramulus*
- 4chi** (*R*)-selective amine transaminase from *Aspergillus fumigatus* at 1.27 Å resolution
- 4uug** The (*R*)-selective amine transaminase from *Aspergillus fumigatus* with inhibitor bound

# Danksagung

An erster Stelle möchte ich mich bei meinem Doktorvater Prof. Winfried Hinrichs für die Betreuung und Bereitstellung dieses spannenden und sehr abwechslungsreichen Themas bedanken. Auch möchte ich mich für das Vertrauen in meine wissenschaftlichen Fähigkeiten und die ständige Unterstützung meiner Arbeit bedanken.

Des Weiteren möchte ich mich bei Prof. Uwe T. Bornscheuer für die tolle Kooperation und Betreuung in biotechnologischen Fragen sowie dem immerwährenden Interesse in meine Arbeit und die Unterstützung bedanken.

Ich danke dem Landesgraduiertenkolleg Mecklenburg-Vorpommern für die finanzielle Unterstützung meiner Promotion.

Der Symrise AG, vor allem den Mitarbeitern Egon Gross, Dr. Torsten Geißler und Dr. Jakob Ley, danke ich für die gute Zusammenarbeit und stete Unterstützung auch auf Gebieten der Grundlagenforschung. Ebenso danke ich dem Arbeitskreis von Prof. Schepel (Gottfried Wilhelm Leibniz Universität Hannover) für die gute Zusammenarbeit im Rahmen des Kooperationsprojektes. Mechthild Gall danke ich für die Vorarbeiten zu diesem Projekt.

Dr. Gottfried J. Palm und Dr. Ioannis V. Pavlidis sei gedankt für die stete Hilfsbereitschaft und Diskussion diverser Problematiken.

Meiner Bachelorstudentin Hanna Kratzat danke ich für ihr großes Engagement und den Einblick in die Welt der Filmproduktion.

Ich danke außerdem allen Mitgliedern der Arbeitskreise „Molekulare Strukturbioologie“ (insbesondere Britta Girhardt und Leona Berndt) und „Biotechnologie & Enzymkatalyse“ für die Unterstützung und allzeit gute und freundliche Arbeitsatmosphäre im Labor.

Für Kaffee- und Teepausen und einen immer gutgefüllten Schokoladenvorrat danke ich besonders Christin Peters, Lilly Skalden, Michael Fibinger, Jan Muschiol, Maria Jenckel, Daniel Möller und Stefan Saß. Ohne euch hätte es nicht so viel Spaß gemacht!

Mein größter Dank aber gehört Fabian, für seine unerschütterliche Zuversicht, ständige Motivation und unendliche Geduld! Ohne deine Unterstützung wäre dies alles nicht möglich gewesen!