TUM School of Natural Sciences
Technical University of Munich

Mechanistic Studies on Polyketide Biosynthesis in Pathogenic Fungi and Bacteria

Maximilian Schmalhofer

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Lukas Hintermann

Prüfer*innen der Dissertation:
1. Prof. Dr. Michael Groll
2. Prof. Dr. Thomas Brück
3. Prof. Dr. Helge B. Bode

Die Dissertation wurde am 31.05.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 19.07.2023 angenommen.

Summary

Polyketides (PKs) are chemically complex and diverse compounds that are widespread among bacteria, fungi, and plants. PKs often possess biological activities and are frequently used as pharmaceuticals. Promising examples of such applications include the cholesterol-lowering drug lovastatin or the antibiotic doxycycline. Thus, the investigation of PKs and their biosynthesis (BS) is of great interest to society.

This work provides mechanistic insights into the biosynthesis of the polyketides aryl polyene lipids (APELs), 1,8-dihydroxynaphthalene (DHN)-melanin, and anthraquinone (AQ). These three PKs are produced by pathogenic bacteria or fungi and, up to date, only little is known about their BS on a structural level.

In the first part, the bacterial APELs, which contain a striking aryl polyene (APE) core moiety, were characterized. APELs are synthesized by a new family of APE type II polyketide synthase systems encoded by one of the most common biosynthetic gene clusters among Gram-negative bacteria. However, little is known about the function, biosynthesis, or biological relevance of APELs. Conspicuously, the biosynthesis of the carbon-skeleton backbone and its processing into the polyene core moiety is driven by seven structurally uncharacterized homo- or heterooligomeric complexes. Therefore, the X-ray crystal structures of all quaternary complexes were determined from pathogenic *Xenorhabdus doucetiae* or *Escherichia coli* O157:H7 (strain Sakai). These include the two ketosynthases ApeR$_2$ and ApeO:C, the ketoreductase ApeQ$_4$, the dehydratase ApeI:P, and the thioesterase-like ApeK$_4$ in the absence and presence of the two acyl carrier proteins (ACPs) ApeK$_4$:E$_4{}^{holo}$ and ApeK$_4$:F$_4{}^{holo}$, respectively. Using the X-ray analysis for structure determination, first insights into the underlying mechanisms, such as sites of catalysis, complex formation, and ACP^{holo} binding, as well as polyene and 4'-phosphopantetheine (Ppant) cavities, are highlighted.

In the second part, the biosyntheses of DHN-melanin and AQs are investigated. DHN-melanin is a virulence factor in pathogenic fungi, and a central modification is the polyketide trimming step during its BS in *Aspergillus fumigatus* and *Exophiala dermatitidis*. The lyases *Af*yg1p and *Wd*yg1p catalyze this cleavage of carbon-carbon bonds. A similar reaction occurs during the anthraquinone BS in *Photorhabdus laumondii*. Here, the homologous lyase AntI performs chain-length shortening and cyclization. The obtained X-ray crystal structures of *Af*yg1p, *Wd*yg1p, and AntI in the presence of products, surrogates, inhibitors, and compounds that mimic the sp^3-hybridized intermediate state were used to analyze the binding and catalysis of the three lyases.

Taken together, the obtained results of the APEL, DHN-melanin, and AQ BS built the foundation for subsequent and ongoing studies for targeted mutagenesis and the design of inhibitors for pathogenic bacteria and fungi.

Zusammenfassung

Polyketide (PKs) sind chemisch komplexe und diverse Substanzen, die in Bakterien, Pilzen und Pflanzen weitverbreitet sind. Aufgrund ihrer biologischen Aktivitäten werden sie oft als Pharmazeutika eingesetzt, wie beispielsweise das cholesterinsenkende Lovastatin oder das Antibiotika Doxycyclin. Daher ist die Erforschung von PKs und ihrer Biosynthese (BS) von großem allgemeinen Interesse.

Die hier vorgelegte Arbeit liefert mechanistische Einblicke in die BS der Polyketide Aryl Polyen Lipide (APELs), 1,8-Dihydroxynaphthalin (DHN)-Melanin und Anthrachinon (AQ). Diese drei PKs sind von zentraler Bedeutung, da sie in pathogenen Bakterien und Pilzen produziert werden, aber bisher nur wenig bekannt ist über die molekulare Zusammensetzung der Biosynthese.

Im ersten Teil dieser Arbeit werden APELs untersucht, die einen Aryl Polyen (APE)-Kern als Hauptmerkmal aufweisen. Sie werden von einer neuartigen Familie der APE Typ II Polyketidsynthasen produziert. Obwohl ihr kodierender Gen-Cluster häufig in gramnegativen Bakterien vorkommt, ist nur wenig über die Funktion, Biosynthese oder biologische Bedeutung dieser PKs bekannt. Der Aufbau des Kohlenstoffgerüsts und dessen Prozessierung in den APE-Kern basiert auf sieben strukturell uncharakterisierten Homo- oder Heterooligomeren, weshalb die Proteinstrukturen aller quaternären Komplexe aus den Pathogenen *Xenorhabdus doucetiae* oder *Escherichia coli* O157:H7 (Stamm Sakai) hier gelöst wurden. Zentrale Merkmale dieser Synthese-Maschinen sind die beiden Ketosynthasen ApeR$_2$ und ApeO:C, die Ketoreduktase ApeQ$_4$, die Dehydratase ApeI:P und das Thioesterase-ähnliche ApeK$_4$ in ungebundener Form, sowie der Komplex mit den beiden Acyl Carrier Proteinen (ACPs) ApeK$_4$:E$_4{}^{holo}$ und ApeK$_4$:F$_4{}^{holo}$. Die Röntgenanalyse der Kristallstrukturen liefert tiefgehende Einblicke in die jeweiligen Mechanismen und Regionen in denen die Katalyse und die Oligomerisierung sowie die Bindung der ACPs, der 4'-Phosphopantetheine (Ppants) oder der Polyene stattfinden.

Im zweiten Teil der vorgelegten Arbeit steht die Untersuchung der BS von DHN-Melanin sowie von AQ im Vordergrund. DHN-Melanin ist ein Virulenzfaktor in pathogenen Pilzen und ein zentraler Naturstoff in *Aspergillus fumigatus* sowie *Exophiala dermatitidis*. Besondere Aufmerksamkeit erweckt die Polyketid-Trimmung, bei der Kohlenstoff-Kohlenstoff-Bindungen durch die Lyasen *Af*yg1p und *Wd*yg1p gespalten werden. Auch in der AQ BS von *Photorhabdus laumondii* findet sich eine homologe Lyase AntI, welche die Kettenverkürzung und eine Zyklisierung katalysiert. Basierend auf den in der Promotionsarbeit gelösten Proteinstrukturen von *Af*yg1p, *Wd*yg1p und AntI gebunden mit verschiedenen Produkten, Surrogaten, Inhibitoren und Verbindungen, die den tetraedrischen Zwischenzustand imitieren, werden Mechanismen der Bindung und der Katalyse in diesen Lyasen auf atomarer Ebene analysiert. Zusammenfassend bilden die kombinierten Ergebnisse der APEL, DHN-Melanin und AQ BS die Grundlage für weiterführende Studien zur Mutagenese oder für die Entwicklung von Inhibitoren zur Behandlung von bakteriellen Infektionen und Mykosen.

Table of contents

1 Introduction

1.1 Importance of natural products

Cellular metabolism comprises all chemical reactions in living cells and is vital for their survival.[1,2] These complex biochemical processes synthesize, destruct, and convert metabolites. While lipids, carbohydrates, amino acids, or nucleic acids are essential for all functions, toxins, pigments, or antibiotics are species selective and necessary to occupy ecological niches (**Fig. 1.1**).[3,4]

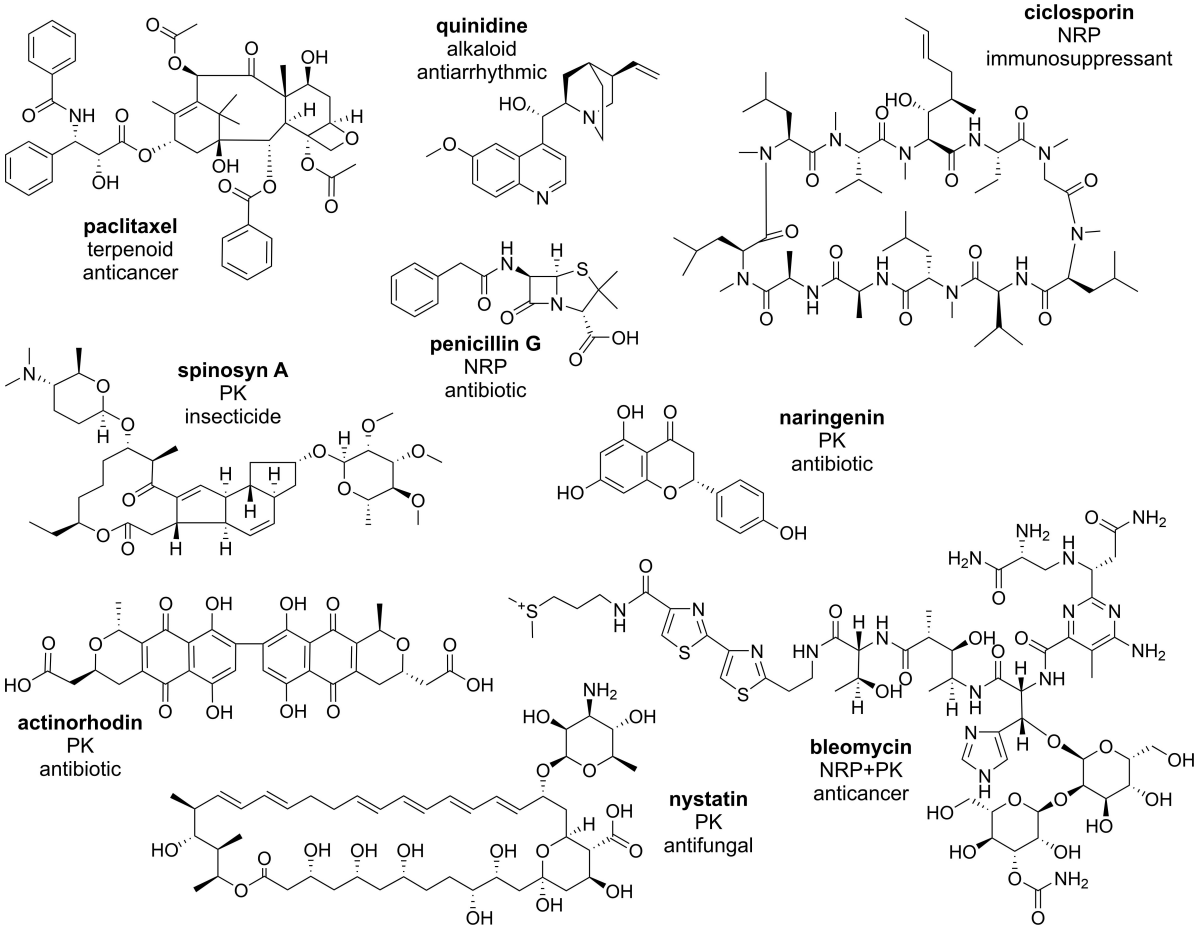

Figure 1.1 Selection of secondary metabolite classes and their possible applications. Shown are the chemical structures of the terpenoid paclitaxel[5,6], the alkaloid quinidine[7], the NRPs penicillin G[8] and ciclosporin[9,10], the PKs spinosyn A[11,12], naringenin[13], actinorhodin[14], and nystatin[15], and the NRP+PK hybrid bleomycin[16]. Figure is adapted from Staunton et al.[17]

Primary metabolites are mostly involved in pathways that sustain and provide energy in cells, like the tricarboxylic acid cycle, glycolysis, or the mitochondrial electron transport

chain.[1,18] Whereas secondary metabolites, termed natural products (NPs) hereafter, are often relevant for interactions with surrounding organisms and environment (ecological functions), improved uptake of nutrients (enhanced growth),[19] or suppression of surrounding organisms (antagonistic interactions)[20]. Predominantly, plants, fungi, and bacteria produce NPs that exhibit diverse bioactivities and are therefore of high interest to humankind.[21] Over the years, many NPs and derivatives were approved as clinical therapeutics, with the antibiotic penicillin as a well-known example.[8] Based on their chemical structure and biosynthesis (BS), common classes of NPs are terpenoids, alkaloids, nonribosomal peptides (NRPs), polyketides (PKs), and hybridized versions (**Fig. 1.1**). This thesis focuses on the BS of polyketides. The following **Section 1.2** presents the logic of type II polyketide synthase (PKS) systems, focusing on the aryl polyenes (APEs). These machinery produce the aryl polyene lipids (APEL) in pathogenic bacteria, like *Xenorhabdus doucetiae* or *Escherichia coli* O157:H7 (strain Sakai). **Section 1.3** introduces the lyases *Af*yg1p, *Wd*yg1p, and *Pl*AntI that participate in the modification of PKs into the bioactive NPs. *Af*yg1p and *Wd*yg1p catalyze polyketide shortening in the BS of 1,8-dihydroxynaphthalene-melanin from pathogenic *Aspergillus fumigatus* and *Exophiala dermatitidis*. In contrast, *Pl*AntI is responsible for polyketide shortening and the final ring cyclization in the anthraquinone production from entomopathogenic *Photorhabdus laumondii*.

1.2 The aryl polyene type II polyketide synthase systems

Generally, PKS systems consist of different functional domains similar to those found in fatty acid synthases (FAS).[22] Considering these domains' overall architecture and assembly, the formed PKS systems are grouped into three types.[23] While type I PKS systems consist of a single peptide chain that includes many domains (multifunctional enzyme), the type II and type III PKSs consist of several discrete peptide chains that form monofunctional proteins.[23] Type I PKS systems are widely distributed among bacterial and fungal species, type II PKSs are exclusively found in bacteria, and type III PKSs occur predominantly in plants. The following presents molecular and mechanistic insights into the natural products actinorhodin (Act)[24,25] and anthraquinone (AQ)[26,27] from *Streptomyces coelicolor* and *Photorhabdus laumondii*, respectively. The corresponding *act* and *ant* genes that encode for the discrete enzymes of these two aromatic (ARO) type II PKS systems assemble in biosynthetic gene clusters (BGCs, **Fig. 1.2a**).[26,28] The biosynthesis of Act and AQ proceeds in three phases: backbone formation, processing, and modification (**Fig. 1.2b**, nomenclature adapted from Cummings et al.[29]).[24,26,27,30] The functional domains involved in the first phase are malonyl-CoA [ACP] transacylases (MCATs), acyl carrier proteins (ACPs), and β-ketoacyl [ACP] synthases (ketosynthases, KSs). MCAT transfers activated precursor molecules (e.g., malonyl-CoA) to ACPs, while the KSs catalyze carbon-carbon (C-C) bond formation between the units by decarboxylative Claisen-condensations. Notably, in type II PKS systems, one KS subunit (KS_α) and a ketosynthase/chain-length factor (KS/CLF, KS_β) form a KS_α:KS_β heterodimer. Iteratively, the skeleton backbone (an octaketide) is formed.

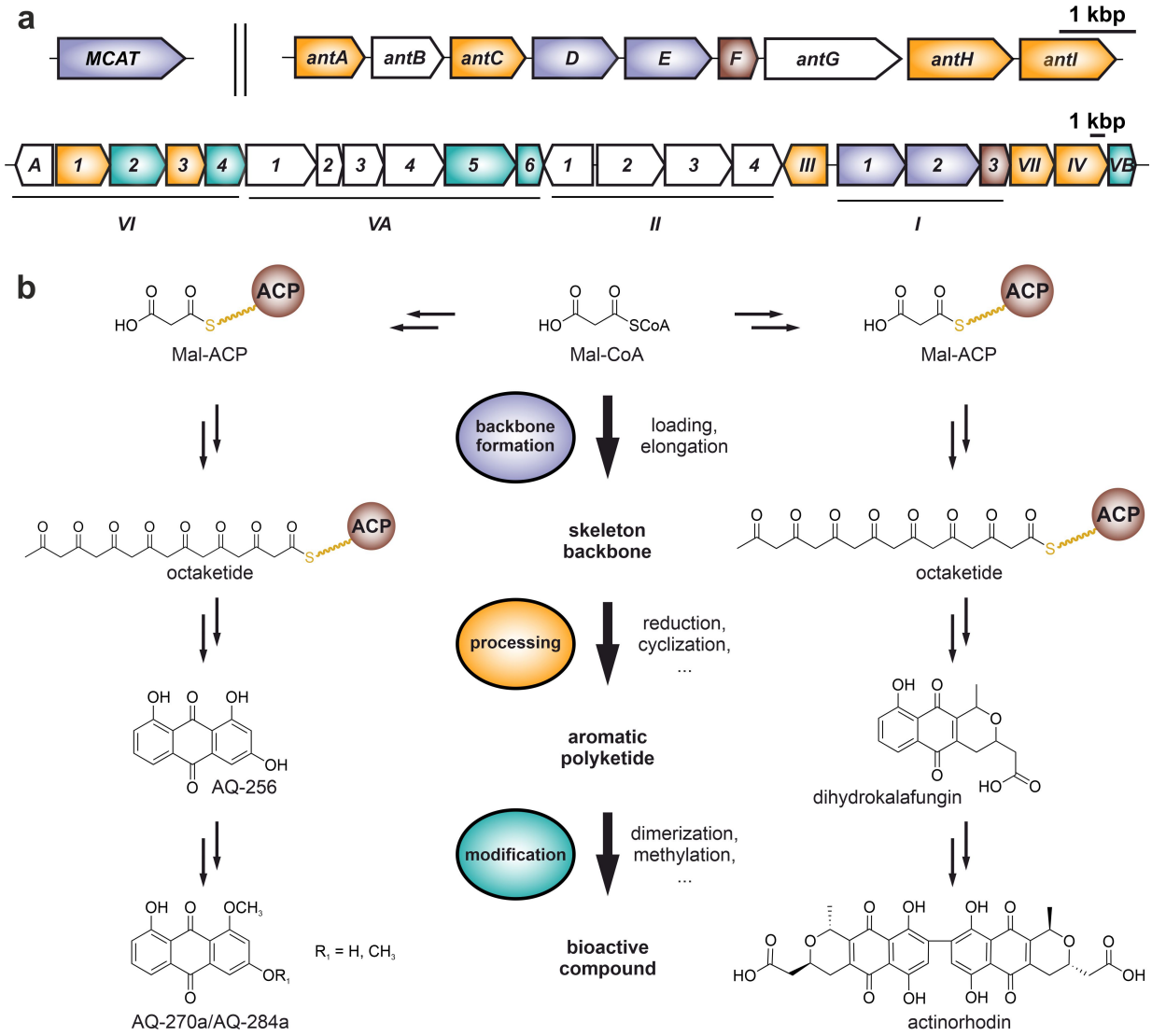

Figure 1.2 Aromatic polyketides anthraquinone and actinorhodin are produced by ARO type II PKS systems. (a) Anthraquinone (top) and actinorhodin (bottom) biosynthetic gene clusters from *Photorhabdus laumondii* and *Streptomyces coelicolor*, respectively. Minimal PKS system is indicated in blue, ACP in brown, and genes involved in processing and modification are colored in yellow and teal, respectively. **(b)** Underlying biosynthetic logic is shown exemplarily for BS of AQ-270a/284a (left panel) and actinorhodin (right panel). Highlighted are the minimal PKS system that is responsible for chain elongation and formation of the carbon-skeleton backbone, the processing phase into the cyclized polyketides, and the modification to finalize the bioactive compound. Figure is adapted from Craney et al.[24] and Bräuer et al.[27]

A central element in the BS of PKs or FAs is the ACP, composed of less than 100 amino acids. A 4'-phosphopantetheine (Ppant)-arm is attached to a conserved serine residue in their *holo*-state. The growing polyketide chain, as well as the starter and elongation units, is tethered to this Ppant-arm of the ACPs as a thioester. The ACPs are responsible for shuttling the cargo to the individual functional domains (e.g., KSs). The three enzymes MCAT, ACP, and KS_α:KS_β form the so-called "minimal PKS system". Once the ACP-tethered core skeleton is formed, the processing into the aromatic polyketide takes place. This procedure includes several reduction, cyclization, and aromatization reactions that are performed by β-ketoacyl [ACP] reductases (ketoreductases, KRs) and aromatases/cyclases (Aro/Cyc). The last phase is the modification of the aromatic to bioactive compounds. This step is in-

dependent of ACPs and often termed 'tailoring'. In the case of the Act BS, dimerization[31] of dihydrokalafungin is carried out to actinorhodin, whereas, in the BS of anthraquinones, AQ-256 is methylated to AQ-270a and AQ-284a (**Fig. 1.2b**).

In addition to the ARO type II PKS systems producing aromatic polyketides, there exist other families as well, including the highly reducing (HR), the aryl polyene (APE), and the cinnamoyl lipid (CL) type II PKS systems. **Figure 1.3a** shows selected natural products assembled by these families, such as ishigamide[32,33] from *Streptomyces* species (sp.) MSC090213JE08 (HR), the aryl polyene (APE_{Xd})[34,35] from *Xenorhabdus doucetiae* (APE), and skyllamycin A[36,37] from *Streptomyces* sp. Acta 2897 (CL). Noteworthy, the BGCs en-

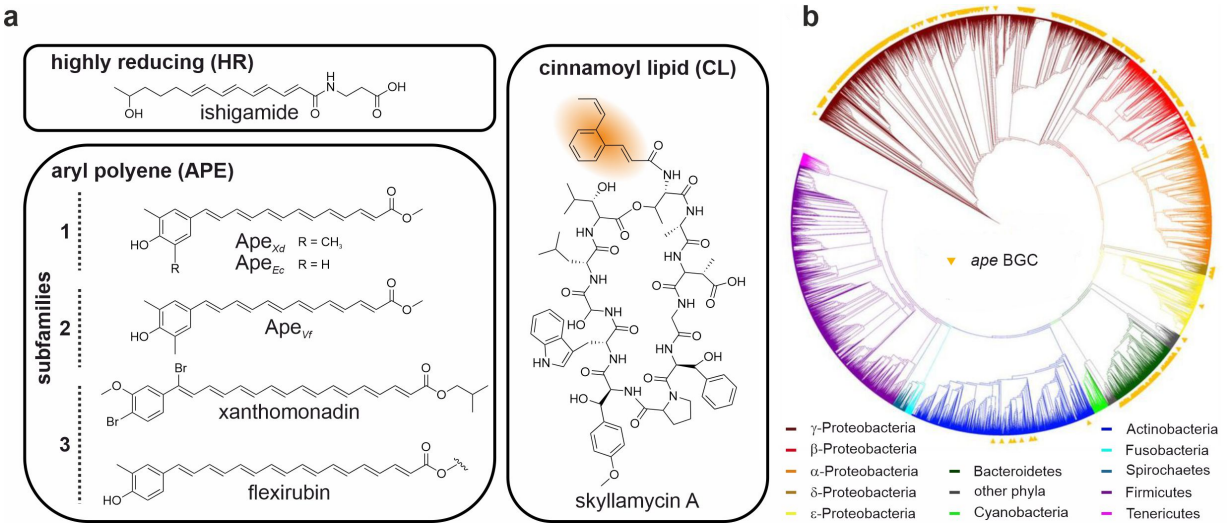

Figure 1.3 Families of polyene type II PKS systems and their products. (a) Shown are the NPs ishigamide, aryl polyenes (Ape_{Xd}, Ape_{Ec}, Ape_{Vf}, xanthomonadin, and flexirubin), and a cinnamoyl lipid (skyllamycin A, cyclized polyene moiety in orange) that contain polyene moieties produced by HR, APE, and CL type II PKS system families, respectively.[34] Skyllamycin A is a depsipeptide produced by an NRPS-PKS hybrid system.[36] **(b)** BGCs encoding for the APE type II PKS subfamily 1 are widespread across most Gram-negative bacterial classes (indicated by yellow triangles). Representation of the BGC is adapted from Grammbitter et al.[38]

coding for the APE type II PKS systems are one of the most common BGCs among Gram-negative bacteria and are often present in pathogenic bacteria (**Fig. 1.3b**).[34] Cimermancic and coworkers classified these APE type II PKS machinery and their products into three subfamilies. The APE type II PKS subfamily 1 produces APEs like Ape_{Xd}[34] from *X. doucetiae* or Ape_{Ec}[34] from *E. coli* CFT073. Subfamily 2 yields exemplarily Ape_{Vf}[34] from *Vibrio fischeri* ES114, and subfamily 3 is involved in the BS of xanthomonadin[39,40] from *Xanthomonas juglandis* strain XJ103 or flexirubin[41] from *Flexibacter elegans* (**Fig. 1.3a**). Interestingly, aryl polyene pigments might protect bacterial cells against light-induced damage and reactive oxygen species.[42–44] Additionally, Ape_{Ec} was assigned to influence the biofilm formation in pathogenic *E. coli* CFT073.[44] Up to now, the full-length products of APE type II PKS systems are unknown as only the hydrolyzed methyl esters (e.g., Ape_{Xd} or Ape_{Ec}) have been isolated and characterized. One exception is the recently analyzed aryl polyene lipid (APEL) from the Gram-negative entomopathogenic bacterium *X. doucetiae*.[38] The chemical structure of APEL is depicted in **Figure 1.4a** and harbors, besides the aryl polyene moiety (highlighted in yellow), a conjugated fatty acid (cFA, highlighted in blue).[35] The underlying BGC was iden-

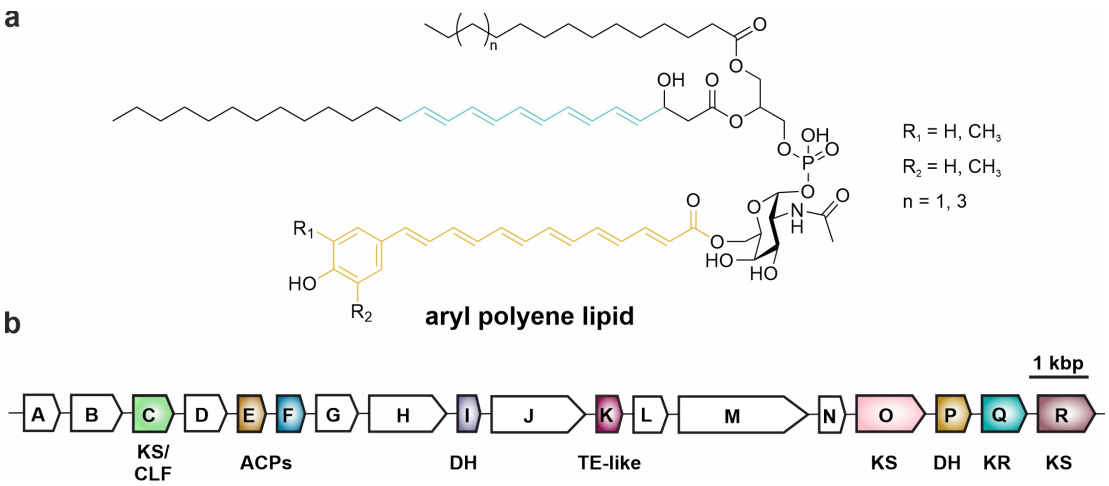

Figure 1.4 The aryl polyene lipid (APEL) from *Xenorhabdus doucetiae*. **(a)** Chemical structure of the full-length aryl polyene lipid (APEL). The APE and cFA chromophores are highlighted in yellow and blue, respectively).[38] **(b)** The APE BGC with the highlighted genes involved in BS of the APE polyene core.[35] Figure is adapted from Grammbitter et al.[38]

tified, and the enzymes involved in the production of the APE polyene core (Ape_{Xd}) were assigned by *in-vitro* reconstitution of the APE BS (**Fig. 1.4b**).[35] Intriguingly, enzyme analyses reveal several fascinating features, some of which are uncommon for known type II PKS systems.[35,38] First of all, the APE BGC encodes for two β-ketoacyl [ACP] synthases (ketosynthases, KSs: ApeO and ApeR), a ketosynthase/chain-length factor (KS/CLF: ApeC), one β-ketoacyl [ACP] reductase (ketoreductase, KR: ApeQ), two β-hydroxyacyl [ACP] dehydratases (dehydratases, DHs: ApeI, ApeP), two acyl carrier proteins (ApeE and ApeF), and one thioesterase-like enzyme of unknown function (TE: ApeK, **Fig. 1.4b**). In addition, native electrospray ionization mass spectrometry (nESI-MS), pull-down, and gel filtration experiments revealed the formation of seven quaternary complexes formed by these nine enzymes, including a KS homodimer ($ApeR_2$) that is common among type II FAS, a KS_α:KS_β heterodimer (ApeO:C, typical for type II PKS), a KR tetramer ($ApeQ_4$), a heterodimeric DH (ApeI:P), and the TE-like enzyme ($ApeK_4$) that forms distinct heterooligomers (*Xd*$ApeK_4$:$F_4{}^{holo}$, *Xd*$ApeK_4$:$E_4{}^{holo}$) in the presence of either ApeE or ApeF.[35] Nevertheless, the proposed biosynthetic pathway of the APE type II PKS system to produce APE_{Xd} is similar to the BS of fatty acids and aromatic polyketides from type II FAS systems and ARO type II PKS systems, respectively (**Fig. 1.5**).[22,35,38] Both ACPs, ApeE and ApeF, harbor the Ppant-arm and shuttle starter and elongation units between the individual domains (**Fig. 1.5a, b**). In the first reaction, the KSs (ApeO:C or $ApeR_2$) catalyze C-C bond formation between the ACP-bound cargo by decarboxylative Claisen condensations (chain elongation, ①, **Fig. 1.5c**). The KR ($ApeQ_4$) reduces the formed β-ketoacyl-ACP to a β-hydroxyacyl-ACP using reduced (red.) nicotinamide adenine dinucleotide phosphate (NADPH) as a substrate (1^{st} reduction, ②). The DHs (ApeI:P) catalyze a condensation and form α,β-trans-enoyl-ACP in the next step (dehydration, ③). In the BS of fatty acids by type II PKS systems, this group is further reduced to the saturated acyl-ACP by the enoyl [ACP] reductases (ERs, utilizing NADPH, 2^{nd} reduction, ④). The α,β-trans-enoyl-ACP or acyl-ACP enters the cycle several times as

a new substrate for ①-③/④ until the final polyene or fatty acid emerges (**Fig. 1.5c**). The polyene is then modified into the bioactive aryl polyene lipid (⑥).

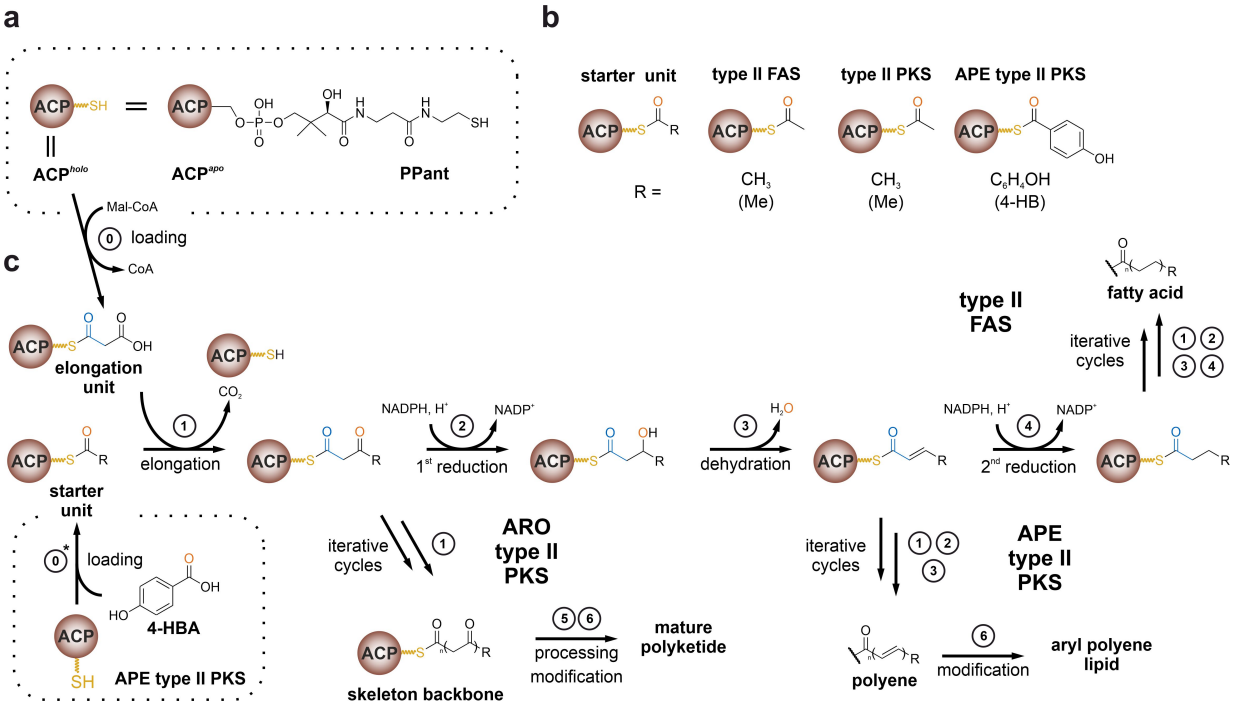

Figure 1.5 Biosynthesis of type II PKS and type II FAS systems. (a) The attachment of a phosphopantetheine arm to a conserved serine residue of ACP^{apo} results in ACP^{holo}. **(b)** Proposed starter units in the APE type II PKS system. 4-HB corresponds to 4-hydroxybenzoyl. **(c)** Overview of the biosynthetic pathways of type II FAS and APE or ARO type II PKS systems, included are ⓪ (loading of elongation units, MCATs) or ⓪* (loading of starter unit 4-hydroxybenzoic acid, ligase ApeH[35]), ① (chain elongation, KSs), ② (1^{st} reduction, KRs), and ③ (dehydration, DHs). Additional steps include ④ (2^{nd} reduction, ERs) or ⑤ (processing, KRs and Aro/Cycs) and ⑥ (modification, e.g., MTs). Figure is adapted from Chen et al.[22]

1.3 Polyketide shortening in the AQ and DHN-melanin biosynthesis

The fungus *Aspergillus fumigatus* is a common airborne pathogen that can cause threatening infections in immunocompromised humans.[45,46] Conidial pigmentation is important for the virulence of *A. fumigatus*, and two major types of pigments are described, the pyomelanin and the 1,8-dihydroxynaphthalene (DHN)-melanin (**Fig. 1.6**).[47,48] The latter plays a role in the recognition of the pathogen and the host's immune response.[49–52] In the fungal DHN-melanin pathway, type I PKS systems are responsible for the production of the central precursor, the pentaketide 1,3,6,8-tetrahydroxynaphthalene (THN). One example is the type I PKS system Pks1 in *Colletotrichum lagenarium*.[53–55] The emerging THN is subsequently converted to 1,8-DHN, which is then polymerized and modified into the chemically diverse DHN-melanin (**Fig. 1.7**).

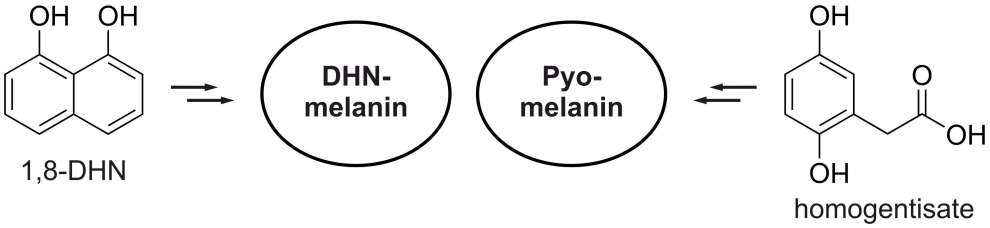

Figure 1.6 1,8-dihydroxynaphthalene (DHN)-melanin and pyomelanin in *Aspergillus fumigatus*. Polymerization and modification of the key precursors 1,8-DHN and homogentisate yield the two major types of melanins in *A. fumigatus*.

In contrast to this pathway, *A. fumigatus* uses an uncommon route. Here, the product of PksP, the type I PKS system, is the naphthopyrone heptaketide YWA1 instead of THN (**Fig. 1.7a**).[56,57] In the next step, this heptaketide is cleaved by *Af*yg1p (in literature often called Ayg1p) to THN, which is further converted to 1,8-DHN by the enzymes Abr1, Abr2, Arp1, and Arp2 (**Fig. 1.7b**).[57–59] There are only a few other examples of type I PKS systems that produce the heptaketide, such as YWA1, e.g., WA[60] in *Aspergillus nidulans*, FwnA[61,62] in *Aspergillus niger*, or PKS12[63] in *Fusarium graminearum*. Another exception is present in the DHN-melanin-producing fungi *Exophiala dermatitidis*. Here, the type I PKS system *Wd*Pks1 produces the hexaketide 2-acetyl-1,3,6,8-tetrahydroxynaphthalene (ATHN) that is deacetylated to THN by the enzyme *Wd*yg1p (**Fig. 1.7a**).[64,65] Sequence analysis showed that the enzymes *Af*yg1p and *Wd*yg1p are α/β-hydrolases containing a catalytic triad (serine, histidine, and aspartate). Enzyme kinetics, inhibition, and mutational studies on both lyases confirmed the substrates (YWA1 or ATHN), the products (THN or THN and acetoacetic acid for *Af*yg1p) as well as the relevance of the active site residues.[58,65] Some homologous enzymes to *Af*yg1p and *Wd*yg1p have been identified in other bacteria as well as fungi, for example, *Pl*AntI from the Gram-negative bacteria *Photorhabdus laumondii* (sequence identity (SI) = 17.4 %, sequence similarity (SS) = 37 % to *Af*yg1p) or the thioesterase/claisen cyclase (TE/CLC) domain from pathogenic *Aspergillus parasiticus* (SI = 8.8 %, SS = 18.2 % to *Af*yg1p, **Fig. 1.7c**). AntI is part of a type II PKS system producing AQ-256 and catalyzes the chain-length shortening (octaketide to heptaketide) as well as its cyclization (see **Section 1.2**).[26] The TE/CLC domain is a more distinct relative and is part of pksA, a type I

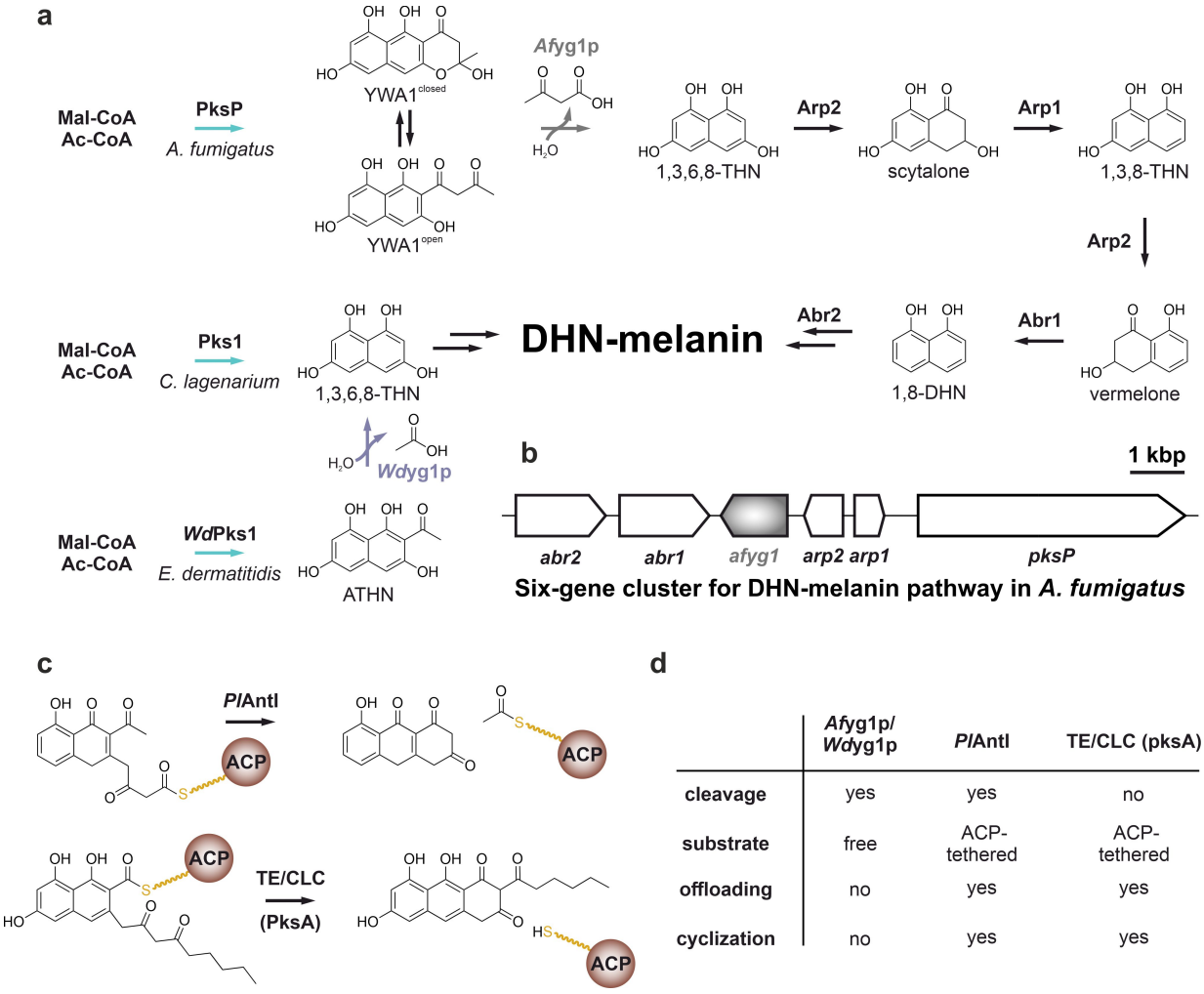

Figure 1.7 The production of DHN-melanin in *Aspergillus fumigatus*. (a) Schematics of the DHN-melanin pathway in *A. fumigatus*. Type I PKS systems (blue arrows) forming THN (Pks1 in *C. lagenarium*), YWA1 (PksP in *A. fumigatus*), or ATHN (WdPks1 in *E. dermatitidis*) are key in product formation. **(b)** Gene cluster involved in BS of DHN-melanin in *A. fumigatus*. Involved genes are *abr1*, *abr2* (laccases), *ayg1* (lyase), *arp2* (hydroxynaphthalene reductase), *arp1* (scytalone dehydratase), and *pksP* (type I PKS system). Six-gene cluster is adapted from Tsai et al.[57] **(c)** In the natural product biosynthesis of AQ-256[26] from *P. laumondii* and of Aflatoxin B_1[66] from *A. parasiticus*, putative chemical reactions are catalyzed by homologous proteins. **(d)** Summary of the similarities and differences of the catalyzed reactions.

PKS system from *Aspergillus parasiticus*, involved in the BS of norsolorinic acid anthrone, a precursor of aflatoxin B_1.[66] Interestingly, the TE/CLC domain catalyzes the sole cyclization and is not involved in polyketide shortening. Additionally, the substrates of *Pl*Antl and the TE/CLC domain are bound to acyl carrier proteins, and part of their reaction is the offloading and release of the bound ACPs. In contrast, the substrates of the lyases *Af*yg1p and *Wd*yg1p are converted in the absence of ACPs (**Fig. 1.7d**).

2 Seven quaternary complexes in aryl polyene lipid core construction

2.1 Objectives

In 2014, Cimermancic and coworkers analyzed APE type II PKS enzymes and assigned them as a new family of PKS systems (**Fig. 1.3**).[34] Despite being encoded by one of the most common BGC among Gram-negative bacteria, only little is known about this enzymatic machinery and its biological relevance. Additionally, no structural information is available on any of the involved enzymes or the bioactive metabolite. Recently, the group of Prof. Dr. Helge B. Bode (Department of Biosciences, Goethe University Frankfurt and Department of Natural Products in Organismic Interactions, Max Planck Institute for Terrestrial Microbiology, Marburg, DE) used the bacterial pathogen *Xenorhabdus doucetiae* as a model organism to identify the full-length natural product, the aryl polyene lipid (APEL), as well as a hydrolyzed methyl ester compound, the aryl polyene (APE, **Fig. 1.4a**).[35,38] Furthermore, they established an *in-vitro* assay that includes all enzymes to produce the APE polyene core (Ape_{Xd}). Intriguingly, the catalysts form seven homo- or heterooligomeric complexes, most of which have not been analyzed so far (e.g., two KS dimers, a heterodimeric DH, or two heterooctameric ACP:TE complexes[35]) and structural information on these complexes is still missing. Therefore, the goal of this Ph.D. thesis was the three-dimensional (3D) structure elucidation of all enzymes involved in the aryl polyene core production. i) The project should start with achieving high purity and yield of the target proteins in their quaternary states by optimizing existing and establishing new protocols. ii) Crystallization and structure determination by X-ray crystallography was the major task. iii) The sites of catalysis, complex formation, and ACP^{holo} binding, as well as the polyene and Ppant cavities, should be characterized in the following. Taken together, the generated data are sought to give insights into the core of the APE type II PKS machinery on a molecular level to allow the future investigation of the biosynthesis of the widespread APEL in pathogenic bacteria with potential for medical applications.

2.2 Results and Discussion

This section presents the X-ray crystal structures of seven quaternary complexes that consist of nine different enzymes involved in the biosynthesis of the APE polyene core. For each of the structures, some aspects will be elaborated with a focus on the intermolecular interactions involved in complex formation, binding sites of ACPs and Ppant arms, as well as comparison with ARO type II PKS and type II FAS systems. All of the obtained X-ray crystal structures were published or will be published as listed below:

"An uncommon type II PKS catalyzes biosynthesis of aryl polyene pigments"[35]

G. L. C. Grammbitter, **M. Schmalhofer**, K. Karimi, Y.-M. Shi, T. A. Schöner, N. J. Tobias, N. Morgner, M. Groll, and H. B. Bode
Journal of the American Chemical Society, 141(42):16615–16623, 2019

"The chemical structure of widespread microbial aryl polyene lipids"

G. L. C. Grammbitter*, **M. Schmalhofer***, Y.-M. Shi*, Y.-N. Shi*, S. P. B. Vemulapalli*, C. Richter, H. Schwalbe, M. Alanjary, A. Schüffler, M. Witt,C. Griesinger, M. Groll, and H. B. Bode
*Equal contributing first author. *Manuscript in preparation*, 2023. First draft without structures was published in *bioRxiv*, 2020[38].

Parts of this section's figures, legends, and text parts are "adapted with permission from G. L. C. Grammbitter, M. Schmalhofer, K. Karimi, Y.-M. Shi, T. A. Schöner, N. J. Tobias, N. Morgner, M. Groll, and H. B. Bode. An Uncommon type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments. *Journal of the American Chemical Society*, 141(42):16615–16623, 2019. Copyright 2019 American Chemical Society. https://doi.org/10.1021/jacs.8b10776"[35]

The overall project is based on a collaboration with the group of Prof. Dr. Helge B. Bode and former staff member Dr. Gina L. C. Grammbitter, who kindly provided some of the expression vectors. Planning, design, and execution of all experiments, as well as analysis and interpretation of all generated results, were conducted with help from Prof. Dr. Michael Groll.

2.2.1 Ketosynthases ApeR and ApeO:C

The β-ketoacyl [ACP] synthases ApeR and ApeO:C are responsible for chain elongation of the polyene moiety by catalyzing the Claisen condensation of malonyl building blocks. The X-ray structures of both APE-ketosynthase complexes were obtained from *Xenorhabdus doucetiae* (**Fig. 2.1a** and **Tab. A.3**). *Xd*ApeR$_2$ (2.8 Å resolution) was solved by molecular replacement using the coordinates of an uncharacterized KS from *Paraburkholderia phymatum* (PDB ID: 4EWG[67], SI = 60 %), and *Xd*ApeO:C (PDB ID: 6QSP, 1.45 Å resolution) starting with coordinations of a mitochondrial type II FAS KS (PDB ID: 2IWZ[68], SI = 37.0 % to ApeO).

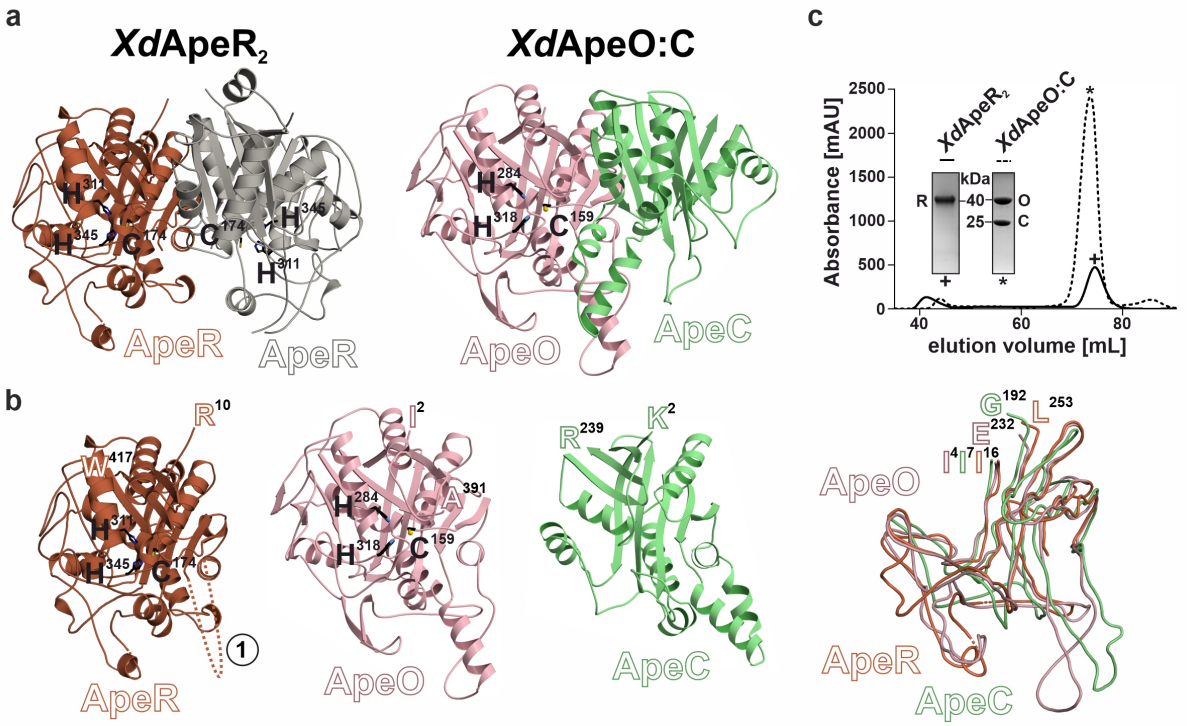

Figure 2.1 Crystal structures of the ketosynthase complexes ApeR$_2$ and ApeO:C from *X. doucetiae*. **(a)** Ribbon representation of the homodimeric ApeR$_2$ (left panel, 2.8 Å resolution) and the heterodimeric ApeO:C complex (right panel, PDB ID: 6QSP, 1.45 Å resolution). ApeR$_2$ harbors two active sites, whereas the catalytic center at the heterodimer is formed exclusively by ApeO. **(b)** Ribbon drawings (from left to right) of ApeR (419 residues), ApeO (391 residues), and ApeC (239 residues), as well as coil representation of the N-terminal domains of ApeO (ile4-glu234) and ApeR (ile16-leu253) structurally imposed onto ApeC (ile7-gly192). All protomers are shown in a similar orientation. Disordered amino acids in region ① (residues 120-153) are highlighted as colored dots. **(c)** Size-exclusion chromatogram and SDS-PAGE analysis show high purity of both ketosynthases. Figure is adapted from Grammbitter et al.[35]

Overall fold and complex formation

Each active ketosynthase subunit, ApeR and ApeO, adopts a double α-β-thiolase fold and shares a similar topology despite their low sequence identity of 28 % (root-mean-square deviation (rmsd) = 1.3 Å over 299 C$_\alpha$-atoms).[69] Notably, the shorter ApeC mimics this topology, as seen in its superposition (residues 7-192) with the N-terminal domains of ApeO (N-domain, residues 4-232, rmsd = 1.9 Å over 118 C$_\alpha$-atoms) and ApeR (residues 16-253, rmsd = 1.9 Å over 103 C$_\alpha$-atoms, **Fig. 2.1b**). In the asymmetric units, ten ApeR protomers

form five homodimeric complexes, whereas one ApeO interacts with one ApeC, which results in an α/β-heterodimeric ketosynthase. This agrees with complementary native electrospray ionization mass spectrometry (nESI-MS) and size-exclusion chromatography (SEC) experiments published in 2019 by Grammbitter et al.[35] The reproduced SEC and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) experiments are depicted in **Figure 2.1c**. Both assemblies show identical arrangements of the respective protomers as the N-terminal domains (N-domains) complement each other in an antiparallel fashion resulting in 10-stranded mixed β-sheets (**Fig. 2.1a** and **Fig. 2.2a**). The formed intermolecular backbone-backbone hydrogen bonds (H-bonds) between the S3 strands of ApeR (exemplary for chain A and chain C, region Ⓐ, **Fig. 2.2a**, left panel) or the S3 strands of ApeO and ApeC (region Ⓐ, **Fig. 2.2a**, right panel) are complemented by salt bridges in the ApeO:C complex (glu134$_{\mathrm{O}}$/arg82$_{\mathrm{C}}$, region Ⓑ and arg164$_{\mathrm{O}}$/glu145$_{\mathrm{C}}$) and in the ApeR$_2$ complex (arg167/glu187). In addition, several hydrophobic clusters facilitate the formation of the complexes. For example, met192$_{\mathrm{O}}$ and leu106$_{\mathrm{O}}$ interact with leu93$_{\mathrm{C}}$, leu96$_{\mathrm{C}}$, and phe107$_{\mathrm{C}}$ in ApeO:C. Furthermore, met250$_{\mathrm{O}}$ forms extensive van der Waals (vdW) interactions with met109$_{\mathrm{C}}$, val111$_{\mathrm{C}}$, and ile121$_{\mathrm{C}}$ (**Fig. 2.2b**). Met250$_{\mathrm{O}}$ is part of a conserved H(M/I)T motif (HMS in ApeO:C, region Ⓒ). This characteristic fold is present in ApeR but is solvent exposed and therefore does not participate in the complex formation (see below).

In the following, the complex interfaces of ApeO:C and ApeR$_2$ will be compared with those formed by homodimeric KSs and heterodimeric KSs from type II FAS and type II PKS systems, respectively. To this end, the online tool bioCOmplexes COntact MAPS (COCOMAPS[70]) was used with a threshold of 6 Å to generate the contact distance plots. Furthermore, PDB in europe Proteins, Interfaces, Structures, and Assemblies (PDBePISA[71]) was applied to calculate the size and stability of the interfaces. The evaluation of the experimentally obtained structural data (**Fig. 2.1**), combined with the contact analysis, shows that the formed contact regions of ApeR$_2$ and ApeO:C are quite similar compared to other KSs (**Fig. 2.2c**). However, residues 120-153 are disordered in all subunits of *Xd*ApeR$_2$ (region ①, **Fig. 2.1b**). This results in a 50 $\%$ reduction of the interplay (1350 Å^2), whereas in FabF and FabB from *E. coli* strain K-12 or ApeR from *Acinetobacter baumannii*, region ① of one subunit interacts with regions ① and ② of the second subunit (**Fig. 2.2c**). As a result, the corresponding computational analysis of FabF and FabB or ApeR calculates interface areas of 2600 Å^2 for *Ec*FabF$_2$ (PDB ID: 6OKG[72]) and 2800 Å^2 for *Ec*FabB$_2$ (PDB ID: 6OKC[72]) and *Ab*ApeR$_2$ (PDB ID: 7F27[73]). Further remarks about disordered regions of *Xd*ApeR$_2$ are presented in **Section 4.2.4**.

The macromolecular interface formed between *Xd*ApeO and ApeC spans 2250 Å^2 and is smaller than type II FAS KS interfaces. Compared to FabF and FabB, the contact region is reduced by 15 $\%$ and 25 $\%$, respectively. This is explained by the smaller size of ApeC (239 residues) compared to the $\sim$400 aa residues long KS protomers ApeO, FabF, FabB, or ActKS. Conspicuously, ApeO:C's interface area deviates even more in comparison with other α/β-heterodimeric KSs of ARO type II PKS (e.g., ActKS:CLF in the biosynthesis of actinorhodin from *Streptomyces coelicolor*, PDB ID: 1TQY[74] or AntKS:CLF in the

biosynthesis of anthraquinone from *Photorhabdus laumondii*, PDB ID: 6SMP[27]). Here, the coordination spheres are increased by 30-40 % to 3200 Å^2 and 2900 Å^2 because ApeC is shortened when compared to other chain-length factors (*Xd*ApeC: 239 residues, ActCLF: 415 residues, AntCLF: 371 residues). Analyzing its structure-based sequence alignment using the Espresso-Server shows the missing regions ③-⑦ (**Fig. A.1a** and **Fig. 2.2c**). While region ⑦ is not involved in the complex formation, regions ③-⑥ are part of the formed interface in *Sc*KS:CLF (**Fig. 2.2c**).

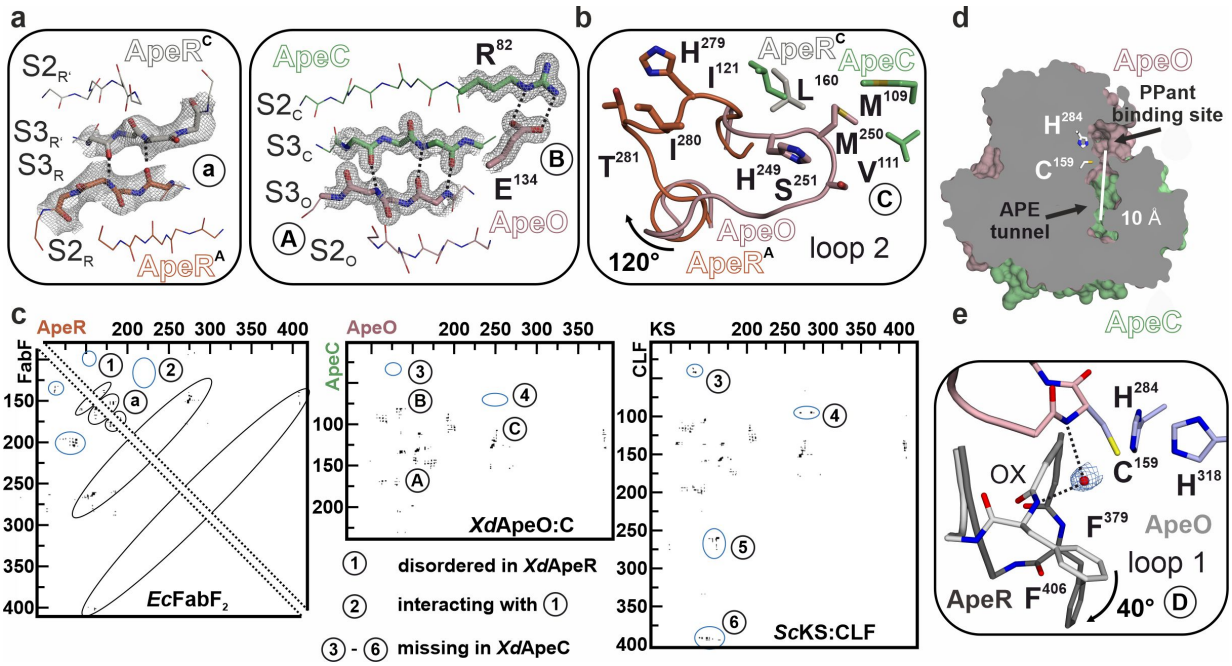

Figure 2.2 The macromolecular interface participates in the specificity pocket and gating. **(a)** Representation of the macromolecular contact regions of *Xd*ApeR$_2$ (left panel) and *Xd*ApeO:C (right panel). The central S3 β-sheets of ApeR (chain A and C, in region Ⓐ) and ApeO:C (in region Ⓐ) are shown as sticks in the 2F$_O$-F$_C$ electron density map (gray mesh, contoured to 1.3σ). The formed H-bonds are highlighted as black dotted lines. **(b)** In ApeO, the conserved H(M/I)T motif (residues his249$_O$, met250$_O$, and ser251$_O$) in loop 2 undergoes extensive vdW interactions with ApeC (residues met109$_C$, val111$_C$, and ile121$_C$). The motif is rotated by 120° and solvent-exposed in the ApeR$_2$ complex (loop 2, residues 275-286, H(M/I)T motif with the residues his279, ile280, and thr281), indicating a role in gating. **(c)** Contact distance plots of *Xd*ApeR$_2$ (left panel), *Xd*ApeO:C (middle panel, PDB ID: 6QSP), and ActKS:CLF (right panel, PDB ID: 1TQY[74]). Interactions formed by one ApeR subunit are substituted with corresponding interactions of *Ec*FabF (PDB ID: 6OKG[72], highlighted by dotted lines). The online tool bioCOmplexes COntact MAPS (COCOMAPS[70]) was used with a threshold of 6 Å to generate the contact distance plots (x- and y-axes show the numbering of the amino acids). **(d)** The sliced surface representation of the heterodimer ApeO:C reveals a prominent entrance pore and a linear APE elongation tunnel (white bar corresponds to 10 Å). **(e)** Atomic view of the active site in *Xd*ApeO:C (blue sticks). The oxyanion hole is formed by the backbone amides of phe379$_O$, located in loop 1 (white sticks and ribbon, residues 377-381), and cys159$_O$. Loop 1 contains the conserved F(A/G)FGG motif (residues 403-407) and is solvent-exposed in *Xd*ApeR (dark-gray ribbon). The 2F$_O$-F$_C$ electron density map is shown as blue mesh and contoured to 1σ with the ligand omitted for phasing. H-bonds are indicated as black dots. Surface representation and atomic view are adapted from Grammbitter et al.[35]

The remaining contact surfaces are conserved despite the reduced length and the missing regions (**Fig. 2.2c**). In addition, the computational analysis using PDBePISA[71] verifies the formation of a stable heterodimeric *Xd*ApeO:C complex with a complexation significance score (CSS) of 0.722 and a gain in solvation free energy (Δ^IG) of -28.3 kcal mol^{-1}.

Specificity pocket and gating

The entrance pore, formed solely by the active KS subunits ApeO and ApeR, is well defined in the $2F_O$-F_C electron density map in both KS structures and points towards the active site. While the polyene elongation tunnel is partly disordered in $ApeR_2$, a specificity pocket could be proposed in ApeO:C. This chamber is predominantly formed by the CLF with a diameter and length of approximately 5 Å and 15 Å, respectively (**Fig. 2.2d**). The active site of both KSs is built exclusively by the active subunits ApeO or ApeR, harboring a nucleophilic cysteine ($cys159_O$ and $cys174_R$) that is coordinated by two histidine residues ($his284_O$, $his318_O$, $his311_R$, and $his345_R$). The catalyzed Claisen-condensation proceeds via an (O^-,S,S)-dithio-orthoester in sp^3-hybridization, which is coordinated by the oxyanion hole (OX) throughout the reaction trajectory.[27] The OX in ApeO:C is formed by the backbone amides of $phe379_O$ and the active site $cys159_O$ (**Fig. 2.2e**).[72,75] So far, this closed-state arrangement occurred in all solved KS structures from ARO and APE type II PKS systems (e.g., *Xd*ApeO:C[35], *Ab*ApeO:C[73], *Ab*$ApeR_2$[73], AntKS:CLF[27], and ActKS:CLF[74]). However, *Xd*$ApeR_2$ presents itself in an open conformation where the oxyanion hole is disrupted (**Fig. 2.2e**). In more detail, loop 1 and 2 rotate by 40° and 120°, respectively (region Ⓓ, Ⓒ, **Fig. 2.2b**, **e**). As a result, both loops are exposed to the surrounding solvent. This gating mechanism has been described on a structural level by Mindrebo and coworkers for the type II FAS KSs *Ec*FabF and *Ec*FabB and is predicted for APE type II PKS KSs based on sequence alignments.[72,75] The closed- and open-state structures *Xd*$ApeR_2$ and *Xd*ApeO:C might indicate a similar gating mechanism for type II FAS and this APE type II PKS (see **Figure A.1b** for sequence alignment of KSs).

Comparison with type II PKS and type II FAS KSs

Comparison of *Xd*ApeO or *Xd*ApeR with active KS-subunits FabB/FabF and ActKS/AntKS of type II FAS and ARO type II PKS systems showed striking similarities (rmsds = 1.1-1.3 Å over 293-333 C_α-atoms). In contrast, the superpositions with *Ab*ApeO or *Ab*$ApeR_2$, the recently solved KSs involved in APE BS from a virulent *Acinetobacter baumannii* strain (PDB ID: 7F28 and 7F27[73]), display slightly higher similarities (*Xd*ApeR:*Ab*ApeR, rmsd = 0.7 Å over 358 C_α-atoms, SI = 60 % and *Xd*ApeO:*Ab*ApeO, rmsd = 1.1 Å over 345 C_α-atoms, SI = 43 %). The chain-length factor *Xd*ApeC is more divergent towards CLFs from ARO type II PKS systems (*Xd*ApeC:*Sc*CLF, rmsd = 1.9 Å over 127 C_α-atoms, SI = 13 % and *Xd*ApeC:*Pl*ApeC, rmsd = 1.8 Å over 116 C_α-atoms, SI = 12 %), while ApeC conforms with *Ab*ApeC (rmsd = 1.5 Å over 165 C_α-atoms, SI = 30 %).

2.2.2 Ketoreductase ApeQ

In the next step, the β-ketoacyl [ACP] reductase *Xd*ApeQ catalyzes the reduction of the keto group in β-position. To obtain the structure of *Xd*ApeQ, different purification and crystallization conditions were screened. However, the crystals only resulted in datasets up to 5 Å resolution. To complete the set of enzymes involved in the APE polyene core biosynthesis,

the X-ray crystal structure of ApeQ from the pathogenic *E. coli* O157:H7 (strain Sakai) was determined at 2.6 Å resolution (**Fig. 2.3a** and **Tab. A.3**). The sequences of the ketoreductase *Ec*ApeQ and the identified ApeQ from *X. doucetiae* share a sequence identity of 82.2 %.

Overall fold and complex formation

Each *Ec*ApeQ subunit adopts the conserved Rossmann fold, which is typical for the superfamily of short-chain dehydrogenases/reductases and consists of seven parallel β-sheets stacked between three helices on each side (**Fig. 2.3b**).[76]

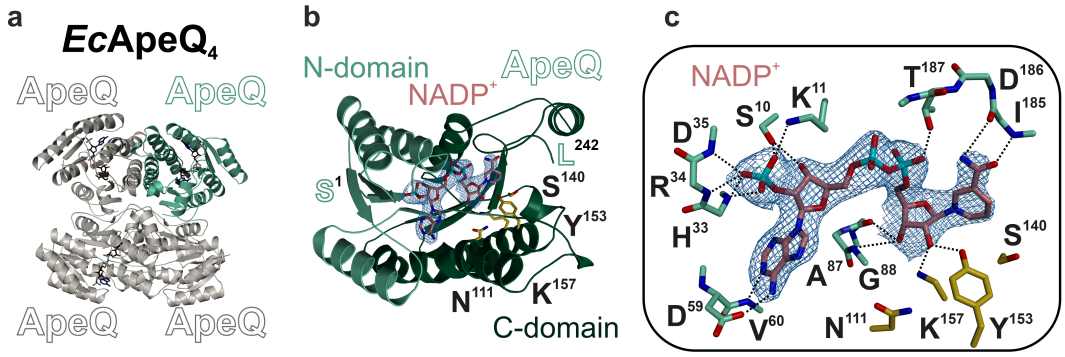

Figure 2.3 Crystal structure of the ketoreductase tetramer ApeQ$_4$ from *E. coli* O157:H7 (strain Sakai). **(a)** Ribbon plot of the ketoreductase with cofactor NADP$^+$ bound to three of the four subunits (2.6 Å resolution). **(b)** Cofactor NADP$^+$ (red sticks) is located between the N-domain (residues 1-91, light green) and the C-domain (residues 92-242, dark green). The active site residues are depicted as yellow sticks. The electron density represents the 2F$_O$-F$_C$ map (blue mesh, contoured to 1.0σ), with the displayed cofactor omitted for phasing. **(c)** Close-up view of the cofactor bound to ApeQ (H-bonds are shown as black dotted lines).

The active site of *Xd*ApeQ is formed by the amino acids asn110, ser140, tyr153, and lys157. Additional signals were observable in the electron density map near the active site. Due to the supplementation of the cofactor during the protein purification and crystallization, this indicates the binding of NADP$^+$ to three subunits, whereas one protomer remains in its apo-state. Additional information about the co-crystallization of *Ec*ApeQ$_4$ with its cofactor is presented in **Section 4.2.4**. The native cofactors perfectly match the 2F$_O$-F$_C$ electron density map and are located between the N-terminal domain (residues 2-91) and the active site. In more detail, the nicotinamide mononucleotide interacts tightly with the active site and loop ile185-thr187 (**Fig. 2.3c**, right), whereas the adenosine-2',5'-bisphosphate forms H-bonding interactions with the N-domain, as seen in the crystal structure (**Fig. 2.3c**, left). As expected, four ApeQ subunits assemble the typical homotetrameric complex of KRs (**Fig. 2.3a** and **Fig. 2.4a**).[35,77] The computational analysis of the oligomer using the PDBePISA tool[71] reveals the formation of two major and stable macromolecular interfaces. The buried surface area of interface I spans 1700 Å (Chain A and C, Δ^IG = -19 kcal mol^{-1}), and interface II spans 1400 Å (Chain A and B, Δ^IG = -21 kcal mol^{-1}). Thereby, the N-terminal halves of each protomer interact predominantly with NADP$^+$, while the remaining protein facilitates the formation of the complex. The contact distance plots show that interface I and II are built

up by the residues 200-243 and 100-175, respectively, and exhibit high similarity with the contact areas of FabG$_4$, the KR in type II FAS from *E. coli* (PDB ID: 1Q7B[78], **Fig. 2.4b**).

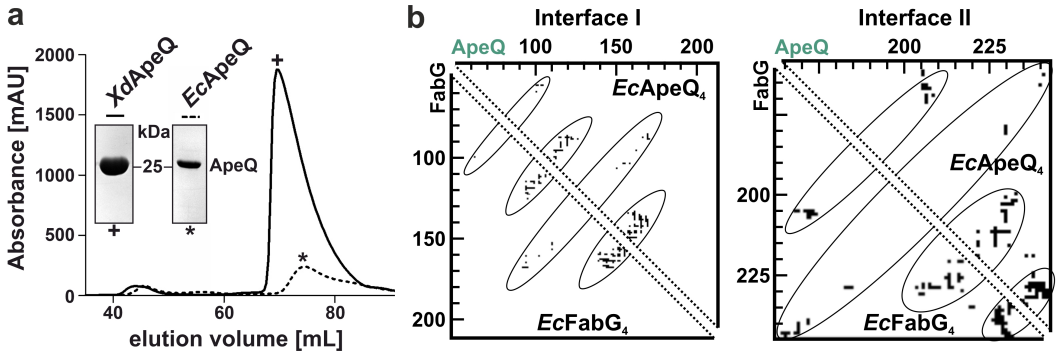

Figure 2.4 Interface regions of ApeQ are conserved. (a) Gel filtration chromatogram and SDS-PAGE analysis of *Xd*ApeQ$_4$ and *Ec*ApeQ$_4$. **(b)** Protein contact plots of interface I (left panel) and interface II (right panel) of *Ec*ApeQ$_4$. Both plots are substituted with corresponding interactions formed by *Ec*FabG$_4$ (PDB ID: 1Q7B[78], highlighted by dotted lines).[70]

Comparison with type II PKS and type II FAS KRs

Comparatively, the solved *Ec*ApeQ structure exhibits a moderate similarity with KRs of type II FAS, e.g., *Ec*FabG with an rmsd of 0.9 Å over 202 Cα-atoms (PDB ID: 1Q7B[78], SI = 41 %) and of ARO type II PKS, e.g., ActKR with an rmsd of 1.3 Å over 224 Cα-atoms (PDB ID: 1X7G[79], SI = 32 %). As expected, the structural similarity to the recently published ApeQ from *Acinetobacter baumannii* (PDB ID: 7CAW[80], SI = 61 %) is even higher (rmsd = 0.7 Å over 235 Cα-atoms). Most strikingly, the *Ec*ApeQ complex structure matches perfectly with a putative KR from *E. coli* O6:H1 (PDB ID: 4IIU[81], SI = 99 %, rmsd = 0.3 Å over 235 Cα-atoms, to be published), which was used as a model for molecular replacement. The analysis of *E. coli* O6:H1 using the antiSMASH[82] bacterial version identifies this KR as a yet unassigned ApeQ that is part of an APE type II PKS system.

2.2.3 Dehydratases ApeI:P

The last step in the iterative BS of the carbon-skeleton backbone and its processing into the APE core moiety is the dehydration reaction of the β-hydroxyacyl group. It is catalyzed by the β-hydroxyacyl [ACP] dehydratase ApeI:P. The structure of ApeI:P is fascinating because of its untypical heterodimeric complex formation identified by Grammbitter and coworkers in 2019 using nESI-MS, pull-down, and SEC experiments.[35] The X-ray crystal structure of the *Xd*ApeI:P complex was solved at a resolution of 1.85 Å (PDB ID: 6QSR, **Fig. 2.5a** and **Tab. A.3**).

Overall fold and complex formation

Both DH subunits adopt a hotdog fold consisting of a central helix H1 surrounded by six antiparallel β-sheets (in the order of S1↓, S2↑, S4↓, S5↑, S6↓, and, S3↑).[83] Additionally, ApeP is extended by the two short β-strands S2a and S3a and a C-terminal helix H2 (**Fig. 2.5b**). Despite their low sequence identity of 12 %, both protomers share a similar topology (rmsd = 1.6 Å over 73 Cα-atoms, **Fig. 2.5b**, **c**).

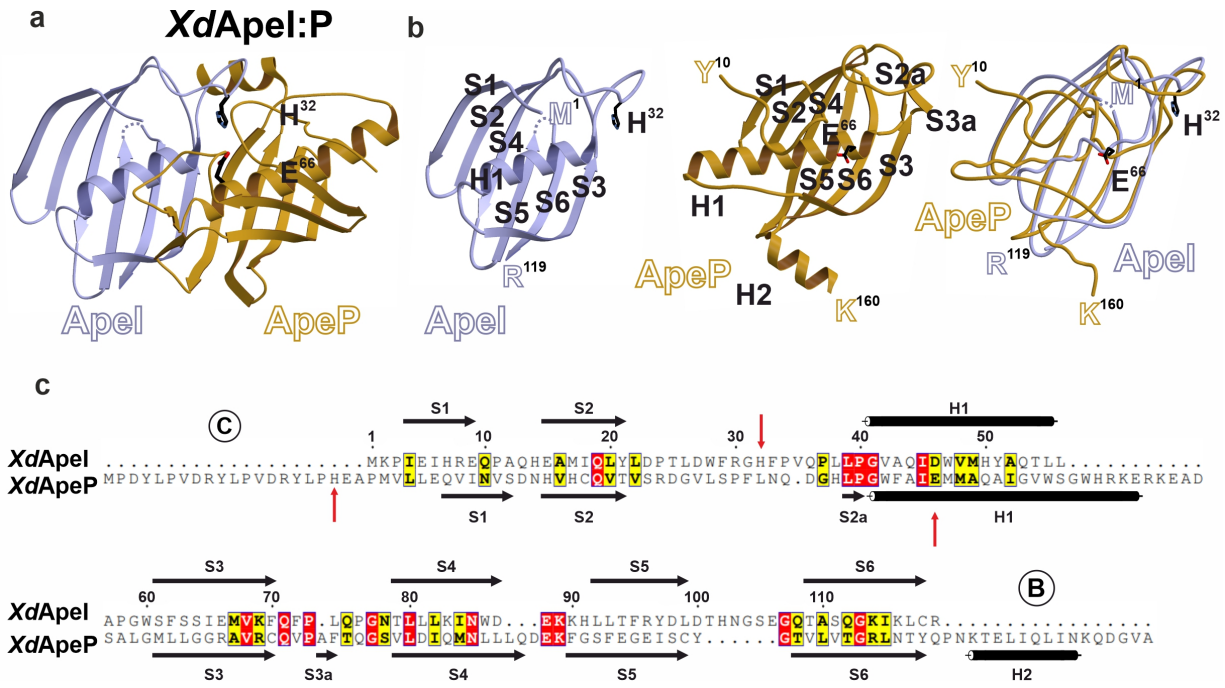

Figure 2.5 Crystal structure of the dehydratase complex ApeI:P from *X. doucetiae*. **(a)** Ribbon plot of the heterodimeric *Xd*ApeI:ApeP (PDB ID: 6QSR, 1.85 Å resolution). **(b)** Ribbon drawings of ApeI (119 residues) and ApeP (165 residues, from left to right) and coil representation of ApeI structurally imposed onto ApeP. The typical hotdog fold of both subunits is shown in similar orientations. **(c)** Both subunits have unique sequences (SI = 12 %), although their domains share a common topology. Secondary structure elements are indicated by black arrows (S, strand) and bars (H, helix). Red arrows highlight His32 and Glu66. ApeI is N- and C-terminally shortened (region Ⓑ and region Ⓒ). Figure is adapted from Grammbitter et al.[35]

As expected, the SEC experiments and the X-ray crystal structure of *Xd*ApeI:P confirm the dimeric oligomer (**Fig. 2.5a** and **Fig. 2.6a**). Its computational analysis using the PDBePISA tool[71] corroborates the formation of a stable complex with a CSS of 0.727 (Δ^IG = -22.2 kcal mol^{-1}). The formed interface spans 1360 Å^2 in total and is comparable to the

interfaces of homodimeric DHs (e.g., *Ec*FabZ$_2$, the type II FAS DH from *E. coli*, PDB ID: 6N3P[84], 1370 Å^2, Δ^IG = -23.5 kcal mol^{-1}, **Fig. 2.6b**). Like FabZ$_2$ and the ketosynthase dimers, *Xd*ApeR$_2$ and *Xd*ApeO:C, the *Xd*ApeI:P complex is mainly stabilized by the alignment of the β-strands forming a 12-stranded sheet (**Fig. 2.5a** and **Fig. 2.6b**). The interacting S3 β-sheets orient in an antiparallel fashion that results in the formation of seven interstrand backbone-backbone H-bonds (S3$_\mathrm{I}$: ser63$_\mathrm{I}$, ile65$_\mathrm{I}$, glu66$_\mathrm{I}$, val68$_\mathrm{I}$, phe70$_\mathrm{I}$ with S3$_\mathrm{P}$: gly96$_\mathrm{P}$, ala98$_\mathrm{P}$, val99$_\mathrm{P}$, leu94$_\mathrm{P}$, cys101$_\mathrm{P}$) and one salt bridge involving the side chains of glu66$_\mathrm{I}$ and arg100$_\mathrm{P}$ (region Ⓐ, **Fig. 2.6b**, **c**). Furthermore, the stability of the heterodimeric complex is based upon a compact and hydrophobic core located in the center of ApeI:P between the aligned S3 strands and the overlapping H1 helices. In particular, an extended network of vdW interactions is built up by the side chains of S3$_\mathrm{I}$: phe62, ile65, val68, phe70 and H1$_\mathrm{I}$: val42, ile45, met49 with S3$_\mathrm{P}$: leu93, val99, cys101 and H1$_\mathrm{P}$: trp62, phe63, ile65 (**Fig. 2.6b**). A third interaction patch can be mapped to the helix H2$_\mathrm{P}$ (interacting residues: lys151$_\mathrm{P}$, leu154$_\mathrm{P}$, leu157$_\mathrm{P}$, ile158$_\mathrm{P}$) and the loop before H1$_\mathrm{I}$ (interacting residues: phe33$_\mathrm{I}$, val35$_\mathrm{I}$, gln36$_\mathrm{I}$, leu38$_\mathrm{I}$, region Ⓑ, **Fig. 2.6b**, **c**). Most strikingly, this helix H2 is absent in ApeI and FabZ (**Fig. 2.5c** and **Fig. 2.7**). Additionally, ApeI is N-terminally shortened, resulting in the absence of region Ⓒ (**Fig. 2.6b** and **Fig. 2.7**). These key players in complex formation are further confirmed by computational analysis using DrugScorePPI[85].

An elongated binding cavity contains the active site

In addition to their role in complex formation, the S3 strands of both protomers are part of the entrance pore, which accepts the incoming Ppant arm as described for ACP:DH complexes.[84,86,87] This formed Ppant binding tunnel ($\approx 7 \times 9 \times 15$ Å^3) connects the surrounding solvent with pro40$_\mathrm{I}$ and the active site (**Fig. 2.6d**). After a 90° kink, the entrance pore is elongated into the hydrophobic polyene binding cavity. In contrast to homodimeric DHs, ApeI:P harbors only one active site that is located at the interface of both subunits. The active site consists of glu66$_\mathrm{P}$ and two histidines (first: his32$_\mathrm{I}$ and second: his19$_\mathrm{P}$, **Fig. 2.5a** and **Fig. 2.6e**). While glu66$_\mathrm{P}$ is H-bonded with a water molecule that is well defined in the electron density map, his32$_\mathrm{I}$ forms a non-prolyl *cis* peptide bond with phe33$_\mathrm{I}$. As a result, the imidazole side chain of his32$_\mathrm{I}$ is locked between pro40$_\mathrm{I}$ and phe33$_\mathrm{I}$ (**Fig. 2.6d**, **e**). These arrangements are also found in FabZ and FabA (second histidine is absent).[88] An X-ray crystal structure of FabA covalently bound to 2-decenoyl-*N*-acetylcysteamine shows his70 and asp84 (correspond to his32$_\mathrm{I}$ and glu66$_\mathrm{P}$ in ApeI:P) near the CHα moiety and β/γ-carbon atoms of the bound mimic, respectively.[83] These insights indicate that glu66$_\mathrm{P}$ and his32$_\mathrm{I}$ act as Brønsted acid and base, respectively. This was further confirmed by nuclear magnetic resonance (NMR) spectroscopy for FabA, showing a deprotonated histidine with a pK_a value below five.[89] Taken together, this allows ApeI:P to catalyze the deprotonation in α-position and reprotonation in β-position, which leads to the elimination of water through acid-base catalysis (**Fig. 2.6e**). As a result, the β-hydroxy thioester is converted to an α,β-*trans*-unsaturated thioester product.

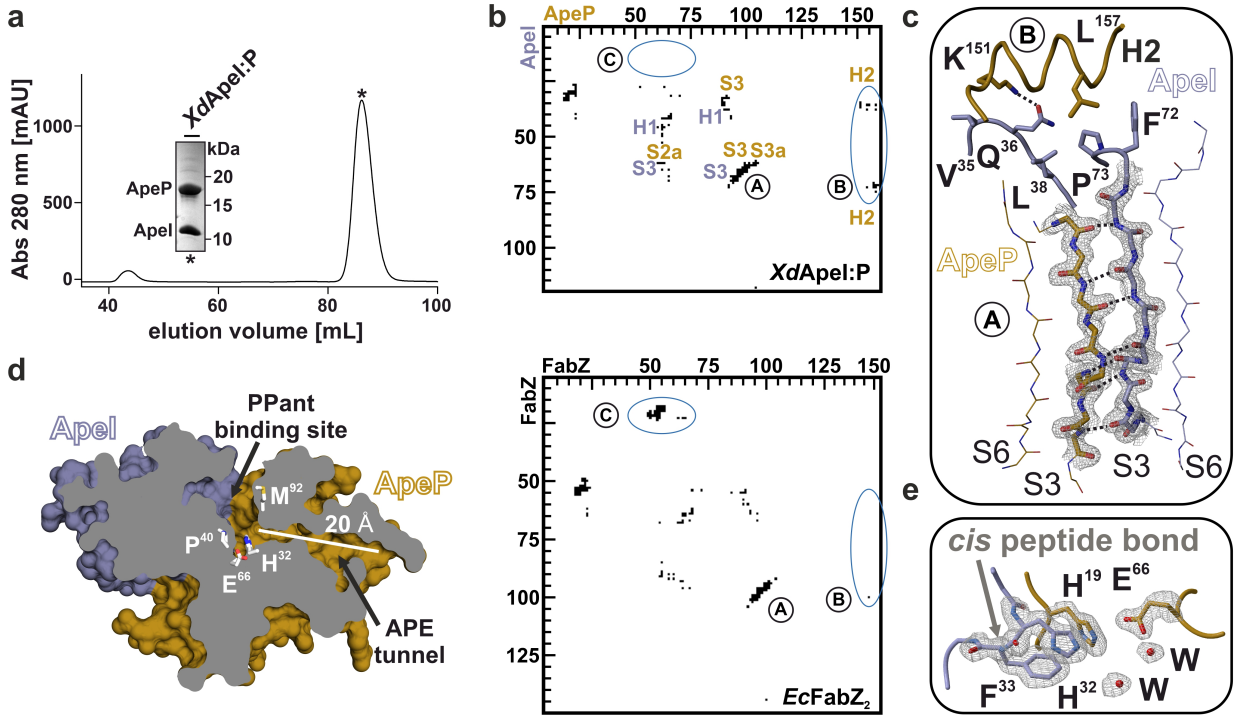

Figure 2.6 The interplay of ApeI and ApeP facilitates the Ppant and polyene binding tunnels. **(a)** Gel filtration and SDS-PAGE analysis confirm heterodimeric complex and high purity. **(b)** Protein contact plots of *Xd*ApeI:P (upper panel) and homodimeric *Ec*FabZ$_2$ (lower panel, PDB ID: 6N3P[84]).[70] The interface of the heterodimeric DH complex shows extensive interaction between both S3 β-strands of ApeI and ApeP (region Ⓐ) and H2 of ApeP (region Ⓑ). Region Ⓒ is partly absent in this complex. **(c)** Representation of the macromolecular interface of ApeI:P. The central β-sheets in the region Ⓐ are shown as sticks in the 2F$_O$-F$_C$ electron density map (gray mesh, contoured to 1.3σ). Dotted lines highlight the formed H-bonds. **(d)** The sliced surface representation illustrates an elongated and narrow APE tunnel (the white bar represents 20 Å). **(e)** Stick representation of the proton relay shuttle. A gray arrow highlights the *cis* peptide bond between His32 and Phe33 in ApeI. The 2F$_O$-F$_C$ electron density map is shown as gray mesh and contoured to 1σ with the amino acid residues omitted for phasing. Representation of the surface and active site are adapted from Grammbitter et al.[35]

Comparison with type II PKS and type II FAS DHs

Due to the low sequence similarity of ApeI:P to known type II FAS/PKS DHs (**Fig. 2.7**), a search for homologous proteins was performed based on their 3D structures using the online tool protein structure comparison server (DALI[90]). The best hits for ApeI and ApeP are DHs from type II FAS systems, e.g., FabZs from *Helicobacter pylori* (PDB ID: 2GLL[91], Z-score = 19.6, rmsd = 1.3 Å over 118 Cα-atoms, SI = 17 % to ApeP and Z-score = 12.8, rmsd = 1.3 Å over 83 Cα-atoms, SI = 24 % to ApeI), *Plasmodium falciparum* (PDB ID: 1Z6B[92], Z-score = 18.5, rmsd = 1.2 Å over 84 Cα-atoms, SI = 24 % to ApeP and Z-score = 14.4, rmsd = 1.3 Å over 115 Cα-atoms, SI = 20 % to ApeI), or *E. coli* (PDB ID: 6N3P[84], Z-score = 18.8, rmsd = 1.2 Å over 82 Cα-atoms, SI = 20 % to ApeP and Z-score = 14.6, rmsd = 1.4 Å over 114 Cα-atoms, SI = 17 % to ApeI). In the case of ApeI, the protein with the highest structural similarity was an uncharacterized protein from *Pectobacterium atrosepticum* (PDB ID: 3ESI[93], Z-score = 10.2, rmsd = 1.7 Å over 69 Cα-atoms, SI = 14 % to ApeP and Z-score = 19.4, rmsd = 0.7 Å over 101 Cα-atoms, SI = 54 % to ApeI). Intrigued by the high sequence similarity to ApeI, the online tool antiSMASH[82] bacterial version was used to an-

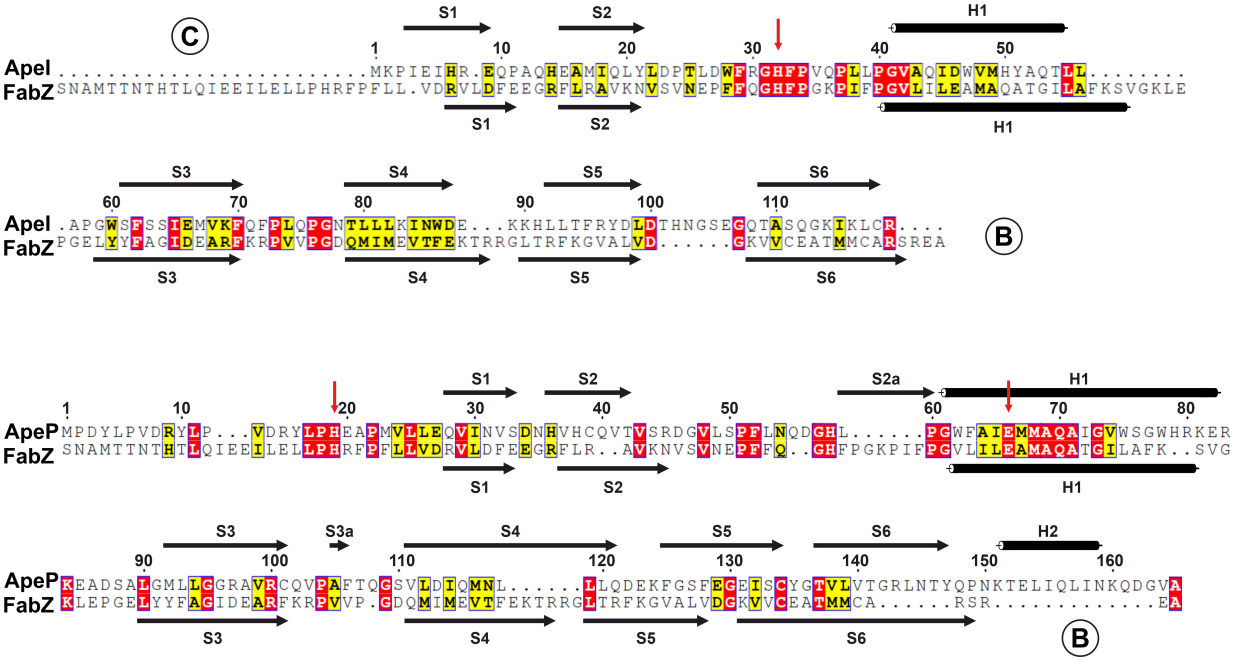

Figure 2.7 Low sequence similarity of ApeI and ApeP towards *Ec*FabZ. Sequence alignment of ApeI (upper panel) and ApeP (lower panel) with FabZ, the type II FAS DH from *E. coli* (red arrows indicate active site residues). ApeI is N-terminally shortened (region Ⓒ), and Helix H2 is missing in ApeI and FabZ (region Ⓑ).

alyze the BGCs from the host organism. As a result, 3ESI was assigned as the respective ApeI in an *ape* BGC from *P. atrosepticum*. This BGC contains a gene for ApeP as well.

2.2.4 ApeK and the acyl carrier proteins ApeE and ApeF

Studies on ApeK were of high interest because i) its exact function is unknown, ii) the enzyme forms a stable tetramer, and iii) it interacts with two different acyl carrier proteins, ApeE and ApeF, that the *ape* BGC encodes. In 2019, Grammbitter and coworkers reported the formation of this tetrameric *Xd*ApeK$_4$ complex (65 kDa) that forms the two 100 kDa assemblies in the presence of the ACPs (MW $\approx$ 9.5 kDa), respectively.[35] Thereby, the heterooligomeric complexes are striking since *Xd*ApeE and *Xd*ApeF exhibit a limited sequence identity (SI = 31.6 %, SS = 70.2 %, **Fig. 2.8a**). To clarify these gaps, stability assays were performed, and the X-ray structures of homooligomeric *Xd*ApeK$_4$, and the heterooligomeric complexes *Xd*ApeK$_4$:E$_4{}^{holo}$ and *Xd*ApeK$_4$:F$_4{}^{holo}$, were determined.

The thioesterase ApeK$_4$ forms two stable heterooctameric complexes with ApeE and ApeF

Since the formation of stable protein:ACP complexes is uncommon, as these complexes typically are only based upon transient interactions, the stability of the heterooctameric complexes was analyzed.[94] Therefore, *Xd*ApeK$_4$:F$_4{}^{holo}$ was applied to an analytical SEC column where both proteins elute as the native complex at pH 8.0 (retention volume (V$_r$) = 12.75 mL, black line, **Fig. 2.8b**). With decreasing pH values (5.0 and 4.7, darker blue lines), the complex shifts towards ApeF (V$_r$ $\approx$ 17 mL, the same volume as a control sample, dotted black line)

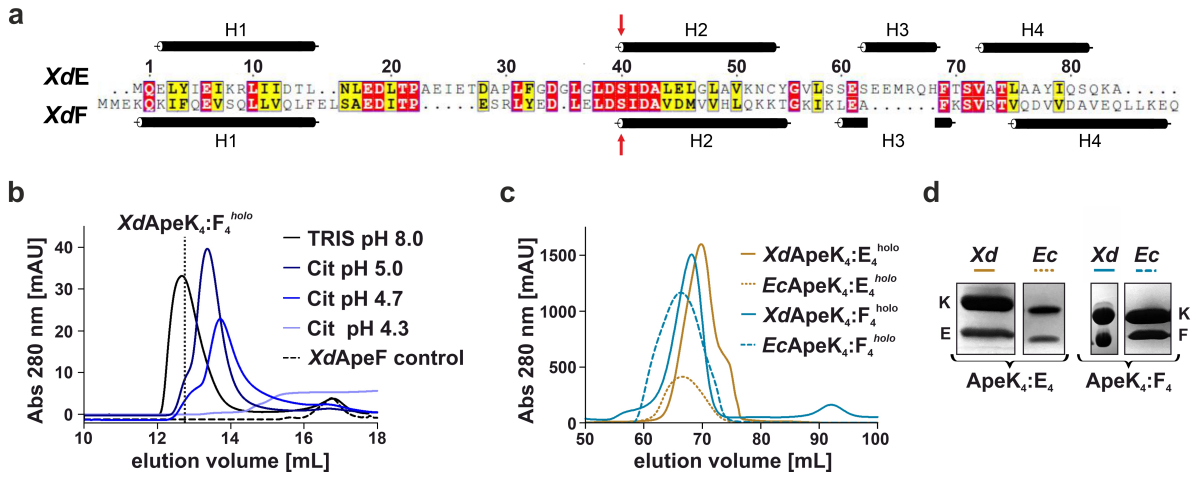

Figure 2.8 Heterooctameric complex assemblies are a feature across species. (a) Sequence alignment of the two ACPs ApeE and ApeF from *X. doucetiae* (red arrows indicate conserved Ppant-binding site). **(b)** Analytical gel filtration experiments show pH-dependent disassembly of the *Xd*ApeK$_4$:F$_4{}^{holo}$ complex. TRIS and Cit (citric acid) correspond to buffers C and C-Cit, respectively (see **Table 4.17**). The dotted line corresponds to the native *Xd*ApeK$_4$:F$_4{}^{holo}$ complex with a retention volume of 12.75 mL. **(c)** Size-exclusion chromatography and **(d)** consecutive SDS-PAGE analysis identify stable complexes of *Xd*ApeK$_4$:E$_4{}^{holo}$ and *Xd*ApeK$_4$:F$_4{}^{holo}$ in *X. doucetiae* as well as in *E. coli* O157:H7 (strain Sakai).

and ApeK$_4$ ($V_r \approx 13.75$ mL). In addition, smaller peaks are observable at retention volumes between 12 mL and 17 mL. These signals, most likely, indicate mixtures of the partly disassembled complexes.[35] The complex falls apart at a pH value of 4.3 (light blue line). Next, the appearance of these two complexes was investigated in species other than *X. doucetiae*. Therefore, the homologous ApeK:E and ApeK:F proteins from *E. coli* O157:H7 (strain Sakai, **Tab. A.1**) were generated, which allowed the identification of similar complex formations in pull-down and SEC experiments (**Fig. 2.8c**), combined with respective SDS-PAGE analysis (**Fig. 2.8d**). In conclusion, these heterooctameric complexes are, most likely, a conserved feature within the *ape*$_{Xd}$ and *ape*$_{Ec}$ BGC subfamily.

Overall fold and complex formation

Since no structures of similar ACP:protein complexes are available, the structure of *Xd*ApeK$_4$ was determined to 2.6 Å resolution (**Fig. 2.9a** and **Tab. A.4**). A 1,4-dihydroxy-2-naphthoyl-CoA thioesterase from *Staphylococcus aureus* (*Sa*DHNA-CoA TE, PDB ID: 6FDG[95], SI = 31 %, and SS = 72.6 %), involved in vitamin K2 biosynthesis, could be used successfully as a model for molecular replacement. Each of the ApeK protomers adopts a hotdog fold.[83] Here, the central helix H1 is surrounded by five antiparallel β-sheets and two additional helices: H2, located at the end of H1, and the C-terminal helix H3. ApeK forms a tetrameric complex and is analyzed below in more detail.

In addition, the crystal structures of the *Xd*ApeK tetramer were determined in complex with the acyl carrier proteins *Xd*ApeE or *Xd*ApeF (*Xd*ApeK$_4$:F$_4{}^{holo}$, 2.45 Å resolution and *Xd*ApeK$_4$:E$_4{}^{holo}$, 2.3 Å resolution, **Fig. 2.9b** and **Tab. A.4**). *Xd*ApeK$_4$:F$_4{}^{holo}$ was solved by molecular replacement using the *Xd*ApeK$_4$ structure combined with an ACP from *Ther-*

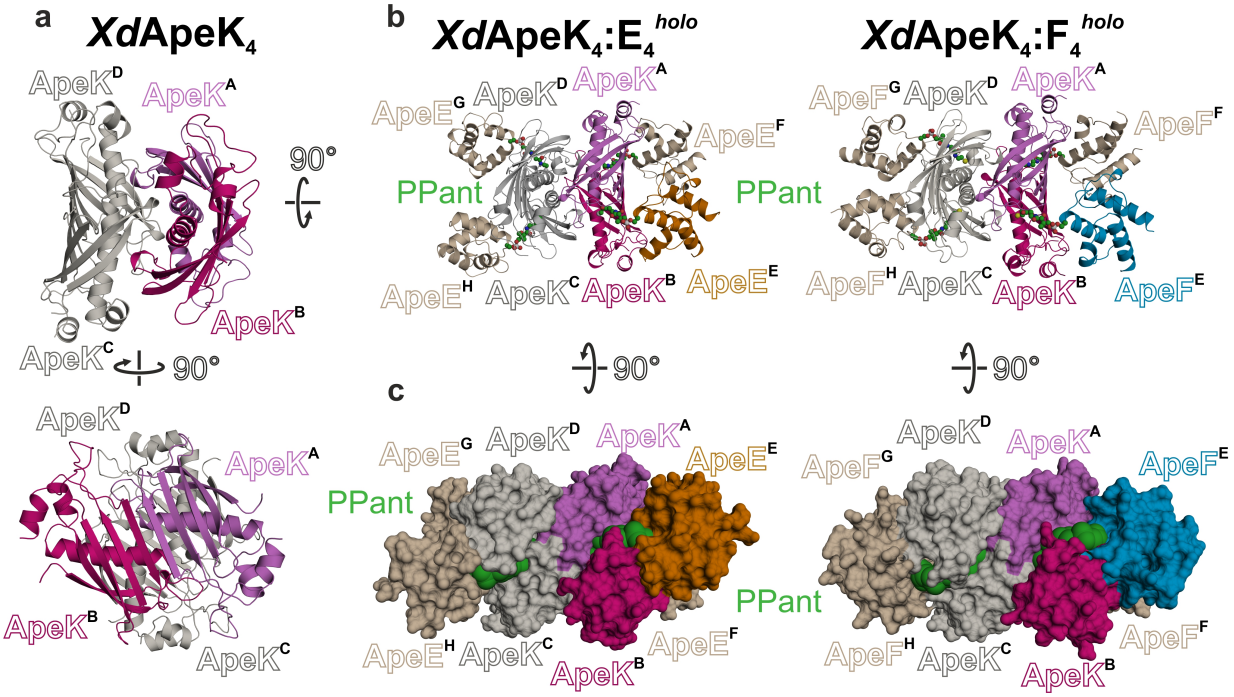

Figure 2.9 Crystal structure of the thioesterase tetramer ApeK$_4$ alone and the heterooctamers *Xd*ApeK$_4$:E$_4{}^{holo}$ and *Xd*ApeK$_4$:F$_4{}^{holo}$ in presence of the two acyl carrier proteins from *X. doucetiae*. **(a)** Side view (upper panel) and top view (lower panel) of *Xd*ApeK$_4$ (2.6 Å resolution) and **(b)** ApeK$_4$ bound to the acyl carrier proteins forming *Xd*ApeK$_4$:E$_4{}^{holo}$ (2.3 Å resolution) and *Xd*ApeK$_4$:F$_4{}^{holo}$ (2.45 Å resolution) complexes. **(c)** Surface rendering of *Xd*ApeK$_4$:E$_4{}^{holo}$ and *Xd*ApeK$_4$:F$_4{}^{holo}$. Ppants are shown as green spheres.

mus thermophilus (PDB ID: 1X3O[96], SI = 33 % to ApeF), whereas coordinates predicted by AlphaFold[97] enabled phases for *Xd*ApeK$_4$:E$_4{}^{holo}$. The quality of the electron density map and the structural data is discussed in **Section 4.2.4**.

*Xd*ApeF and *Xd*ApeE adopt an identical ACP fold that includes the H1, H2, and H4 helices and the more flexible H3 (**Fig. 2.9b**).[98] Despite their low sequence similarity (SI = 31.6 % and SS = 70.2 %, **Fig. 2.8a**), both ACPs are structurally similar to each other (rmsd = 1.5 Å over 67 Cα-atoms) as well as to complexed ACPs from type II FAS (e.g. AcpP, PDB ID: 6N3P[84], Z-score = 11.8, SI = 27.1 %, rmsd = 1.4 Å over 65 Cα-atoms to ApeF and Z-score = 10.9, SI = 26.6 %, rmsd = 1.4 Å over 71 Cα-atoms to ApeE) or type II PKS (e.g. AntF, PDB ID: 6SMP[27], Z-score = 11.7, SI = 17.6 %, rmsd = 1.2 Å over 62 Cα-atoms to ApeF and Z-score = 10.9, SI = 24.1 %, rmsd = 1.3 Å over 68 Cα-atoms to ApeE). Computational analysis of the ApeK structure using PDBePISA[71] suggests the formation of a stable tetrameric assembly (Δ^IG = -54 kcal mol^{-1}), consistent with published results based on nESI-MS and SEC experiments.[35] Taken together, ApeK$_4$ consists of the two homodimers ApeK$^{\mathrm{A}}$:K$^{\mathrm{B}}$ (chain A and B) and ApeK$^{\mathrm{C}}$:K$^{\mathrm{D}}$ (chain C and D) that arrange in an ApeK$^{\mathrm{AB}}$:K$^{\mathrm{CD}}$ complex following D2-symmetry (**Fig. 2.9a**). When using ApeK$^{\mathrm{A}}$ (chain A) as exemplary origin, three interfaces can be assigned: interface I (ApeK$^{\mathrm{A}}$:K$^{\mathrm{B}}$ with a total protein surface area of 840 Å^2), interface II (ApeK$^{\mathrm{A}}$:K$^{\mathrm{C}}$ with a total protein surface area of 650 Å^2), and interface III (ApeK$^{\mathrm{A}}$:K$^{\mathrm{D}}$ with a total protein surface area of 410 Å^2, **Fig. 2.10a**). To gain further insights into the ApeK$_4$ complex, each interface was analyzed. Similar to the previously

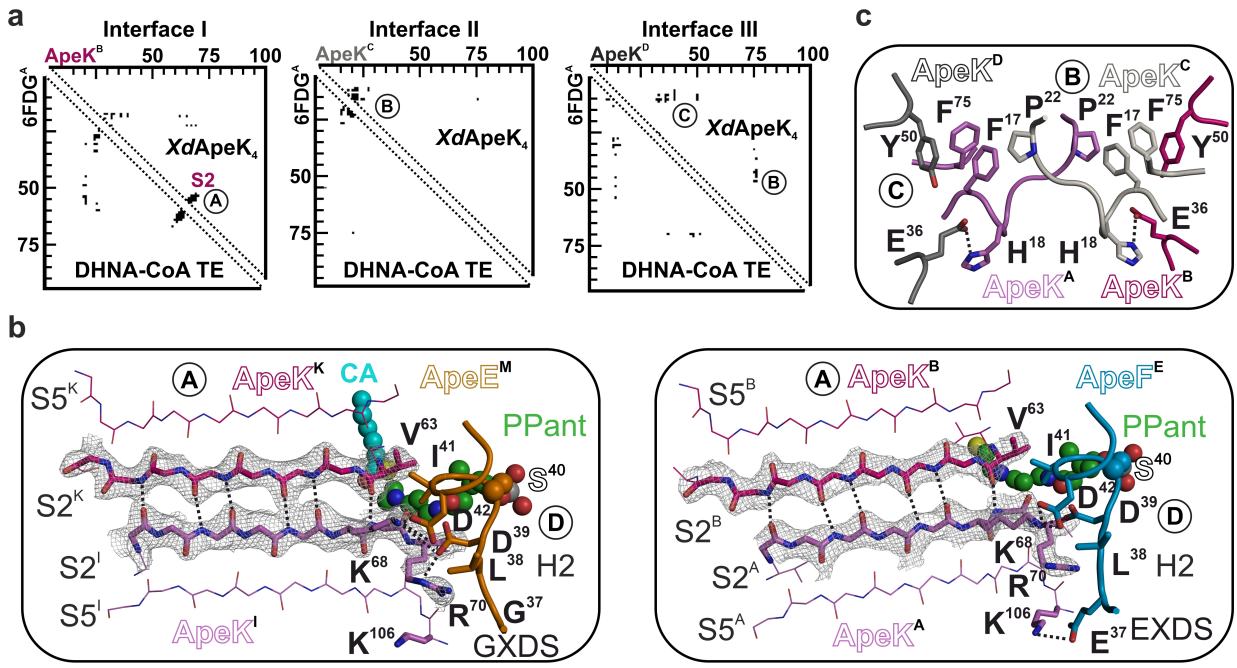

Figure 2.10 Macromolecular interface formed by ApeK, ApeE, and ApeF. (a) Protein contact plots of ApeK^{A}, ApeK^{B}, and ApeK^{C}. Plots are substituted with corresponding interactions formed by *Sa*DHNA-CoA TE$_4$ (PDB ID: 6FDG[95], highlighted by dotted lines).[70] **(b)** Representation of the macromolecular interface of ApeK^{AB}:E (left panel) and ApeK^{AB}:F (right panel). The central β-sheets (S2^{A} and S2^{B}) in the region Ⓐ are shown as sticks in the 2F$_O$-F$_C$ electron density map (gray mesh, contoured to 1.3σ). The formed H-bonds are highlighted as dotted lines, and Ppant is depicted as green spheres. **(c)** Atomic view of one of the two ApeK^{A}:K^{B}:K^{C}:K^{D} cross sections shows an extensive network of vdW and H-bonding interactions (regions Ⓑ and Ⓒ).

described complex structures ApeR$_2$, ApeO:C, and ApeI:P, ApeK^{A}:K^{B}-homodimer formation is predominantly facilitated by antiparallel alignment of the strands S2^{A} and S2^{B}, which results in an extended 10-stranded sheet with a network of backbone-backbone H-bonding interactions (region Ⓐ, **Fig. 2.10b**). The two homodimers, ApeK^{A}:K^{B} and ApeK^{C}:K^{D}, are stacked perpendicularly onto each other and are interlocked at two cross sections. One is shown in **Figure 2.10c** and consists of extensive vdW (region Ⓑ, tyr50^{D}, phe75^{A}, phe17^{A}, pro22^{C} and pro22^{A}, phe17^{C}, phe75^{C}, tyr50^{B} are most striking) as well as H-bonding interactions (region Ⓒ, glu36^{D}, his18^{A} and glu36^{B}, his18^{C}). The overall interface is highly similar to *Sa*DHNA-CoA TE$_4$ (PDB ID: 6FDG[95], **Fig. 2.10a**). Each ApeK homodimer (ApeK^{AB} and ApeK^{CD}) is coordinated by two ACPs, which results in the formation of the *Xd*ApeK$_4$:E$_4{}^{holo}$ and *Xd*ApeK$_4$:F$_4{}^{holo}$ complexes (**Fig. 2.9b**), that are summarized in the following.

Structural superposition of the unbound and ACP-complexed states show no major rearrangements in ApeK$_4$ upon binding of ApeE (rmsd = 0.8 Å over 522 Cα-atoms) or ApeF (rmsd = 0.7 Å over 526 Cα-atoms). The ACP^{E}s (chains E used as an example) interact with both ApeK protomers of one ApeK^{AB} homodimer, and the overall binding regions can be mapped to the helices H2 and H3 in *Xd*ApeE or *Xd*ApeF and several strands and loops of ApeK^{AB}, including the strand S2^{A} and loop L3^{A} (loop after S3) of ApeK^{A} and S2^{B}, S5^{B}, L3^{B}, and L5^{B} of ApeK^{B} (**Fig. 2.11a**).

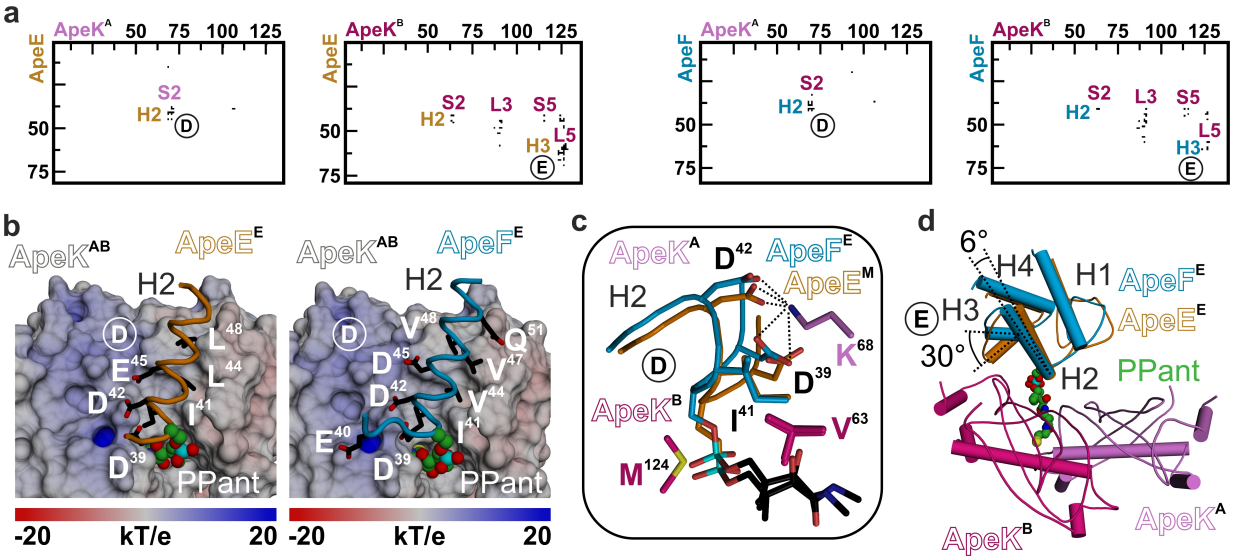

Figure 2.11 Complex formation relies on helix H2 and the modified serine. (a) Protein contact plots show the interaction of ApeK^{A} and ApeK^{B} with the ACPs (*Xd*ApeE, left panel and *Xd*ApeF, right panel).[70] **(b)** Surface representation of the electropositive and hydrophobic binding patch of ApeK that binds the helix H2 of the ACPs. The interacting residues are shown as sticks and the Ppant as spheres. **(c)** Structural overlay of ApeE and ApeF bound to ApeK reveals an ideal match of protein side chain residues and Ppant (rmsd = 1.5 Å over 67 Cα-atoms, SI = 31.6 %). **(d)** Superposition of ApeK$_4$ shows shifted helices H2 and H3.

As seen in the contact distance plots, the interface areas formed with ApeK^{A} are small (ApeE^{E}: 138 Å^2 and ApeF^{E}: 176 Å^2), while the interfaces formed with ApeK^{B} are significantly increased in their sizes (ApeE^{E}: 520 Å^2 and ApeF^{E}: 481 Å^2). Nevertheless, prominent interactions are formed in the ACP^{E}:ApeK^{A} interfaces, including the recognition helices H2 of the ACPs that harbor several acidic and hydrophobic amino acids. The formed electronegative/hydrophobic patches are typical for ACPs and complement the electropositive/hydrophobic binding sites of ApeK^{AB} that enclose the helices H2, as seen in the calculated electrostatic surface potentials of ApeK (**Fig. 2.11b**).[99,100] The amino acids in H2 are mostly conserved for ApeE and ApeF (**Fig. 2.8a**) resulting in formation of nearly identical interactions (region Ⓓ, **Fig. 2.11c**). The key interactions are formed between the ACPs (H2: asp39^{E}, asp42^{E}) and ApeK^{A} (S2: lys68^{A}, arg70^{A}). The complex is further stabilized by an isoleucine (ile41^{E}) that pierces into a hydrophobic cavity formed by ApeK^{AB} (lys68^{A}, arg70^{A}, val63^{B}, val115^{B}, and met124^{B}). The bigger ACP^{E}:ApeK^{B} interfaces contain additional hydrophobic interactions formed between the helices H3^{E} (ApeE^{E}: met65^{E}, arg66^{E}, and phe69^{E} or ApeF^{E}: leu60^{E}, phe63^{E}, and lys64^{E}) and L5^{B} (ApeK^{B}: phe126^{B} and val127^{B}, region Ⓔ). Conspicuously, the structural superposition of ApeK$_4$ in the *Xd*ApeK$_4$:E$_4{}^{holo}$ and *Xd*ApeK$_4$:F$_4{}^{holo}$ complexes shows that one ACP is tilted by 6 ° and shifted approximately 3 Å. Remarkably, helix H3 is rotated by 30 °, which explains the deviations in the interacting amino acids located in region Ⓔ and the transient complex formation (**Fig. 2.11a**, **d**).

The Ppants as prosthetic groups and the hydrophobic polyene tunnel

The Ppants are oriented along the ApeK^{AB} interfaces. While the phosphate groups are solvent-exposed and mobile, the pantothenic moieties and cysteamine residue are buried

deep inside the protein complex (**Fig. 2.9b**, **c**). As a result, the cofactor significantly contributes to the binding, as has been described for other protein:Ppant oligomers.[27] Taken together, computational analysis using PDBePISA[71] shows that the Ppants account for 406 Å^2 and 403 Å^2 of the contact area between ApeK^{AB} and ApeE^{holo} or ApeF^{holo}, respectively. This corresponds to an increase in the interface areas by 60 %. The prosthetic group fits perfectly in the electron density map and adopts identical conformations in both complex structures and all subunits, indicating rigid and tight binding (**Fig. 2.12a**). Thereby, the nitrogen and oxygen atoms of the ligand form strong H-bonding interactions with the backbone atoms of tyr69^A, arg70^A, ser72^A, and val62^B, as well as with a highly coordinated water molecule. Additionally, the aliphatic moieties of val25^A, his28^A, tyr69^A, ser72^A, pro61^B, val63^B, val115^B, val117^B, and met124^B form hydrophobic interactions with the Ppant residue and direct the sulfur atom toward the active sites of ApeK^{AB} (**Fig. 2.12a**).

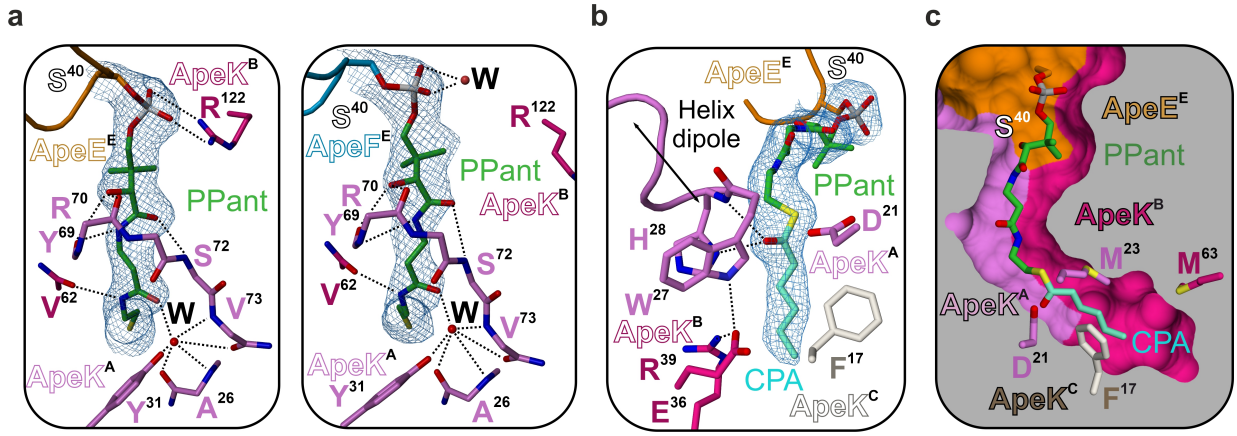

Figure 2.12 The prosthetic group Ppant is a prerequisite for complex formation and catalytic activity. **(a)** Interactions of Ppant in the heterocomplexes *Xd*ApeK$_4$:E$_4{}^{holo}$ (left panel) and *Xd*ApeK$_4$:F$_4{}^{holo}$ (right panel). The 2F$_O$-F$_C$ electron density map is shown with the displayed ligand omitted for phasing (blue mesh, contoured to 1.0σ). H-bonds are highlighted as black dots. **(b)** The thioester and attached caproic acid (CPA) are displayed as sticks in the *Xd*ApeK$_4$:E$_4{}^{holo}$ complex. The 2F$_O$-F$_C$ electron density map is shown with the displayed ligand omitted for phasing (blue mesh, contoured to 1.0σ). H-bonds and helix dipole are highlighted as black dots and arrow, respectively. **(c)** Sliced surface of *Xd*ApeK$_4$:E$_4{}^{holo}$. The CPA cargo is bound to the Ppant and occupies the polyene channel.

Most strikingly, the F$_O$-F$_C$ map depicts additional electron density near the cysteamine tip in the *Xd*ApeK$_4$:E$_4{}^{holo}$ complex (**Fig. 2.12b**). Most likely, this modification is a fatty acid, e.g., caproic acid (CPA), that is transferred onto ApeE^{holo} during heterologous expression of *Xd*ApeE:K in *E. coli* as has been described for the ant BS before.[27] While the acyl chain of CPA forms hydrophobic interactions with met23^A, trp27^A, met53^B, and phe17^C, the oxygen atom of the thioester linkage is coordinated by a helix dipole moment of H1^A and partially fixated in an oxyanion hole (**Fig. 2.12b**, **c**). His28^A attributes one strong H-bonding interaction (2.8 Å), but the distances to the indole nitrogen of trp27^A (3.9 Å) and the amide atom of the his28^A backbone (3.4 Å) are too large for correct orientation of the artificial CPA cargo in the putative oxyanion hole. Thus, full occupation of the OX is most likely achieved when ApeE or ApeF delivers the correct substrate. Conspicuously, next to the active site asp21^A and the OX, a polar network is formed by his28^A, trp27^A, glu36^B, and arg39^B (**Fig. 2.12b**).

To assign the polyene binding tunnel, a search was conducted for homologous structures that contain a bound ligand. One promising hit is DynE7, an enzyme from the dynemicin biosynthesis with a bound polyene compound (PE, PDB ID: 2XEM[101], rmsd = 1.1 Å over 92 Cα-atoms, SI = 16 %). Structural superposition and analysis of the complexes with DynE7 show a similar-sized cavity built up exclusively by ApeK^{B} ($\approx 5 \times 5 \times 17$ Å^3, **Fig. 2.13a**). Analysis of the surrounding amino acid residues highlights the hydrophobic cavity formed by the side chains of leu42^{B}, met43^{B}, met53^{B}, ile85^{B}, leu92^{B}, and ile94^{B} (**Fig. 2.13b**).

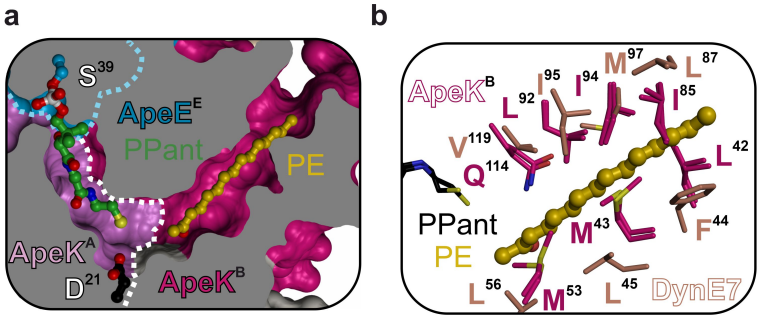

Figure 2.13 Hydrophobic polyene binding cavity formed by *Xd*ApeK. (a) Sliced surface representation of *Xd*ApeK$_4$:F$_4{}^{holo}$ shows an elongated tunnel next to the active site asp21^{A}. The polyene binding channel is formed by ApeK^{B} and fits the framework of polyenes (Ppant and a modeled polyene moiety (gold) are illustrated as balls-and sticks). Protein boundaries of ApeF to both ApeK monomers (light pink, pink) are emphasized by dotted lines in cyan and white, respectively. **(b)** *Xd*ApeK$_4$:F$_4{}^{holo}$ and *Xd*ApeK$_4$:E$_4{}^{holo}$ (pink) imposed onto DynE7 complexed with polyene (DynE7:PE, orange, 2XEM[101]) displays similar features of the hydrophobic polyene cavity.

2.3 Conclusion

In this chapter, production and purification of high yield and purity were achieved for all enzymes involved in the BS of the APE polyene core by complementing existent and establishing new protocols.[35] In addition, the crystallization and determination of the X-ray crystal structures was accomplished for all nine enzymes in seven of their oligomeric states, including the homooligomeric complexes: ApeR$_2$ (KS), ApeQ$_4$ (KR), ApeK$_4$ (TE-like) and the heterooliogomeric complexes: ApeI:P (DHs), ApeO:C (KSs), ApeK$_4$:E$_4{}^{holo}$ (ACP1^{holo} in complex with TE-like), and ApeK$_4$:F$_4{}^{holo}$ (ACP2^{holo} complexed with TE-like). The X-ray structure analysis provided first insights into the nine enzymes that are responsible for the production of the APE polyene core on a structural level. For example, these structures were used to characterize the complex formation of seven quaternary structures. In addition, the solved α/β-heterodimeric ketosynthase is the first and second KS structure in APE and type II PKS systems, respectively, and enabled confirmation of the proposed KS gating mechanism in the APE BS. The X-ray crystal structure ApeI:P resembles the first structure of a heterodimeric DH, with its active site formed by two distinct protomers. The TE-like ApeK tetramer in complex with ApeE and ApeF revealed interesting ACP binding sites and Ppant and polyene cavities. The KR ApeQ from the human pathogen *E. coli* O157:H7 (strain Sakai) visualizes the bound cofactor NADP$^+$ and bridges to the potential relevance in medical applications. To sum up, this work provides first insights on a structural level into the APE type II PKS systems, and the here presented data forms the foundation for subsequent and ongoing studies for targeted mutagenesis and the design of inhibitors.

3 Lyases *Af*yg1p, *Wd*yg1p and *Pl*AntI catalyze polyketide shortening

3.1 Objectives

The lyases *Af*yg1p from *Aspergillus fumigatus* and *Wd*yg1p from *Wangiella* (*Exophiala*) *dermatitidis* are enzymes involved in the biosynthesis of 1,8-dihydroxynaphthalene (DHN)-melanin, a virulence factor in pathogenic fungi. *Af*yg1p and *Wd*yg1p cleave an hepta- and hexaketide, respectively, and the resulting pentaketide 1,3,6,8-tetrahydroxynaphthalene (THN) is the key precursor in the BS of DHN-melanin. Interestingly, an ARO type II PKS system, responsible for the BS of AQ-256 in *Photorhabdus laumondii*, contains *Pl*AntI, a lyase that is homologous to *Af*yg1p and *Wd*yg1p.[26] It catalyzes the last step in the AQ-BS and cleaves an octaketide to the final heptaketide. Conspicuously, AntI is additionally responsible for ring cyclization and formation of the third ring of AQ-256. While the 3D structure of *Pl*AntI is known (PDB ID: 6HXA[26]), no structures are deposited for either *Af*yg1p or *Wd*yg1p. Therefore, the structures of these two enzymes should be determined by X-ray crystallography by establishing the generation of expression plasmids and the recombinant production and purification of the target proteins. Since no protein:ligand complexes of any of these three lyases are available, which are involved in chain-length shortening of polyketides, the overall goal is to illuminate mechanistic insights into the underlying modes of binding and catalysis. Therefore, molecular insights into *Af*yg1p, *Wd*yg1p, and *Pl*AntI bound with mono-, di-, and tetrahydroxynaphthalenes, or covalent inhibitors, were envisioned. In addition, the inhibition of AntI by cinnamic acid (CA), a recently identified endogenous inhibitor of *Pl*AntI with biological relevance in *Photorhabdus laumondii*, is a major goal of this Ph.D. thesis.

3.2 Result and Discussion

This section presents the X-ray crystal structures of the lyases *Af*yg1p, *Wd*yg1p, and *Pl*AntI involved in the biosynthesis of 1,8-DHN melanin in pathogenic fungi and the biosynthesis of anthraquinones, respectively. Moreover, all three enzymes were crystallized in the presence of a diverse range of ligands. All of the obtained X-ray crystal structures will be published in the manuscripts listed below:

"Lyases drive polyketide trimming and cyclization in the biosynthesis of dihydroxynaphthalene-melanins and anthraquinones"

M. Schmalhofer, A. L. Vagstad, H. B. Bode, and M. Groll
Manuscript in preparation, 2023

"Anthraquinone production is influenced by cinnamic acid"

G. L. C. Grammbitter*, **M. Schmalhofer***, H. Vural, N. J. Tobias, M. Groll, and H. B. Bode.
*Equal contributing first author. *Manuscript in preparation*, 2023

Planning, design, and execution of all experiments, as well as analysis and interpretation of all generated results, were conducted with help from Prof. Dr. Michael Groll.

3.2.1 X-ray crystal structures of *Wd*yg1p and *Af*yg1p

The lyases *Af*yg1p and *Wd*yg1p catalyze polyketide shortening in the BS of DHN-melanin. The established purification protocols yielded both proteins in high quantity and purity (**Fig. 3.1a**). They were successfully crystallized and used to collect native datasets of *Af*yg1p and *Wd*yg1p to a resolution of 1.7 Å and 1.85 Å, respectively (**Fig. 3.1b** and **Tab. A.5**). All attempts to derive phase information by applying molecular replacement failed. This may be because of the low similarity of *Af*yg1p and *Wd*yg1p with the homologous structures used as a model (PDB ID: 2JBW[102], 2,6-dihydroxy-pseudo-oxynicotine hydrolase, SI = 24 %; PDB ID: 6HXA[26], *Pl*AntI, SI = 23 %). Therefore, single-wavelength anomalous diffraction (SAD) was performed to obtain experimental phases for *Af*yg1p (see **Subsection 4.2.4**). *Af*yg1p$^{\mathrm{SeMet}}$ was successfully used as a model for MR of the native X-ray crystal structures of *Af*yg1p and *Wd*yg1p (**Fig. 3.2a** and **Tab. A.5**).

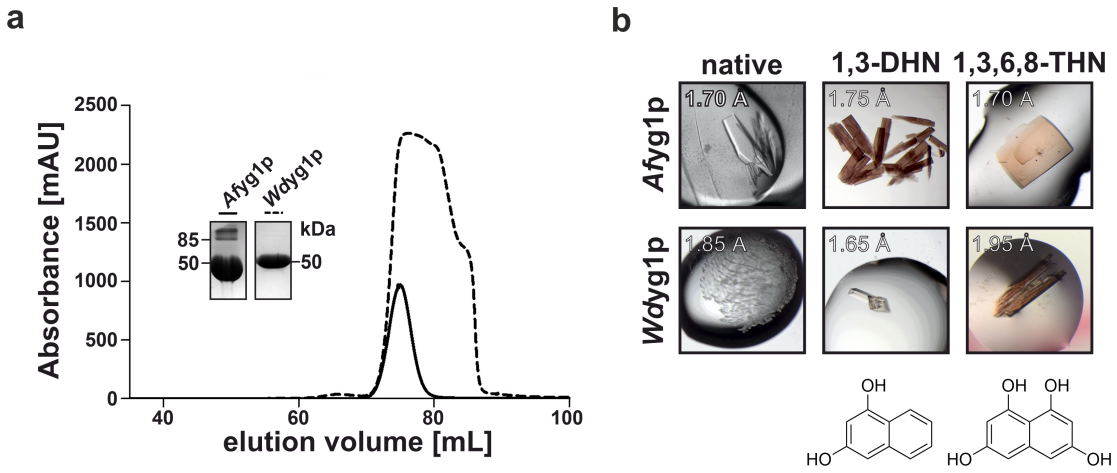

Figure 3.1 Crystallization of the lyases involved in polyketide shortening. (a) SEC chromatogram and SDS-PAGE analysis of *Af*yg1p (black line) and *Wd*yg1p (dashed line). **(b)** Light microscope images of *apo* state crystals (left panel) and co-crystallized with the ligand 1,3-dihydroxynaphthalene (1,3-DHN, middle panel) or the product 1,3,6,8-tetrahydroxynaphthalene (1,3,6,8-THN, right panel).

The data showed that both enzymes consist of an N-terminal domain (N-domain) and a C-terminal domain (C-domain) that are connected with a short linker. The N-domains exhibit an α-helix bundle, and the C-domains adopt an α/β-hydrolase fold.[103] The amino acid sequences of *Af*yg1p and *Wd*yg1p share a sequence identity of 40 %, resulting in a high similarity of the overall fold as well as the active sites (rmsd = 1.1 Å over 382 Cα-atoms, **Fig. 3.2b** and **Fig. A.2**). Despite the low sequence identity of both enzymes towards *Pl*AntI (SI < 20 %), the similarities are striking, as seen in the superpositions (rmsd = 1.6 Å over 244 Cα-atoms and rmsd = 1.8 Å over 247 Cα-atoms, respectively, **Fig. 3.2b**). In all three enzymes, the N- and C-domains are oriented in a back-to-back fashion forming an interface that encloses the intramolecular binding cavity next to the active site (**Fig. 3.2a**, **b**). In these enzymes, the active sites are formed exclusively by the α/β-hydrolase domain and consist of the nucleophilic serine, located in the GXSXG motif, and the amino acid residues histidine and aspartate. With distances in the range of 2.5 Å to 2.8 Å between the active

site residues, the resulting catalytic triads are perfectly aligned and similar to active sites of serine proteases.[104,105]

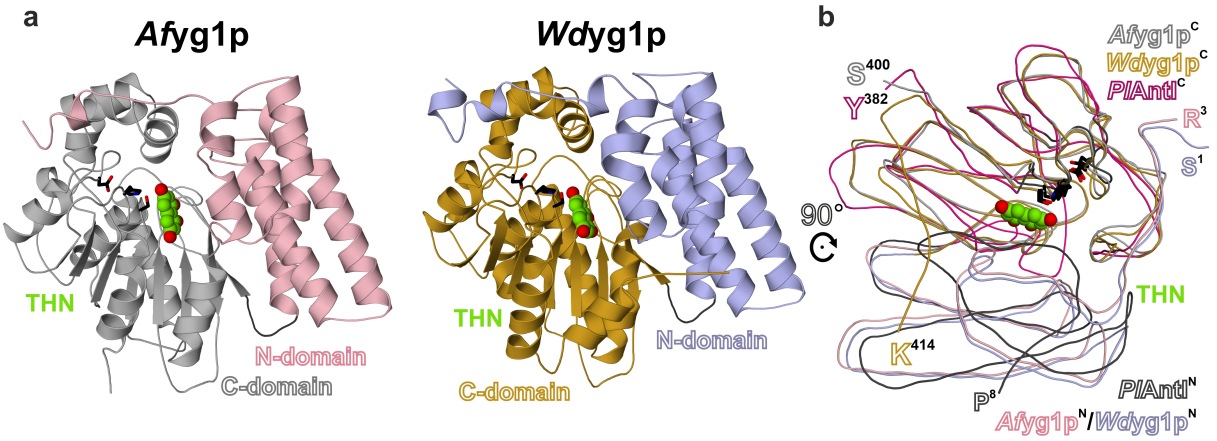

Figure 3.2 Crystal structures of lyases involved in the biosynthesis of 1,8-DHN melanin. (a) Ribbon drawings of *Af*yg1p from *Aspergillus fumigatus* (left panel, 1.7 Å resolution) and *Wd*yg1p from *Exophiala dermatitidis* (right panel, 1.85 Å resolution). The α-helix bundle (N-domain, residues 3-140 (*Af*yg1p) and residues 1-139 (*Wd*yg1p)) is colored in pink and blue, while the C-terminal α/β-hydrolase (C-domain, residues 145-394 (*Af*yg1p) and residues 144-393 (*Wd*yg1p)) is colored in gray and yellow, respectively. The product of *Af*yg1p and *Wd*yg1p THN (green spheres) is depicted in the binding cavity. **(b)** Ribbon plot of *Af*yg1p (residues 3-400) and *Wd*yg1p (residues 1-393) superimposed onto *Pl*AntI from *Photorhabdus laumondii* (residues 7-381, depicted in black and red, PDB ID: 6HXA[26]).

3.2.2 Activity and oxyanion hole

To investigate the active sites of *Af*yg1p, *Wd*yg1p, and *Pl*AntI, sulfonyl fluorides (SF) and chloromethylketones (CMKs) were used. Both warheads are highly reactive toward serine proteases and other enzymes that contain ser-his-asp triads. They react covalently with the active site serine.[106–108] As a result, the structure of *Af*yg1p was obtained complexed with phenylmethane-SF (PMSF, *Af*yg1p$^{\mathrm{PMS}}$, 1.8 Å resolution, **Fig. 3.3a** and **Tab. A.5**). The modification of the active site is clearly seen in the $2F_O$-F_C electron density map in all four monomers of the AU. The nucleophilic ser257O$^{\gamma}$ of *Af*yg1p reacts with SF and is converted to a sulfonate ester that is fully defined in the $2F_O$-F_C electron density map. Comparison with the unbound state shows no structural rearrangements upon ligand binding (rmsd = 0.2 Å over 397 Cα-atoms to *Af*yg1p$^{\mathrm{apo}}$) and, despite the covalent attachment, the catalytic triad remains perfectly aligned with distances of 2.9 Å (ser257O$^{\gamma}$ and his380N$^{\epsilon}$) and 2.5 Å (his380NH$^{\delta}$ and asp352O$^{\delta}$). Most strikingly, binding of the ligand leads to the formation of a tetrahedral orientation of the sulfonate ester and, therefore, mimics parts of a late state reaction intermediate (**Fig. 3.3b**). Thus, the oxyanion hole of *Af*yg1p coordinates the sulfonate ester via the backbone amides of leu187 (3.0 Å) and ala258 (3.0 Å, **Fig. 3.3a**). A helix dipole provided by H8 further intensifies the stabilization of the oxyanion. While the oxygen atoms of the sulfonate ester are highly stabilized by the OX and therefore rigid, the pi-pi-stacking interactions between the benzyl group of PMSF and tyr37, met381, and tyr383 of *Af*yg1p are less pronounced. This results in a mobile aryl moiety with an 80° rotation in-plane, as seen in the four monomer states (**Fig. 3.3a**). Therefore, the lyases were crys-

tallized with phenyl-SF (PSF), lacking the methylene group (-CH_2-) between the sulfur atom and the aromatic ring system.

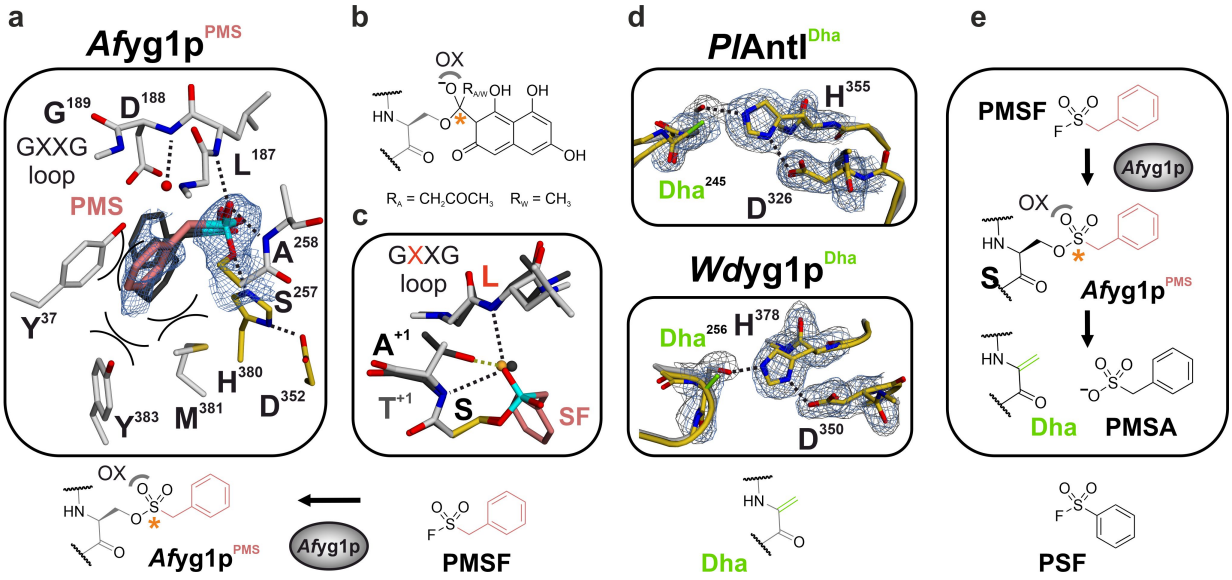

Figure 3.3 Active sites of *Af*yg1p, *Wd*yg1p, and *Pl*AntI. (a) Atomic view of modified active site serine in *Af*yg1p$^{\mathrm{PMS}}$ (Chain B, 1.8 Å resolution). The benzyl group and phosphate atom are depicted as red and cyan sticks, respectively. Different orientations of the benzyl moiety in four monomers of the AU are shown as gray sticks. The $2F_O$-F_C electron density map (depicted as blue mesh) is contoured to 1σ (residues were excluded for phase calculation, H-bonds are shown as black dots, and pi-pi-stacking interactions as black lines). **(b)** Putative intermediate bound to *Af*yg1p. The sp^3-hybridized atom is highlighted (orange asterisks) and the O^- is located in the oxyanion hole (OX, gray curve). **(c)** Active site of *Af*yg1p$^{\mathrm{PMS}}$ (white sticks) superimposed onto *Wd*yg1p$^{\mathrm{apo}}$ (gray sticks) highlights differences in the oxyanion holes. H-bonding interactions of thr257 in *Wd*yg1p are shown as green dots. In the wild-type enzymes, the oxyanion hole is occupied by water molecules (*Wd*yg1p, gray sphere and *Af*yg1p as orange sphere). **(d)** Active site of *Pl*AntI (upper panel, *Pl*AntI$^{\mathrm{Dha}}$, 1.7 Å resolution) and *Wd*yg1p (lower panel, *Wd*yg1p$^{\mathrm{Dha}}$, 1.85 Å resolution) including a dehydroalanine (Dha)-modification superimposed onto the *apo* state catalytic triad. The $2F_O$-F_C electron density map is contoured to 1σ with residues excluded for phasing (Dha-containing as blue mesh and *apo* as gray mesh). H-bonds are shown as black dots. **(e)** Schematic illustration of Dha formation. Mode of action includes esterification of the active serine nucleophile followed by β-elimination of the leaving group phenylmethanesulfonic acid (PMSA).

Remarkably, when using PSF or after a prolonged incubation period with PMSF, strong negative density emerged in the F_O-F_C difference map located at serOH$^{\gamma}$ (*Wd*yg1p$^{\mathrm{Dha}}$, 1.85 Å and *Pl*AntI$^{\mathrm{Dha}}$, 1.7 Å resolution, **Fig. 3.3d**, **Tab. A.6**, and **Tab. A.7**). Thus, the ligand induces a β-elimination at the active site serine resulting in a dehydroalanine (Dha) moiety (**Fig. 3.3e**).[109,110] To exclude spontaneous base-catalyzed elimination in an aqueous solution (crystallization of *Wd*yg1p occurs at pH 8.0-9.0), the structures of crystals older than 30 days were determined. The serine residues in these structures are still present (Ayg1p$^{\mathrm{old}}$, 1.8 Å and *Wd*yg1p$^{\mathrm{old}}$, 1.75 Å resolution, **Tab. A.6**). The catalytic triads adopt the active orientation identical to the *apo* state structures. Notably, the backbone amides leu186NH (2.8 Å) and thr257NH (3.4 Å), as well as the side chain of thr257OH$^{\gamma}$ (2.8 Å), form the oxyanion hole in *Wd*yg1p$^{\mathrm{apo}}$, which is coordinated by a water molecule (**Fig. 3.3c**).

Co-crystallization of AntI with phenacyl chloride (PAC) yielded the X-ray crystal structure *Pl*AntI:PAC at 1.65 Å resolution (**Fig. 3.4a** and **Tab. A.8**). The reactive warhead CMK did not react with the ser245OH$^{\gamma}$ nucleophile. Instead, the reactive ligand is locked in the binding

cavity of *Pl*AntI. Here, the keto group forms H-bonding interactions with asp175NH (3.3 Å), the OX (leu174NH, 3.1 Å), and ser245OH$^{\gamma}$ (3.6 Å) and the chloride atom interacts with ser245OH$^{\gamma}$ (3.4 Å) and asp175O$^{\delta}$ (3.4 Å). The conjugated ring-system undergoes cation-pi-, pi-pi-stacking, and vdW interactions with arg24, phe27, and leu358, respectively (**Fig. 3.4a**). The successful conversion of *Af*yg1p, *Wd*yg1p, and *Pl*AntI with SFs that resulted in the structures *Af*yg1p$^{\mathrm{PMS}}$, *Wd*yg1p$^{\mathrm{Dha}}$, and *Pl*AntI$^{\mathrm{Dha}}$ led to the conclusion that all three enzymes are in an active state with perfectly aligned catalytical triads and that the active sites are accessible in the crystalline form. The X-ray crystal structure of AntI with the trapped PAC highlights the need for further investigation of the binding cavity.

3.2.3 Naphthalene binding pocket

To investigate the specificity pocket and the underlying binding modes of *Af*yg1p, *Wd*yg1p, and *Pl*AntI, conjugated naphthalenes substituted with one (1- and 2-N) or two hydroxy-groups (1,3-, 1,6-, 1,8-, and 2,7-DHN) were used. Thus, the X-ray crystal structure of *Pl*AntI in complex with 1-N was determined (*Pl*AntI:1-N, 2.05 Å resolution, **Fig. 3.4b** and **Tab. A.8**) and of *Af*yg1p and *Wd*yg1p in complex with 1,3-DHN (*Af*yg1p:1,3-DHN, 1.75 Å resolution and *Wd*yg1p:1,3-DHN, 1.65 Å resolution; **Fig. 3.4c**, **d**, and **Tab. A.9**). Upon binding of these

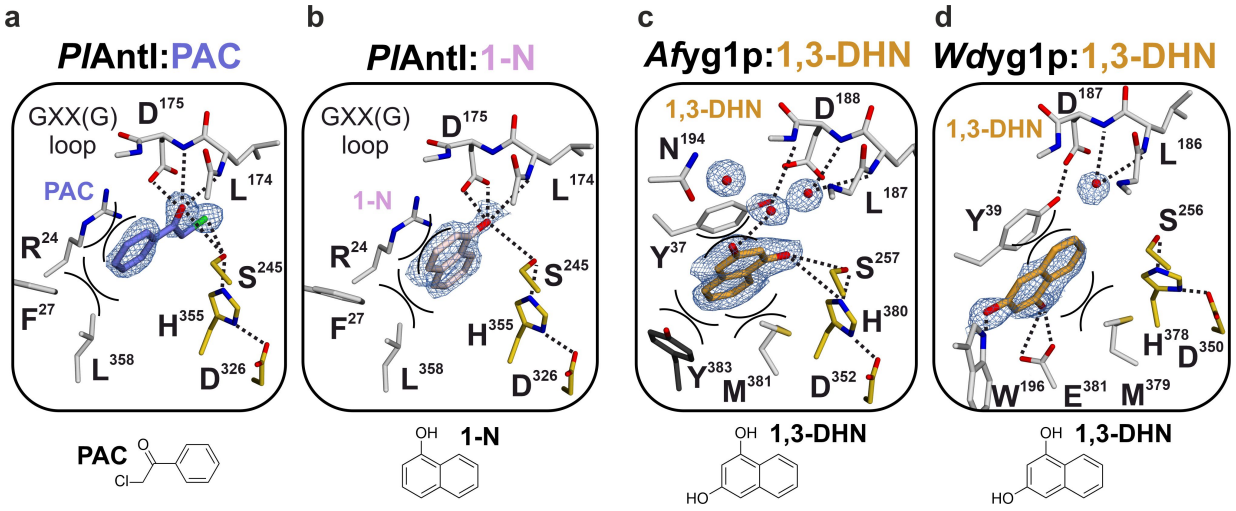

Figure 3.4 *Pl*AntI, *Af*yg1p, and *Wd*yg1p in complex with hydrophobic ligands. Atomic view of *Pl*AntI's active site complexed with **(a)** phenacyl chloride (*Pl*AntI:PAC, 1.65 Å resolution, blue sticks) and **(b)** 1-naphthol (*Pl*AntI:1-N, 2.05 Å resolution, light pink sticks). The $2F_O$-F_C electron density map is shown as blue mesh and contoured to 1σ with the ligand excluded for phasing. H-bonds are indicated as black dots, and cation/pi-pi-stacking interactions as black lines. 1,3-DHN (orange sticks) bound to the specificity pocket of **(c)** *Af*yg1p (*Af*yg1p:1,3-DHN, 1.75 Å resolution, rotated tyr383 is colored in dark gray) and **(d)** *Wd*yg1p (*Wd*yg1p:1,3-DHN, 1.65 Å resolution), respectively.

ligands, there are no major structural rearrangements in *Pl*AntI, *Af*yg1p, or *Wd*yg1p, and the catalytic triads remain perfectly aligned (*Pl*AntI:1-N: rmsd = 0.4 Å over 374 Cα-atoms and others: rmsds $\leq$ 0.2 Å over 375-412 Cα-atoms, respectively). The ligands are buried in the cavity at the interface between N- and C-domain in all three structures (**Fig. 3.4** and **Fig. 3.2a**). The planar naphthalene ring systems are pi-stacked with arg24 (*Pl*AntI, same in-plane orientation when compared to *Pl*AntI:PAC) or between met381/379 and tyr37/39

(*Af*yg1p/*Wd*yg1p). In *Pl*AntI:1-N, the hydroxy group of the ligand interacts with the active site (ser245OH$^{\gamma}$, 3.4 Å) and partly occupies the GLDS motif. This motif adopts a β-turn and consists of the residues gly173-ser176, while 1-OH forms interactions with asp175OH$^{\delta}$ (3.0 Å) and leu174NH (OX, 3.5 Å, **Fig. 3.4b**). In contrast, the β-turns in *Af*yg1p (GLDG motif, residues gly186-gly189) and *Wd*yg1p (GLDG motif, residues gly185-gly188) are occupied by solvent water molecules. The hydroxy groups of 1,3-DHN are oriented towards the active site of *Af*yg1p but form only loose H-bonding interactions with ser257O$^{\gamma}$ and his380N$^{\epsilon}$ (3.7 Å and 3.9 Å, respectively, **Fig. 3.4c**). These minor interactions with the active site are supported by co-crystallization experiments of 1,3-DHN and *Af*yg1p^{S257A}. Here, 1,3-DHN binds in an identical orientation. In *Wd*yg1p:1,3-DHN, the ligand is located at the edge of the binding cavity, and 1-OH is H-bonded to glu381O$^{\epsilon}$ (2.3 Å, 3.5 Å), whereas 3-OH interacts with the indol nitrogen of trp196 (3.1 Å, **Fig. 3.4d**). The side chains of these two amino acid residues define the border of the specificity pocket. The cavity border in *Af*yg1p:1,3-DHN is enclosed by tyr383 that, in comparison to the *apo* state *Af*yg1p structure, rotates 90°. As a result, it undergoes additional pi-pi-stacking interactions with 1,3-DHN (3.7 Å, **Fig. 3.4c**).

The successful incorporation of different hydrophobic compounds in all three enzymes shows that the interactions are based predominantly on pi-pi-stacking and minor H-bonding interactions. Due to the different orientations in *Pl*AntI, *Af*yg1p, and *Wd*yg1p, the aim was to co-crystallize these enzymes in complex with their respective products AQ-256 or THN. While it was not possible to obtain a *Pl*AntI:product structure due to the autooxidation of AQ-256, the structure of *Af*yg1p in complex with THN was successfully determined (*Af*yg1p:THN, 1.7 Å resolution, **Fig. 3.5a** and **Tab. A.10**). The ligand binds identically in all four monomers of the AU and is fully defined in its $2F_O$-F_C electron density map. The naphthalene ring system performs weak pi-pi-stacking with tyr37 and met381, and the binding mode shifts towards an extensive network of H-bonding interactions between *Af*yg1p and the four hydroxy groups of THN (**Fig. 3.5a**). The coordination formed by the binding GLDG motif, the active site, and the amino acids at the specificity pocket's edge are most prominent. Hereby, 1-OH interacts with gly186NH (3.7 Å), leu187NH (3.2 Å), asp188NH (3.3 Å), gly189NH (3.0 Å), and the carboxylic side chain of asp188O$^{\delta}$ (3.7 Å), 8-OH interacts with gly186NH (3.7 Å), gly189NH (3.7 Å), and with the side chain of asn194 via a water molecule. Furthermore, 3-OH is H-bonded with side chains of ser257OH$^{\gamma}$ (3.5 Å), his380N$^{\epsilon}$ (3.6 Å), asp188O$^{\delta}$ (3.6 Å), and tyr37OH$^{\eta}$ (3.4 Å). The 6-OH forms tight H-bonds with the side chain of arg197N$^{\eta}$ (2.9 Å) and tyr383OH$^{\eta}$ (2.3 Å), which was formerly involved in pi-pi-stacking interactions with the naphthalene ring of 1,3-DHN.

To sum up the binding modes of *Af*yg1p, the three complex structures with PMS, 1,3-DHN, and THN reveal that binding is based upon pi-stacking and H-bonding interactions between *Af*yg1p and its ligands (**Fig. 3.5b**). As seen in *Af*yg1p:THN, only the correct hydroxy groups facilitate proper orientation of the naphthalene ring and are a prerequisite for tight coordination by the binding GLDG motif (β-turn, residues gly186-gly189). In *Af*yg1p:1,3-DHN, the ligand is rotated 90° in-plane and translated $\approx$ 3.6 Å towards the solvent. This visualizes the open and spacious binding cavity between the active site and the surrounding solvent

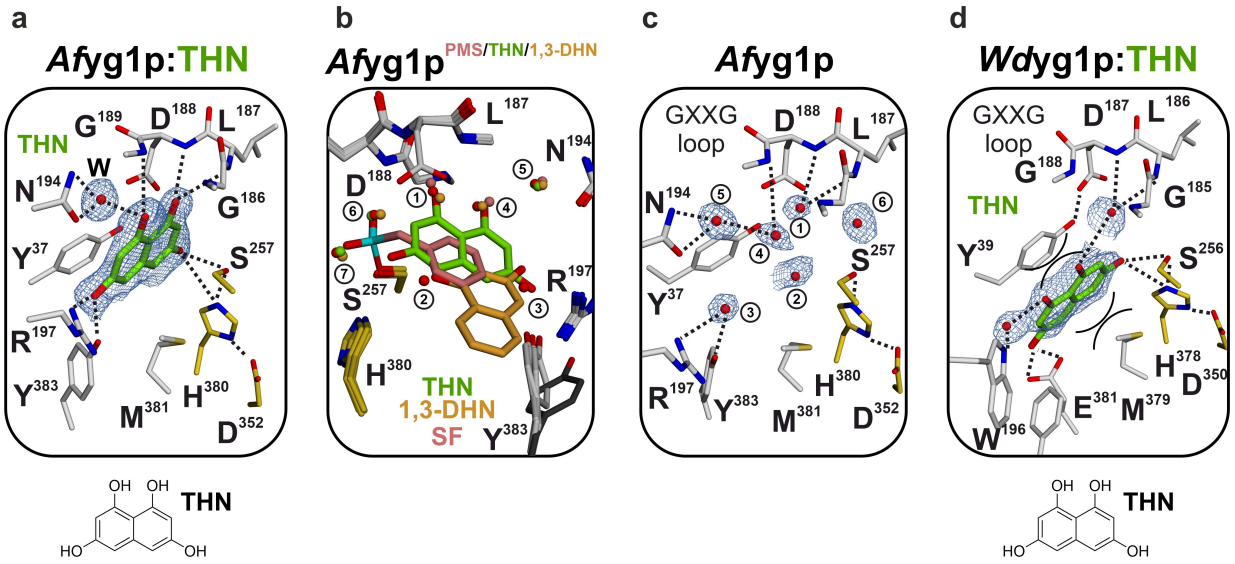

Figure 3.5 Product THN bound to the specificity pocket of *Af*yg1p and *Wd*yg1p. (a) Atomic illustration of THN (green sticks) bound to the binding cavity of *Af*yg1p (*Af*yg1p:THN, 1.7 Å resolution). The $2F_O$-F_C electron density map is shown as blue mesh and contoured to 1σ with the ligand excluded for phasing. H-bonds are indicated as black dots, and cation/pi-pi-stacking interactions as black lines. The electron density map shows additional signals at C4 and C5 of THN. Since THN is prone to autooxidation and forms flaviolin, the occupancy of THN was set to 50 %, and the flaviolins were added (25 % each). For reasons of simplicity, only THN is shown in this figure. THN does not induce any structural rearrangements in *Af*yg1p (rmsd = 0.2 Å over 398 Cα-atoms). **(b)** Superposition of *Af*yg1p in complex with the three ligands PMSF, 1,3-DHN, and THN depicted as red, orange, and green sticks, respectively. Water molecules are shown as spheres and colored according to the respective ligands. Tyr382 rotates by 90° in the presence of dehydroxylated surrogates in the *Af*yg1p:1,3-DHN complex structure and is highlighted in gray. **(c)** Stick representation of the active site in the X-ray structure of *Af*yg1p^{apo}. **(d)** Atomic view of THN (green sticks) bound to the binding cavity of *Wd*yg1p (*Wd*yg1p:THN, 1.95 Å resolution). THN does not induce any structural rearrangements in *Wd*yg1p (rmsd = 0.2 Å over 410 Cα-atoms).

that spans $\approx 7 \times 10 \times 12$ Å^3 (see below). Interestingly, in *Af*yg1p^{apo}, this specificity pocket is filled with solvent water molecules, and each position matches with the hydroxy groups 1-OH, 6-OH, and 8-OH of THN and 1-OH of 1,3-DHN ($\approx$ 3-OH in THN), respectively. Thus, substrate binding is also entropically favored by releasing these water molecules (**Fig. 3.5c**).

Next, the crystal structure of *Wd*yg1p was determined in complex with its product THN (*Wd*yg1p:THN, 1.95 Å resolution, **Fig. 3.5d** and **Tab. A.10**). In contrast to the highly coordinated THN in *Af*yg1p, the binding of THN in *Wd*yg1p is less pronounced and indicates product release. Nevertheless, the electron density map displays the pi-pi-stacking of THN with try39 and met379. Additionally, the 3-OH is H-bonded with ser256OH$^{\gamma}$ (3.2 Å), his378N$^{\epsilon}$ (3.2 Å), and tyr39OH$^{\eta}$ (3.6 Å), while 1-OH is coordinated via a water molecule to gly185NH. As seen for 1,3-DHN, THN is located at the edge of the binding cavity, and its 6-OH interacts with glu381O$^{\epsilon}$ (2.3 Å and 2.5 Å), whereas 8-OH coordinates with the indol nitrogen atom of trp196 (via a water molecule, **Fig. 3.5d**). In *Wd*yg1p, glu381 and trp196 define the border of the specificity pocket, which is more spacious than the binding pocket of *Af*yg1p (border-defining residues: arg197 and tyr383) and spans $\approx 7 \times 12 \times 12$ Å^3. Taken together, the data of the lyases in complex with mono-, di-, and tetrahydroxynaphthalenes show a fine-tuned balance of ligand binding. The combination of pi-pi-stacking with the naphthalene ring

and the H-bonding with the hydroxy groups is responsible for orienting the substrates inside the naphthalene binding cavity.

3.2.4 Acetyl and acetoacetyl binding pocket

After identification of the naphthalene specificity pocket in *Af*yg1p and *Wd*yg1p, the acetyl and acetoacetyl binding sites were analyzed (**Fig. 3.6a**, **b**). *Af*yg1p cleaves the heptaketide YWA1^{open}. Therefore, an acetoacetyl is positioned next to the active site. The sliced surface representation shows that, in addition to the naphthalene specificity pocket, *Af*yg1p provides a hydrophobic tunnel next to the active site that spans $\approx 7 \times 8 \times 14$ Å^3 (**Fig. 3.6a**). This

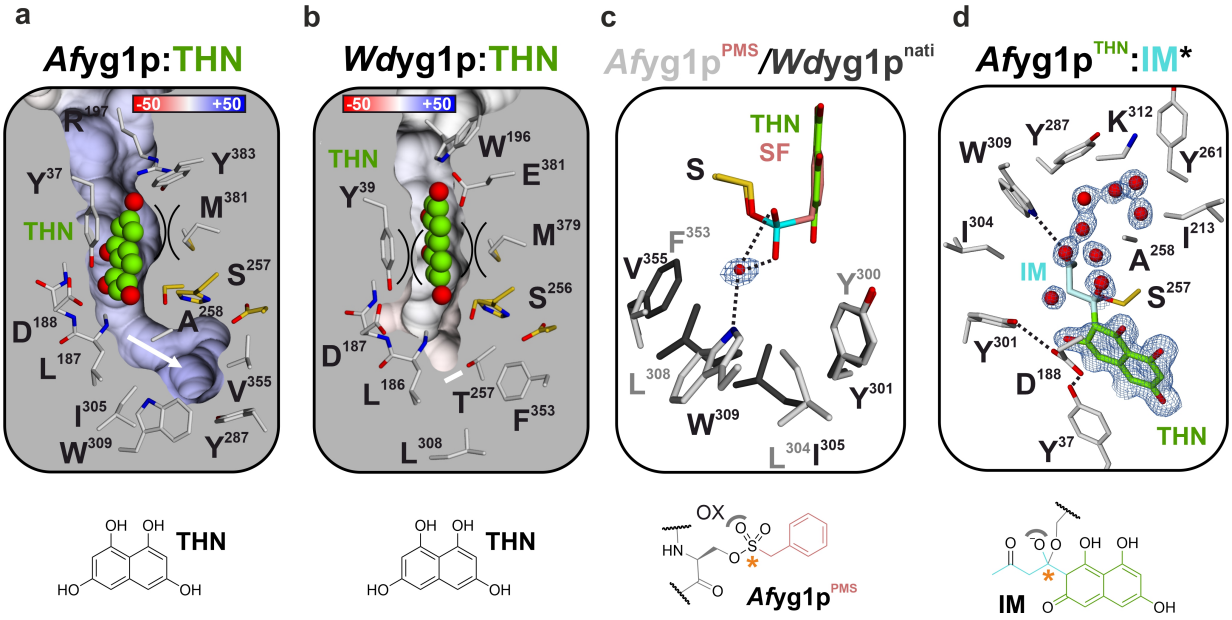

Figure 3.6 The naphthalene and (aceto)acetyl specificity pockets of *Af*yg1p and *Wd*yg1p. Sliced surface representation with electrostatic potentials of THN bound to **(a)** *Af*yg1p and **(b)** *Wd*yg1p. THN is shown as green spheres, pi-stacking interactions as semicircles, and a white arrow/line indicates (aceto)acetyl-tunnel. **(c)** Superposition of *Wd*yg1p$^{\mathrm{apo}}$ (gray sticks) and *Af*yg1p$^{\mathrm{PMS}}$ (white sticks). The $2F_O$-F_C electron density map is shown as blue mesh and contoured to 1σ with the water molecule excluded for phasing. H-bonds are indicated as black dots, and THN is modeled as green sticks. **(d)** Schematic view of (aceto)acetyl tunnel illustrating the *Af*yg1p:THN structure with modeled (*) tetrahedral intermediate state (IM, blue and green sticks).

chamber is confined by the amino acid residues ala258, tyr287, ile305, trp309, and val355. A striking feature in the tunnel of *Af*yg1p is trp309, which is bound to a well-defined water molecule located next to the OX. Notably, this water is also present in *Af*yg1p$^{\mathrm{apo}}$, *Af*yg1p$^{\mathrm{PMS}}$, *Af*yg1p:1,3-DHN, and *Af*yg1p:THN (**Fig. 3.6c**). These results indicate that trp309 stabilizes the second oxo-group in the ketide, which is absent in the substrate of *Wd*yg1p. Consistent with the shorter substrate, the cavity in *Wd*yg1p is more hydrophobic and significantly reduced in size by thr257, leu304, leu308, and phe353 ($\approx 8 \times 7 \times 7$ Å^3, **Fig. 3.6b**, **c**). Thus, the obtained structures show major differences in the acetyl and acetoacetyl cavities of the lyases *Pl*AntI, *Af*yg1p, and *Wd*yg1p.

3.2.5 Proposed enzyme mechanism

The determined structures of *Af*yg1p highlight several key features: 1) the orientation of the conjugated pi-system in the naphthalene specificity pocket and tight coordination of the four hydroxy groups (*Af*yg1p:THN). 2) Mimic of the sp^3-hybridized intermediate state allowed assignment of the OX (*Af*yg1p$^{\mathrm{PMS}}$). 3) Trp309 stabilizes the acetoacetyl moiety (*Af*yg1p$^{\mathrm{apo/PMS}}$ and *Af*yg1p:1,3-DHN/THN). With these experimental results, the tetrahedral intermediate was positioned in its keto form (IM*[58]) bound to *Af*yg1p using the locations of THN, the identified tetrahedral orientation, and the water molecule next to trp309 as anchor points (**Fig. 3.6d**). Based on this modeling, the following mechanism is proposed (**Fig. 3.7**). First,

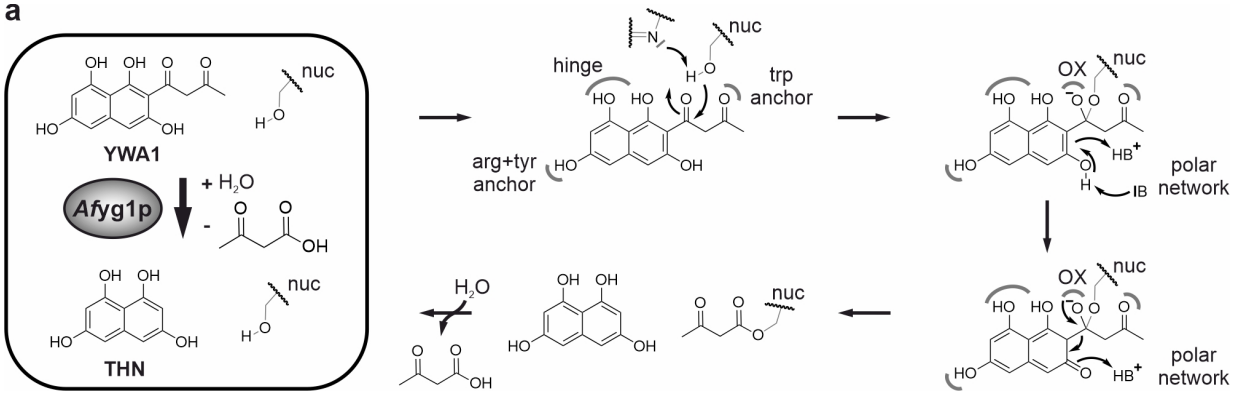

Figure 3.7 Proposed enzyme mechanism of *Af*yg1p. (a) Carbon-carbon bond breakage proceeds via a tetrahedral intermediate, nuc = active site nucleophile (ser), trp anchor = pocket that stabilizes the aceto moiety of YWA1, hinge = β-turn that stabilizes the 1- and 8-hydroxy groups, OX = oxyanion hole, polar network = responsible for keto-enol tautomerism, B = base, and BH^+ = protonated base.

the substrate YWA1 binds to *Af*yg1p by tight coordination of all four hydroxy groups of the naphthalene ring system. Asp352 and his380 of the catalytically active triad deprotonate the nucleophilic ser256OH$^{\gamma}$. The resulting ser256O$^{\gamma-}$ attacks the electrophilic keto group (1'-carbonyl), and the emerging negatively charged oxygen atom is stabilized in the oxyanion hole built up by the backbone amides of leu187 and ala258. In contrast, the polar network (tyr37OH$^{\eta}$, asp188O/OH$^{\delta}$, and tyr301OH$^{\eta}$) enables the conversion of the enol- to the keto-form. This network is in proximity to 3-OH and C2 (3.2 Å and 3.0 Å to asp188O/OH$^{\delta}$, respectively) and facilitates the deprotonation of 3-OH and the successive protonation at the C2 carbon atom. The keto-form of the IM allows the reshuffling of the electrons back into the ring system, which is responsible for the carbon-carbon (C-C) bond breakage. Last, the protonation of O3 takes place via the polar network and the formation of the fully conjugated ring product THN. Hydrolysis of the bound acetoacetyl moiety restores the active state of the enzyme. This proposed mechanism is most likely transferable to *Wd*yg1p due to their high similarity in the active site cavity that includes the β-turn (**Fig. 3.5** and **Fig. A.2**).

3.2.6 Cinnamic acid - an endogenous inhibitor of *Pl*AntI

The group of Prof. Dr. Helge B. Bode recently identified cinnamic acid (CA) as an endogenous inhibitor of *Pl*AntI. In this subsection, the inhibition mode was investigated using crystallography. The AntI:CA complex is key to revealing insights into the molecular mechanism. Therefore, the AntI:CA X-ray structure was determined at a high-resolution of 1.4 Å (**Fig. 3.8a** and **Tab. A.10**). The structural comparison of AntI:CA with the previously obtained

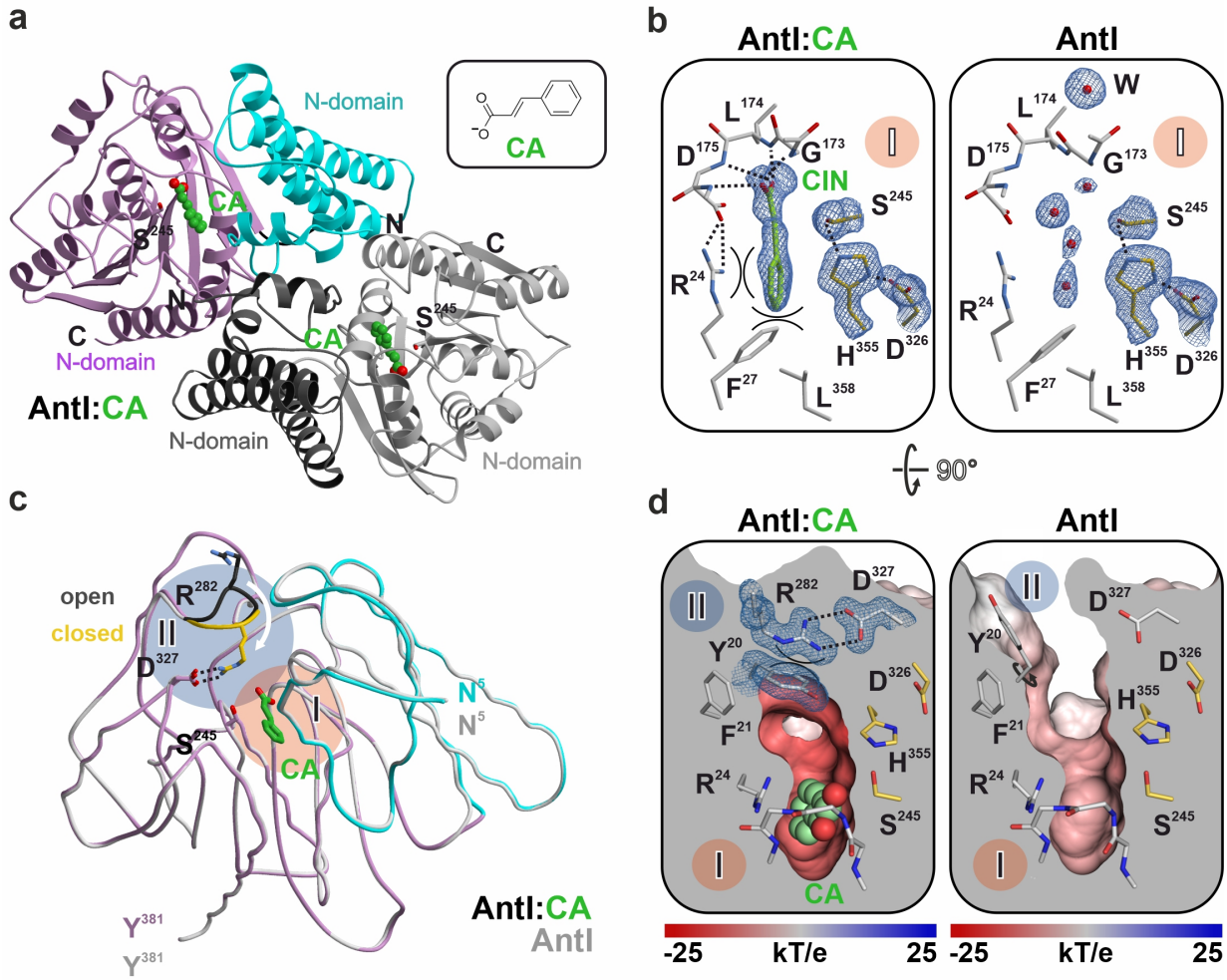

Figure 3.8 The binding mode of CA to AntI. (a) Ribbon representation of dimeric AntI in complex with the inhibitor CA (*Pl*AntI:CA, 1.4 Å resolution, carbon atoms in green). **(b)** Atomic view of the active site in *apo* (right panel) and the CA-bound structure (left panel). Yellow sticks indicate the aligned triad; surrounding residues (gray) form salt bridges, H-bonds (dashed lines), and pi-stacking interactions (semicircles) with CA (green sticks). The F_O-F_C electron density map (blue mesh) is contoured to 3.0σ (CA, water molecules, and the active site residues were excluded for phase calculations). **(c)** The superposition of $AntI^{apo}$ (gray, PDB ID: 6HXA[26]) with AntI:CA depicts the inhibitor bound in a central cavity (region II) and structural rearrangements of the substrate binding channel (region I). The loop connecting residues pro281-arg283 (gold) is rotated by 90°. **(d)** Sliced surface representations of the closed (ligand-bound) and open (*apo*) states with electrostatic potentials. The $2F_O$-F_C electron density map (blue mesh, contoured to 1.0σ) of the asp327 and arg282 forming the gate and tyr20 is illustrated.

AntI structures shows preservation of the homodimer, whereby both subunits maintain their identical conformation (rmsd = 0.3 Å over 374 Cα-atoms).[26] However, the binding of CA in the cavity in proximity to the active site ser245 (6 Å in width and 9 Å in length) results in

rearrangements of the enzyme (see below). Although the catalytic triad is aligned in AntI:CA (region I), there is no additional electron density between ser245 and the ligand (distance: 3.3 Å, **Fig. 3.8b**). Thus, CA acts as a non-covalent inhibitor. Most prominently, the styryl moiety of CA performs strong pi-pi-stacking with arg24 and phe27, locking the ligand in the binding cavity, whereas H-bonds with the backbone amides of the GLDS motif (β-turn, residues gly173, leu174, asp175, and ser176) stabilize its carboxyl tail (**Fig. 3.8b**). The styryl moiety of CA mimics the quadrupole of the two-ring system in the AntI substrate, which requires precise orientations for C-C bond breakage and subsequent cyclization. Moreover, the carboxylate interacts with the β-turn, which plays a key role in stabilizing the tetrahedral intermediates.[26] AntI$^{\mathrm{apo}}$'s active site is filled with water molecules that exhibit high mobility (**Fig. 3.8b**). Upon binding of CA, desolvation leads to a favorable change in entropy, whereas CA's planar and rigid nature favors ligand binding. A comparison of the surrounding amino acids in the binding pocket of AntI:CA with the *apo* state depicts similar orientations of all residues except gly173 (**Fig. 3.8b**). The backbone oxygen of this glycine forms an H-bond with a well-defined water molecule (W, 2.8 Å) in the unbound state. However, upon ligand binding, the carboxyl tail results in a steric clash between gly173CO and CA (2.8 Å). Intriguingly, the carboxylate forces a 37° rotation of the glycine keto group and increases the distance to 3.8 Å. In contrast, the rotation of gly173 misaligns the binding GLDS motif, resulting in an unfavored structural orientation. Additionally, protein-ligand binding includes four H-bonds formed between the amides in the backbone of the motif and the carboxylate oxygen atoms. Three of them are in close distances up to 2.6 Å, whereas the leu174NH:CA interaction (2.2 Å) induces repulsive forces that destabilize the inhibitor in the binding site counteracting the enthalpic and entropic favored coordination. Taken together, the structural analysis discovered a non-perfect inhibitor that acts as an endogenous modulator. Fixation of CA via pi-stacking in the reaction cavity of AntI leads to a high degree of tension in the binding β-turn. This spring-like mechanism creates a fine-tuned regulator of AntI. The superposition of AntI$^{\mathrm{apo}}$ and AntI:CA demonstrates the high similarity of the overall structures (rmsd = 0.3 Å over 374 Cα-atoms, **Fig. 3.8c**), while the binding tunnel undergoes major rearrangements in the presence of CA (region II). In the *apo*-state, the loop (pro281, arg282, and arg283) points outwards, and the gatekeeper residue tyr20 is located at the outer end of the binding tunnel.[26] These orientations provide an accessible and elongated substrate binding cleft connecting the catalytic center to the protein surface (5 Å in width and 15 Å in length, **Fig. 3.8d**). After ligand binding, the loop rotates 90°, and the side chain of arg282 moves 180° inwards. This shift facilitates the formation of a β-turn and a new salt bridge interaction between arg282 and asp327 (**Fig. 3.8c**). Notably, arg282 is also stabilized by pi-stacking with tyr20. As a result, these residues seal AntI:CA and form a reaction cavity with a width of 6 Å and a length of 9 Å (**Fig. 3.8d**). The data illustrates that AntI uncouples the final enzymatic reactions from the rest of anthraquinone biosynthesis. Thereby, AntI controls and acts in a chronological sequence with a step-wise progression. First, the C-C bond breaks and releases AntF, followed by major rearrangements in the binding tunnel that closes the enzyme for cyclization. In the last step, the product is released and oxidized to AQ-256.

3.3 Conclusion

In the second part of this work, the lyases *Af*yg1p, *Wd*yg1p, and *Pl*AntI were analyzed. The first two enzymes are involved in the BS of 1,8-dihydroxynaphthalene-melanin, whereas the lyase AntI is part of the anthraquinone biosynthesis. The crystal structures of *Af*yg1p and *Wd*yg1p were successfully determined and characterized. Sulfonyl fluorides were used to mimic the tetrahedral intermediate state of the reaction. This allowed visualization of the active site triad in action and the stabilization of the oxyanion. The X-ray analyses of these lyases in complex with inhibitors, product, and surrogates revealed a fine-tuned balance of entropic and enthalpic effects on the binding mode. The assigned ketide tunnel of *Af*yg1p forms prominent H-bonding interactions with the acetoacetyl of YWA1^{open}, whereas the ketide tunnel of *Wd*yg1p is blocked. These results allowed the proposition of an enzyme mechanism for *Af*yg1p and *Wd*yg1p. In addition, the structural investigation of AntI in complex with its endogenous inhibitor cinnamic acid discovered an interesting spring-like mechanism as the mode of action. Furthermore, open- and closed-state structures of AntI suggest a gating mechanism that ensures the correct ring formation to the anthraquinone product. To this end, the determined structures of *Af*yg1p, *Wd*yg1p, and *Pl*AntI in the presence of surrogates, products, inhibitors, and mimics provide a foundation for the design of inhibitors. These molecules might be particularly interesting for *Af*yg1p, which is involved in DHN-melanin biosynthesis of the pathogenic fungi *Aspergillus fumigatus*.

4 Materials and Methods

4.1 Materials

4.1.1 Chemicals and enzymes

Unless stated otherwise, all chemicals were purchased from Sigma-Aldrich (St. Louis, US), AppliChem (Darmstadt, DE), Serva (Heidelberg, DE), Carl Roth (Karlsruhe, DE), Merck (Darmstadt, DE), or VWR (Radnor, US).

Table 4.1 Ligands and cofactors

Ligands	
1,3,6,8-tetrahydroxynaphthalene (THN)	ChemScene, Monmouth Junction, US
Phenylmethane sulfonyl fluoride (PMSF)	AppliChem, Darmstadt, DE
1,3-dihydroxynaphthalene (1,3-DHN)	Sigma-Aldrich, St. Louis, US
1,6-dihydroxynaphthalene (1,6-DHN)	Sigma-Aldrich, St. Louis, US
1,8-dihydroxynaphthalene (1,8-DHN)	Sigma-Aldrich, St. Louis, US
2,7-dihydroxynaphthalene (2,7-DHN)	Sigma-Aldrich, St. Louis, US
1-naphthol (1-N)	Sigma-Aldrich, St. Louis, US
2-naphthol (2-N)	Sigma-Aldrich, St. Louis, US
Phenacyl chloride (PAC)	Sigma-Aldrich, St. Louis, US
Cinnamic acid (CA)	Sigma-Aldrich, St. Louis, US
Cofactors	
Red. nicotinamide adenine dinucleotide phosphate (NADPH)	BioCatalytics, Pasadena, US
Coenzyme A (CoA)	Sigma-Aldrich, St. Louis, US

Table 4.2 Enzymes

BamHI-HF, PstI-HF, and NcoI-HF(20 U/µL)	New England Biolabs, Ipswich, US
NdeI and XhoI (20 U/µL)	New England Biolabs, Ipswich, US
DpnI (20 U/µL)	New England Biolabs, Ipswich, US
Q5-HF DNA polymerase (2 U/µL)	New England Biolabs, Ipswich, US
PfuTurbo DNA polymerase (2.5 U/µL)	Agilent, Santa Clara, US
T4 DNA Ligase (400 U/µL)	Agilent, Santa Clara, US
SUMO-protease (Ulp1, *Saccharomyces cerevisiae*)[111]	prepared according to Institute protocols
tobacco etch virus (TEV)-protease (S219P-mutant)[112]	prepared according to Institute protocols

4.1.2 Standards and kits

Table 4.3 Standards and kits

DNA standard	
peqGOLD DNA ladder 1 kb (0.1-10 kb)	Avantor, Radnor, US (peqlab VWR)
Protein standard	
Unstained protein standard, broad range (10-200 kDa)	New England Biolabs, Ipswich, US
Color prestained protein standard, broad range (10-250 kDa)	New England Biolabs, Ipswich, US
Molecular cloning and mutagenesis kits	
Wizard Plus SV Minipreps DNA Purification Systems	Promega, Fitchburg, US
Wizard SV Gel and PCR Clean-Up System	Promega, Fitchburg, US
QuickChange II Site Directed Mutagenesis Kit	Agilent, Santa Clara, US

4.1.3 Laboratory equipment and tools

Table 4.4 General instruments

Thermomixer comfort	Eppendorf, Hamburg, DE
Varioclav	HP Labortechnik, Oberschleißheim, DE
Autoklave Prestige Medical 2100 Classic	Soleni, Schenkendöbern, DE
Magnetic Stirrer Hei-Tec	Heidolph, Schwabach, DE
Analytical balance LA124i	VWR, Radnor, US
Precision balance M-power AZ4101	Sartorius, Göttingen, DE
UV germicidal lamp VL-315G	Vilber, Collégien, FR
Tecan infinite 200Pro	Tecan, Männedorf, CH
96-well plate flat bottom transparent polystyrene	Greiner, Kremsmünster, AT
MyCycler Thermal Cycler	Bio-Rad, Hercules, US
Drying oven ED 53	Binder, Tuttlingen, DE
CFX96 Touch Real-Time PCR Detection System	Bio-Rad, Hercules, US
G:Box UV/White Light Gel Documentation System	Syngene, Bangalore, IN
pH meter InoLab WTW series pH 720	Xylem Analytics , Rye Brook, US
pH meter electrode WTW Sentix 91	Xylem Analytics , Rye Brook, US

Table 4.5 Centrifuges

Refrigerated floor-standing centrifuge Sigma 8K	Sigma, Osterode, DE
Refrigerated centrifuge Sigma 4K15	Sigma, Osterode, DE
Refrigerated high speed centrifuge Sigma 3-30K	Sigma, Osterode, DE
Refrigerated microcentrifuge Sigma 1-14K	Sigma, Osterode, DE

Table 4.6 Electrophoresis and electroporation

Electrophoresis	
Agarose electrophoresis chamber and casting modules	Appligene, Heidelberg, DE
VXR basic Vibrax	Ika, Staufen im Breisgau, DE
Mini-PROTEAN Tetra Vertical Electrophoresis Cell	Bio-Rad, Hercules, US
Mini-PROTEAN Tetra Cell Casting Module	Bio-Rad, Hercules, US
Light panel prolite basic 2	Kaiser Fototechnik, Buchen, DE
DNA Stain G	Serva, Heidelberg, DE
Electrophoresis Power Supply EPS 600	Pharmacia Biotech, Uppsala, SE
Electroporation	
Eporator	Eppendorf, Hamburg, DE
MicroPulser Electroporation Cuvettes (2 mm gap)	Bio-Rad, Hercules, US

Table 4.7 Recombinant protein expression and purification

Protein expression	
Infors HT Multitron	Infos HT, Bottmingen, CH
3 L Fernbach with vent cap	Corning, New York, US
Ultrospec 10 Cell Density Meter	Cytiva, Marlborough, US
Single-use cuvettes polystyrene, semi-micro 1.5 mL	Brand, Wertheim, DE
Protein purification	
Digital sonifier 250	Branson Ultrasonics, Brookfield, US
Äkta pure	Cytiva, Marlborough, US
Äkta prime	Cytiva, Marlborough, US
Äkta purifier	Cytiva, Marlborough, US
Fraction Collector F9-R	Cytiva, Marlborough, US
Amicon Ultra-15 Centrifugal Filter Units	Cytiva, Marlborough, US
Dialysis tubing Membra-Cel cellulose	Carl Roth, Karlsruhe, DE
Nanodrop 2000c	Thermo Fisher Scientific, Waltham, US
Pefabloc SC	Roche, Basel, CH
Chromatography columns	
HisTrap^{HP} 5 mL	Cytiva, Marlborough, US
StrepTrap^{HP} 5 mL	Cytiva, Marlborough, US
Superdex 75 HiLoad 16/600 pg	Cytiva, Marlborough, US
Superdex 200 HiLoad 16/600 pg	Cytiva, Marlborough, US
Superdex 200 Increase 10/300 GL	Cytiva, Marlborough, US
Resource Q 6 mL	Cytiva, Marlborough, US

4.1.4 Crystallization tools and screens

Table 4.8 Crystallization tools

22 mm Siliconized Glass Cover Slides	Hampton Research, Aliso Viejo, US
Crystalgen SuperClear 24 well plates pregreased	Jena Bioscience, Jena, DE
Intelli-Plate 96-3 LVR	Art Robbins Instruments, Sunnyvale, US
MICROLAB STARlet	Hamilton, Reno, US
Magnetic CryoVials	Molecular Dimensions, Newmarket, UK
Magnetic CryoCap	Molecular Dimensions, Newmarket, UK
Mounted CryoLoop, (20 micron)	Hampton Research, Aliso Viejo, US
Universal Puck Kit	Crystal Positioning Systems, Jamestown, US
Quick Combi Sealer Plus	HJ-Bioanalytik, Erkelenz, DE
JANSi UVEX Imaging Systems	JAN Scientific, Seattle, US; SWISSCI, Neuheim, CH
CrystalWand Magnetic	Hamilton, Reno, US
Crystal Phoenix Liquid Handling System	Art Robbins Instruments, Sunnyvale, US
Crystal Gryphon Liquid Handling System	Art Robbins Instruments, Sunnyvale, US
Oryx4 Protein Crystallization Robot	Douglas Instruments, East Garston, UK
Cooled Incubator Series 3000	RUMED Rubarth Apparate, Laatzen, DE
Research Stereomicroscope System SZX10	Olympus, Tokio, JP
SC50 microscope camera	Olympus, Tokio, JP
Box of microtools	Molecular Dimensions, Newmarket, UK
Vial Tongs (45 and 90°)	Molecular Dimensions, Newmarket, UK
Cryogenic Sample Storage HC20, HC35	Worthington Industries, Columbus, US
Cryo Exchange CX100	Worthington Industries, Columbus, US
Transport container CS 100 SK	Cryo Anlagenbau, Wilnsdorf, DE
Foam dewar	Spearlab, San Francisco, US

Table 4.9 NeXtal Tubes Suites for Crystallization screens

pHClear Suite	Qiagen, Hilden, DE
pHClear II Suite	Qiagen, Hilden, DE
Classics Suite	Qiagen, Hilden, DE
Classics II Suite	Qiagen, Hilden, DE
JCSG+ Suite	Qiagen, Hilden, DE
PEGs Suite	Qiagen, Hilden, DE
PEGs II Suite	Qiagen, Hilden, DE
Protein Complex Suite	Qiagen, Hilden, DE
$AmSO_4$ Suite	Qiagen, Hilden, DE

4.1.5 Genes, plasmids, and primers

Table 4.10 Genes

Construct	Gene ID of encoding gene	Origin organism
***Ec*ApeQ**[a]	915807	*Escherichia coli* O157:H7 (strain Sakai)
***Ec*ApeE**[a]	916270	*Escherichia coli* O157:H7 (strain Sakai)
***Ec*ApeF**[a]	916293	*Escherichia coli* O157:H7 (strain Sakai)
***Ec*ApeK**[a]	916364	*Escherichia coli* O157:H7 (strain Sakai)
***Af*yg1p**[a]	AF116902.1 (GenBank)	*Aspergillus fumigatus* (strain B-5233)
***Wd*yg1p**[a]	20306951	*Exophiala dermatitidis* (strain NIH/UT8656)
***Xd*ApeC**	34866642	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeE**	34867408	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeF**	34867338	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeI**	34869804	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeK**	34869256	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeO**	34867359	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeP**	34868720	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeQ**	34866380	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Xd*ApeR**	34868138	*Xenorhabdus doucetiae* FRM16 (DSM 17909)
***Pl*AntI**	24170538	*Photorhabdus laumondii* subsp. *laumondii* TTO1

[a]Genes were used for molecular cloning as part of this work. Codon-optimized gene sequences are listed in **Tab. A.1**.

Table 4.11 Expression vectors

Basic vectors		
pRSET A^{Amp}	Bacterial Expression Vector	Thermo Fisher Scientific, Waltham, US
$pCDF^{Spec}$/$pACYC^{Cm}$/$pCOLA^{Kan}$	Duet Vectors for Co-expression	Novagen, Darmstadt, DE
Generated expression vectors	**(Co-)expression of**	
$pCDF^{Spec}$-TEV-Strep-*Xd*ApeIP	Strep-TEV-*Xd*ApeP and ApeI	Bode group[a]
$pACYC^{Cm}$-TEV-Strep-*Xd*ApeOC	Strep-TEV-*Xd*ApeO and ApeC	Bode group[a]
$pCOLA^{Kan}$-TEV-Strep-*Xd*ApeR	Strep-TEV-*Xd*ApeR	Bode group[a]
$pACYC^{Cm}$-TEV-Strep-*Xd*ApeQ	Strep-TEV-*Xd*ApeQ	Bode group[a]
pRSET A^{Amp}-TEV-His_6-*Ec*ApeQ	His_6-TEV-*Ec*ApeQ	This work
$pCATI4^{Kan}$-TEV-His_6-*Xd*ApeF	His_6-TEV-*Xd*ApeF	Bode group[a]
$pCOLA^{Kan}$-TEV-Strep-*Xd*ApeEK	Strep-TEV-*Xd*ApeE and ApeK	Bode group[a]
$pCOLA^{Kan}$-TEV-Strep-*Xd*ApeFK	Strep-TEV-*Xd*ApeF and ApeK	Bode group[a]
$pCDF^{Spec}$-TEV-Strep-*Ec*ApeEK	Strep-TEV-*Ec*ApeE and ApeK	This work
$pCDF^{Spec}$-TEV-Strep-*Ec*ApeFK	Strep-TEV-*Ec*ApeF and ApeK	This work
pRSET A^{Amp}-TEV-His_6-*Af*Ayg1p	His_6-TEV-*Af*Ayg1p	This work
$pCDF^{Spec}$-TEV-Strep-*Wd*yg1p	Strep-TEV-*Wd*yg1p	This work
$pCOLA^{Kan}$-TEV-His_6-*Pl*AntI	His_6-TEV-*Pl*AntI	Bode group[a]

[a]These expression vectors were kindly provided by Dr. Gina L. C. Grammbitter and Prof. Dr. Helge B. Bode.

Table 4.12 Primer list

Sequencing primer	
T7-terminator-R	GCT AGT TAT TGC TCA GC GG
ACYCDuetUP1-F	GGA TCT CGA CGC TCT CCC T
DuetUP2-F	TTG TAC ACG GCC GCA TAA TC
DuetDOWN1-R	GAT TAT GCG GCC GTG TAC AA
pRSET RP-R	ATG CTA GTT ATT GCT CAG C
Molecular cloning primer	
*Af*yg1p-BamHI-F	**CCA GGA TCC** CCG CGT TGG ATC TTA GGT G
*Af*yg1p-PstI-R	**CCA CTG CAG TTA** TTA GTT TTT GGT TTT AGA CGG AG
*Wd*yg1p-BamHI-F	**CCA GGA TCC** GCT ACC GAA AAA TAC TAC ATC CG
*Wd*yg1p-PstI-R	**CCA CTG CAG TTA** TTA TTT GGT ACG AGA CGG GAT C
*Ec*ApeK-NdeI-F	**CCA CAT ATG** GTT CTG AAC GAC CCG CGT TTC
*Ec*ApeK-XhoI-R	**CCA CTC GAG TTA** CGG TTT AAC ACC CAT ACG TTC G
*Ec*ApeE-BamHI-F	**CCA GGA TCC** CAG GCT CTG TAC CTG GAA ATC
*Ec*ApeE-PstI-R	**CCA CTG CAG TTA** AGC ACG CTG AGC AGC
*Ec*ApeF-BamHI-F	**CCA GGA TCC** ACC GAA CAG CAG ACC GTT TAC C
*Ec*ApeF-PstI-R	**CCA CTG CAG TTA** CGC TTC TTG CAG CAG ACG TTC
*Ec*ApeQ-BamHI-F	**CCA GGA TCC** TCT CGT TCT GTT CTG GTT ACC
*Ec*ApeQ-PstI-R	**CCA CTG CAG TTA** CAG CAT ACC ACC GTT GAT AGA G
Mutagenesis primer	
*Xd*ApeF$^{\mathrm{leu60pro}}$-F	CGG ACG CAA AAT CAA ACC GGA AGC ATT CAA GTC TG
*Xd*ApeF$^{\mathrm{leu60pro}}$-R	CAG ACT TGA ATG CTT CCG GTT TGA TTT TGC GTC CG
*Xd*ApeK$^{\mathrm{asp21asn}}$-F	GCC ATT TCA TGA TGC TAA TCC CAT GGG CGT GG
*Xd*ApeK$^{\mathrm{asp21asn}}$-R	CCA CGC CCA TGG GAT TAG CAT CAT GAA ATG GC
*Af*yg1p$^{\mathrm{ser256ala}}$-F	GGT TGT TTG GGG TCT GGC TGC TGG TGG TTA CTA
*Af*yg1p$^{\mathrm{ser256ala}}$-R	TAG TAA CCA CCA GCA GCC AGA CCC CAA ACA ACC
*Af*yg1p$^{\mathrm{asp351asn}}$-F	GCT GAA CGG TGT TGA CAA CGG TGT TGT TCC GAT
*Af*yg1p$^{\mathrm{asp351asn}}$-R	ATC GGA ACA ACA CCG TTG TCA ACA CCG TTC AGC

Primer overhangs are shown in bold (contain sequence for binding site of restriction enzymes and and stop codon in reverse primers). Modified codons are underlined.

4.1.6 Bacterial strains and growth-media

Table 4.13 Bacterial strains

XL10-Gold	*E. coli* K Tet$^{\mathrm{r}}$ Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*) *173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacI*$^{\mathrm{q}}$*Z*Δ*M15* Tn*10* (Tet$^{\mathrm{r}}$) Amy Cam$^{\mathrm{r}}$].	Agilent, Santa Clara, US
BL21(DE3)	*E. coli* B F^{-} *dcm ompT hsdS*(r_{B}^{-} m_{B}^{-}) *gal* λ(DE3)	Novagen, Darmstadt, DE
BL21-Gold(DE3)	*E. coli* B F^{-} *ompT hsdS*(r_{B}^{-} m_{B}^{-}) *endA* Hte *dcm*$^{+}$ Tet$^{\mathrm{r}}$ *gal* λ(DE3)	Agilent, Santa Clara, US

Table 4.14 Bacterial growth-media

LB medium	Peptone	1 % (w/v)
	Yeast extract	0.5 % (w/v)
	NaCl	0.5 % (w/v)
	(Agar)	2 % (w/v)
SOC medium	Peptone	2 % (w/v)
	Yeast extract	0.5 % (w/v)
	Glucose	20 mM
	$MgSO_4 \cdot 7\,H_2O$	10 mM
	$MgCl_2 \cdot 6\,H_2O$	10 mM
	NaCl	10 mM
	KCl	2.5 mM
M9 minimal medium	$Na_2HPO_4 \cdot 2\,H_2O$	42 mM
	KH_2PO_4	22 mM
	NaCl	8.5 mM
	NH_4Cl	9.3 mM
	$MgSO_4 \cdot 7\,H_2O$	2 mM
	Glucose	0.4 % (w/v)
	Vitamins (1,000×, **Tab. 4.15**)	0.1 % (v/v)
	Trace elements (100×, **Tab. 4.15**)	1 % (v/v)

Table 4.15 Additives for recombinant expression of selenomethionine-labelled protein

1L vitamins pH 7.0 (1,000×)	EDTA	5 g
	$FeCl_3$	0.8 g
	$ZnCl_2$	0.05 g
	$CuCl_2$	0.01 g
	$CoCl_2$	0.01 g
	H_3BO_3	0.01 g
	$MnCl_2$	1.6 g
	Ni_2SO_4	some
	molybdic acid	some
1L trace elements (100×)	riboflavin	1 g
	niacinamide	1 g
	pyridoxine monohydrate	1 g
	thiamine	1 g
Feed-back inhibition amino acid mix for 3L	lysine, threonine, phenylalanine	0.3 g
	leucine, isoleucine, valine	0.15 g
	L(+) selenomethionine	0.15 g

4.1.7 Buffers and solutions

Table 4.16 Buffers A and B for affinity chromatography

Buffer A-S pH 8.0	100 mM tris 500 mM NaCl 5 mM DTT	Standard buffer for Strep-tagged proteins: ***Wd*yg1p**, ***Ec*ApeE:K**, and ***Xd*ApeR**
Buffer A-LS pH 8.0	100 mM tris 100 mM NaCl 5 mM DTT	Low-salt buffer for Strep-tagged protein: ***Xd*ApeO:C**
Buffer A-MS pH 8.0	100 mM tris 500 mM NaCl 5 mM $MgCl_2 \cdot 6\,H_2O$ 5 mM DTT	Magnesium buffer for Strep-tagged proteins: ***Xd*ApeK**, ***Xd*ApeF:K**, and ***Xd*ApeI:P**
Buffer A-MLS pH 8.0	100 mM tris 100 mM NaCl 5 mM $MgCl_2 \cdot 6\,H_2O$ 5 mM DTT	Low-salt magnesium buffer for Strep-tagged proteins: ***Xd*ApeE:K**, ***Xd*ApeF:K**^{D21N}, and ***Ec*ApeF:K**
Buffers B-S, B-LS, B-MS and B-MLS	same as buffers A above +2.5 mM *d*-Desthiobiotin	Elution buffers for Strep-tagged proteins
Buffer A-H pH 7.5	100 mM tris 500 mM NaCl 20 mM imidazole 5 mM ß-ME	Standard buffer for His_6-tagged proteins: ***Af*yg1p** and ***Xd*ApeF**
Buffers A-MH pH 7.5/8.0	100 mM tris 500 mM NaCl 20 mM imidazole 5 mM $MgCl_2 \cdot 6\,H_2O$ 5 mM ß-ME	Magnesium buffer for His_6-tagged proteins: ***Xd*ApeQ** (pH 7.5) ***Ec*ApeQ** (pH 8.0)
Buffers A-LH pH 8.0	100 mM tris 300 mM NaCl 20 mM imidazole 10 % (v/v) glycerol 2 mM ß-ME	Low-salt buffer for His_6-tagged protein: ***Pl*AntI**
Buffers B-H and B-MH pH 7.5/8.0	same as buffers A-H/MH, +200 mM imidazole	Elution buffers for His_6-tagged proteins: ***Xd*ApeQ** (pH 7.5) and ***Ec*ApeQ** (pH 8.0)
Buffer B-LH pH 8.0	same as buffer A-LH +500 mM imidazole	Elution buffer for His_6-tagged protein: ***Pl*AntI**

Buffer conditions were extensively screened. Only optimized buffer conditions used for successful purification, crystallization, or other experiments are listed. A = wash buffer, B = elution buffer, H = buffer for His_6-tagged protein, S = buffer for Strep-tagged protein, L = low-salt buffer, M = Mg^{2+}-buffer.

Table 4.17 Buffers C for size-exclusion chromatography

Buffer C pH 7.5/8.0	20 mM tris 100 mM NaCl 2 mM DTT	Standard buffers for size-exclusion chromatography: ***Af*yg1p** and ***Pl*AntI** (pH 7.5) ***Wd*yg1p**, ***Ec*ApeE:K**, and ***Xd*ApeR** (pH 8.0)
Buffer C-R pH 8.0	20 mM tris 20 mM NaCl 2 mM DTT	Low-salt buffers for size-exclusion chromatography: ***Xd*ApeO:C**
Buffer C-M pH 7.5/8.0	20 mM tris 100 mM NaCl 1 mM $MgCl_2 \cdot 6\,H_2O$ 2 mM DTT	Magnesium buffers for size-exclusion chromatography: ***Xd*ApeQ** (pH 7.5) ***Ec*ApeQ**, ***Xd*ApeK**, ***Xd*ApeF:K**, and ***Xd*ApeI:P** (pH 8.0)
Buffer C-MR pH 8.0	20 mM tris 20 mM NaCl 1 mM $MgCl_2 \cdot 6\,H_2O$ 2 mM DTT	Low-salt magnesium buffers for SEC: ***Xd*ApeE:K**, ***Xd*ApeF:K**$^{\mathrm{D21N}}$, and ***Ec*ApeF:K**
Buffer C-Cit pH 4.0/4.3/4.7/5.0	20 mM Citric acid 100 mM NaCl 2 mM DTT	Buffers for size-exclusion chromatography: Stability test of ***Xd*ApeF:K**

Buffer conditions were extensively screened. Only optimized buffer conditions used for successful purification, crystallization, or other experiments are listed. C = SEC buffer, R = low-salt buffer, M = Mg^{2+}-buffer, Cit = citric acid buffer.

Table 4.18 Electrophoresis buffers and solutions

Agarose electrophoresis		
TAE running buffer (50×)	tris/Ac, pH 8.5	2 M
	EDTA	100 mM
tris-glycine SDS-PAGE		
tris-glycine running buffer (10×)	tris	250 mM
	Glycine	1.92 M
	SDS	1 % (w/v)
tris-glycine stacking buffer	tris/HCl, pH 6.8	0.5 M
	SDS	0.4 % (w/v)
tris-glycine separation buffer	tris/HCl, pH 8.8	1.5 M
	SDS	0.4 % (w/v)
tris-tricine SDS-PAGE		
tris-tricine anode running buffer (10×)	tris/HCl, pH 8.9	1 M
tris-tricine cathode running buffer (10×)	tris	1 M
	tricine	1 M
	SDS	1 % (w/v)
tris-tricine gel buffer (3×)	tris/HCl, pH 8.45	3 M
	SDS	0.3 % (w/v)
SDS sample buffer (5×)	tris/HCl, pH 6.8	60 mM
	glycerol	30 % (v/v)
	sucrose	10 % (w/v)
	SDS	5 % (w/v)
	ß-ME	3 % (v/v)
	bromophenol blue	0.02 % (w/v)
Staining solution	coomassie brilliant blue	0.05 % (w/v)
	isopropanol	25 % (v/v)
	acetic acid	10 % (v/v)
Destaining solution	acetic acid	10 % (v/v)

4.1.8 Software and online tools

Table 4.19 Software

The PyMOL Molecular Graphics System Version 2.4.0	Schrödinger, New York, US
CorelDRAW 2019 Version 21.0.0.593	Corel Corporation, Ottawa, CA
ChemDraw Professional Version 19.1.0.8	PerkinElmer, Waltham, US
Coot Version 0.8.8	[113]
CCP4i Version 7.1	[114]
CCP4i2 Version 1.1.0	[115]
REFMAC Version 5.8.0267	[116]
GraphPad Prism 5 Version 5.02	GraphPad Software, San Diego, US
XDS/XSCALE package	[117,118]
MAIN 2019	[119]
Phaser Version 2.8.3	[120]
ARP/wARP Version 8	[121,122]
A plasmid Editor Version 2.0.61	[123]
PHENIX	[124,125]
CFX Maestro 2.2	Bio-Rad, Hercules, US
Adobe Acrobat Reader Version 2022.003.20282	Adobe, San José, US
OriginPro 2020 Version 9.7.0.185	OriginLab, Northampton, US
CRANK2	[126,127]
MolScript Version 2	[128]
Topological and atomic comparison program (TOPP)	[129]

Table 4.20 Online tools

Basic Local Alignment Search Tool (Blast)	https://blast.ncbi.nlm.nih.gov/Blast.cgi[130]
Expasy ProtParam tool	https://web.expasy.org/protparam/[131]
DrugScore$^{\mathrm{PPI}}$ Version 2.2	Heinrich-Heine-University, Düsseldorf, DE[85]
PDBePISA	https://www.ebi.ac.uk/pdbe/pisa/[71]
JPRED 4	https://www.compbio.dundee.ac.uk/jpred/[132]
DALI protein structure comparison server	http://ekhidna2.biocenter.helsinki.fi/dali/[90]
antiSMASH 6.1.1 bacterial version	https://antismash.secondarymetabolites.org/[82]
Alphafold2 Colab	ColabFold[97]
ESPript Version 3.0	https://espript.ibcp.fr/[133]
ENDscript Version 2.0	https://endscript.ibcp.fr/[133]
UNIversal PROTein database (UniProt)	https://www.uniprot.org[134]
Alignments Simple(/Structural) T-Coffee (Expresso)	https://tcoffee.crg.eu/[135]
Sequence Manipulation Suite: Ident and Sim	www.bioinformatics.org/sms2/ident_sim.html[136]
bioCOmplexes COntact MAPS (COCOMAPS)	www.molnac.unisa.it/BioTools/cocomaps/[70]
RCSB Protein Data Bank (PDB)	https://www.rcsb.org/[137]
NEB Ligation Calculator	New England Biolabs, Ipswich, US
T_m Calculator	Thermo Fisher Scientific, Waltham, US[138]
MolProbity Version 4.5.1	Richardson's lab, Duke University, US[139]
wwPDB Validation System	https://validate-rcsb-1.wwpdb.org/[140]
QuikChange Primer Design tool	Agilent, Santa Clara, US
Java Codon Adaptation Tool (JCAT)	http://www.prodoric.de/JCat[141]
ShareLaTeX @ TUM	https://sharelatex.tum.de/

4.2 Methods

4.2.1 Molecular biology

Polymerase chain reaction

All gene fragments used for restriction enzyme-based cloning were codon optimized using the online tool JCAT[141] and synthesized by Eurofins Genomics (Ebersberg, DE). The genes were amplified by polymerase chain reaction (PCR[142]) using primers that contained overhang sequences with restriction enzyme binding sites (**Tab. 4.12**). The PCR program for the thermal cycler and PCR reaction mix were as follows:

Table 4.21 PCR reaction program (left) and reaction mixture (right)

Step	T	t	cycles
denaturation	98°C	30 s	1
denaturation	98°C	10 s	
annealing[a]	T_A	30 s	25
elongation[b]	72°C	t_E	
elongation	72°C	10 min	1
elongation	10°C	∞	1

amount	reagent
5 µL	primer forward[c](100 pmol µL^{-1})
5 µL	primer reverse[c](100 pmol µL^{-1})
1 µL	DNA template[d](10 ng µL^{-1})
2 µL	dNTPs (10 mmol L^{-1})
20 µL	Q5 reaction buffer (5×)
1 µL	Q5-HF polymerase (2 U µL^{-1})
66 µL	ddH_2O

[a]Annealing temperature was calculated using the T_m Calculator (Thermo Fisher Scientific, Waltham, US) [b]Time of elongation was adjusted according to length of DNA template (30 s kb^{-1}) [c]Primers are listed in **Tab. 4.12** [d]Genes are listed in **Tab. 4.10**, **Tab. A.1**, and **Tab. A.2**. T = temperature and t = time.

The quality and quantity of PCR products were analyzed by agarose gel electrophoresis and purified using Wizard SV Gel and PCR Clean-Up System Kit (**Tab. 4.3**) following the manufacturer's instructions.

Agarose gel electrophoresis

DNA samples were spiked with gel loading dye (purple, 6×, New England Biolabs, Ipswich, US) and loaded onto 1 % (w/v) agarose gels in TAE buffer that contained 1 µL Serva DNA Stain G (Serva, Heidelberg, DE, **Tab. 4.18**). Electrophoresis was performed in TAE buffer (1×) for 20 minutes at 120 V. Gels were analyzed using a G:Box Documentation System (Syngene, Bangalore, IN).

Restriction enzyme digestion

Reaction mix for analytical digestion of plasmid DNA or preparative digestion of purified PCR-product or plasmid is summarized in **Tab. 4.22**. The reaction mixtures were incubated at 37°C for 20 minutes (analytical) or 2 hours (preparative) and purified using Wizard SV Gel and PCR Clean-Up System. Preparative samples were used for ligation, whereas analytical samples were subsequently analyzed by agarose gel electrophoresis and sequenced if necessary.

Table 4.22 Restriction enzyme digestion

preparative	analytical	reagent
1-2 µg	∼0.2 µg	PCR product or plasmid
2 µL	0.2 µL	restriction enzyme I (20 U/µL)[a]
2 µL	0.2 µL	restriction enzyme II (20 U/µL)[a]
7 µL	1 µL	CutSmart Buffer (10×)
Ad 70 µL	Ad 7 µL	ddH_2O

[a] Restriction enzymes listed in **Tab. 4.2**.

Ligation

The amounts for a molar 7:1 ratio of digested insert to 50 ng of vector DNA were calculated using the Ligation Calculator (New England Biolabs, Ipswich, US). The total volume of 8.5 µL (addition of ddH_2O) was incubated at 55°C for 10 minutes, followed by incubation on ice for 5 minutes. After the addition of 1 µL T4 DNA Ligase Reaction Buffer (10×) and 0.5 µL T4 DNA Ligase, the reaction mixture was incubated at 16°C overnight, and the next day heat deactivated by incubation at 65°C for 10 minutes. Afterward, 1 µL was transformed in electrocompetent XL10-Gold *E. coli* cells (**Tab. 4.13**).

Transformation

0.5 ng of plasmid DNA was mixed with electrocompetent *E. coli* cells (prepared by technicians) and transferred in MicroPulser Electroporation Cuvettes (2 mm gap, **Tab. 4.6**). After the electrical pulse (≈ 5 ms, 2500 V)[143,144] the reaction mixture was diluted with 1 mL of fresh SOC medium (**Tab. 4.14**) and incubated at 37°C and 125 revolutions per minute (rpm) for 1 hour. Afterward, the cells were plated on LB-agar plates with appropriate antibiotics (see below).

E. coli cultivation and storage

All experiments with *E. coli* cells involved were performed under sterile conditions. Additionally, LB medium, minimal medium, and agar plates contained the appropriate antibiotics (50 µg mL^{-1} kanamycin A, Kan; 25 µg mL^{-1} chloramphenicol, Cm; 180 µg mL^{-1} ampicillin, Amp, or 50 µg mL^{-1} spectinomycin, Spec). *E. coli* single clones were used to inoculate 5 mL of LB medium, and the cultures were incubated at 37°C and 125 rpm overnight to reach saturation. Typically, glycerol stocks were generated for long-term storage. Here, after centrifugation at 16,000 g for 30 seconds, the pellet of 1 mL culture was resuspended in 1 mL LB medium containing 30 % (v/v) glycerol, transferred into cryo vials, flash frozen, and stored in liquid nitrogen. The remaining culture was used to inoculate large-scale cultures or to isolate vector DNA (see below).

Isolation of vector DNA from *E. coli*

The Wizard Plus SV Minipreps DNA Purification Systems Kit (**Tab. 4.3**) was used according to the manufacturer's instructions.

Site-directed mutagenesis

Mutagenesis primers were designed with the QuikChange Primer Design tool (Agilent, Santa Clara, US) and synthesized by Eurofins Genomics (Ebersberg, DE). The site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit according to the manufacturer's instructions. The mutations were verified by sanger-sequencing (GATC, Eurofins Genomics, Ebersberg, DE).

Determination of DNA concentration

To determine the mass concentration ρ_{DNA} of double-stranded (ds-DNA) its absorbance at a wavelength of 260 nm (A_{260}) was measured using a nanodrop 2000c spectrophotometer. After baseline correction at 340 nm, the concentration ρ_{DNA} in $\mathrm{ng\,\mu L^{-1}}$ was calculated according to Beer–Lambert law using the average extinction coefficient for ds-DNA (ϵ = 0.02 $\mathrm{\mu L\,ng^{-1}\,cm^{-1}}$) and the path length b = 1 cm:

$$\rho_{\mathrm{DNA}} = \frac{\mathrm{A}_{260}}{\epsilon \times \mathrm{b}} \tag{4.1}$$

4.2.2 Protein chemistry

Protein concentration

The absorbance of a protein sample was measured at a wavelength of 280 nm (A_{280}) using a nanodrop 2000c spectrophotometer. After a baseline correction (340 nm), the molar concentration $\mathrm{c}_{\mathrm{protein}}$ in $\mathrm{mol\,L^{-1}}$ was calculated according to **Eqn. 4.2** with a path length of b = 1 cm. The protein's molar extinction coefficient $\epsilon_{\mathrm{molar}}$ in $\mathrm{L\,mol^{-1}\,cm^{-1}}$ and molecular weight M in $\mathrm{g\,mol^{-1}}$ were calculated using the Protparam tool[131].

$$\mathrm{c}_{\mathrm{protein}} = \frac{\mathrm{A}_{280}}{\epsilon_{\mathrm{molar}} \times \mathrm{b}} \tag{4.2}$$

The mass concentration ρ_{protein} in $\mathrm{mg\,mL^{-1}}$ was calculated by multiplying $\mathrm{c}_{\mathrm{protein}}$ with the molecular weight:

$$\rho_{\mathrm{protein}} = \mathrm{c}_{\mathrm{protein}} \times \mathrm{M} \tag{4.3}$$

SDS-PAGE analysis

Quantity and purity of proteins were monitored by denaturing SDS-PAGE according to Laemmli et al.[145] or Schägger et al.[146] The samples were mixed with SDS sample buffer (5×), heated at 95°C for 5 minutes, and loaded onto SDS-gels. Electrophoresis of SDS-PAGE gels was

performed in respective running buffers (1×) at a constant current of 35 mA per gel for tris-glycine gels or at a constant voltage of 30 V for stacking and 190-270 V for separation per tris-tricine gel. The preparation of buffers and gels is summarized in **Tab. 4.18** and **Tab. 4.23**. Gels were transferred in a staining solution, boiled, and incubated for at least 30 minutes. Afterward, gels were boiled in a destaining solution and incubated at RT overnight before analyzing with a G:Box UV/White Light Gel Documentation System (Syngene, Bangalore, IN).

Table 4.23 Preparation of SDS gels[a]

tris-glycine SDS-PAGE		
tris-glycine separating gel	tris-glycine separation buffer (4×)	5 mL
	Acrylamide[a]	10 % gel: 5 mL
		12 % gel: 6 mL
		15 % gel: 7.5 mL
	ddH_2O	Ad 20 mL
	APS (10 % w/v)	100 µL
	TEMED	10 µL
tris-glycine stacking gel	tris-glycine stacking buffer (2×)	5 mL
	Acrylamide[a]	1 mL
	ddH_2O	Ad 10 mL
	APS (10 % w/v)	100 µL
	TEMED	10 µL
tris-tricine SDS-PAGE		
tris-tricine separating gel	tris-glycine gel buffer (3×)	10 mL
	Acrylamide[a]	10 % gel: 6 mL
	glycerol	3 g
	ddH_2O	Ad 30 mL
	APS (10 % w/v)	150 µL
	TEMED	15 µL
tris-tricine stacking gel	tris-glycine gel buffer (3×)	3 mL
	Acrylamide[a]	10 % gel: 1 mL
	ddH_2O	Ad 12 mL
	APS (10 % w/v)	100 µL
	TEMED	10 µL

[a]Solutions and buffers for electrophoresis are listed in **Tab. 4.18**. [a]Acrylamide (40 %, Mix 29:1, AppliChem, Darmstadt, DE).

Standard protein expression in *E. coli*

5 µL of glycerol stocks were used to inoculate 150 mL of LB medium and then grown at 37°C and 125 rpm overnight to reach saturation. The pre-culture was diluted 1:60 in 3 L of fresh LB medium in Fernbach culture flasks and incubated at 37°C and 125 rpm. The growth of the main-cultures was monitored by measuring the optical density at λ = 600 nm (OD_{600}). At an OD_{600} of 0.6-0.8, the main-cultures were directly induced with isopropyl-β-D-1-

thiogalactopyranoside (IPTG, final concentration of 1 $\mathrm{mmol\,L^{-1}}$) to induce protein expression or cooled down at 4°C for 30 minutes and induced thereafter. The *E. coli* cells were grown at 37°C for another 4 hours (directly induced) or at 16-25°C for 16-20 hours (cooled prior induction), harvested by centrifugation (20°C, 8,000 g, 20 minutes), washed with 0.9 % (w/v) NaCl solution, and stored at -20°C.

Recombinant expression of selenomethionine-labeled protein in *E. coli*

Selenomethionine-labeled protein was produced using the methionine feed-back inhibition method[147]. Briefly, a 5 mL pre-culture (LB medium) was used to inoculate 1:1,000 a 300 mL pre-culture in minimal medium (**Tab. 4.14** and **Tab. 4.15**). After incubation at 37°C overnight, the main-culture (minimal medium) was inoculated 1:100, and at an OD_{600} of 0.6-0.8, the feed-back inhibition amino acids mix (**Tab. 4.15**) and after another 15 minutes, IPTG was added. The expression of the target proteins and the harvest of *E. coli* cells were performed as described above.

Protein purification

Typically, a 5 g cell pellet was thawed on ice and resuspended in 50 mL buffer A (respective buffers of all constructs are listed in **Tab. 4.24**). After addition of a spatula tip('s) of pefabloc (Roche, Basel, CH), the cells were lysed on ice by sonication (5 minutes, 70 % amplitude, 1 second on/off). The cells were centrifuged at 40,000 g for 20 minutes at 4°C, and the supernatant was applied to an affinity chromatography column (HisTrap$^{\mathrm{HP}}$ or StrepTrap$^{\mathrm{HP}}$ 5 mL, Cytiva, Marlborough, US) that had been previously equilibrated with buffer A (**Tab. 4.16**). The bound protein was washed with buffer A and eluted with the respective buffer B (**Tab. 4.16**). After the addition of SUMO- or TEV-protease (**Tab. 4.2**), the pooled eluates were dialyzed at 4°C overnight against buffer A. The tag-free protein was isolated by reverse affinity chromatography using the same buffers and methods as described above. Here, the flow-through fractions were pooled, dialyzed against buffer C at 4°C overnight, and loaded onto a size-exclusion chromatography column (Superdex HiLoad 16/600 pg 75 or 200, Cytiva, Marlborough, US) that had been previously equilibrated with buffer C (**Tab. 4.17**).

Table 4.24 Protein purification of all constructs

Construct[a]	buffer A/B[b]	cleavage	buffer C[c]	additional remarks
Strep-tagged:				
***Wd*yg1p**	A-S/B-S pH 8.0	no	C pH 8.0	same conditions for mutants, no pefabloc added
***Ec*ApeE:K**	A-S/B-S pH 8.0	yes	C pH 8.0	
***Xd*ApeR**	A-S/B-S pH 8.0	yes	C pH 8.0	
***Xd*ApeO:C**	A-LS/B-LS pH 8.0	no	C-R pH 8.0	
***Xd*ApeK**	A-MS/B-MS pH 8.0	yes	C-M pH 8.0	***Xd*ApeK** was obtained from ***Xd*ApeE:K** construct
***Xd*ApeF:K**	A-MS/B-MS pH 8.0	yes	C-M pH 8.0	
***Xd*ApeI:P**	A-MS/B-MS pH 8.0	no	C-M pH 8.0	
***Xd*ApeE:K**	A-MLS/B-MLS pH 8.0	yes	C-MR pH 8.0	
***Xd*ApeF:K**	A-MLS/B-MLS pH 8.0	yes	C-MR pH 8.0	
***Ec*ApeF:K**	A-MLS/B-MLS pH 8.0	yes	C-MR pH 8.0	
His_6-tagged:				
***Af*yg1p**	A-H/B-H pH 7.5	no	C pH 7.5	same conditions for mutants, no pefabloc added
***Xd*ApeF**	A-H/B-H pH 7.5	no	C pH 7.5	
***Xd*ApeQ**	A-MH/B-MH pH 7.5	yes	C-M pH 7.5	NADPH-addition before size-exclusion chromatography
***Ec*ApeQ**	A-MH/B-MH pH 8.0	no	C-M pH 8.0	(2.5 mM final concentration)
***Pl*AntI**[d]	A-LH/B-LH pH 8.0	no	C pH 7.5	

[a]Constructs listed in **Tab. 4.11**, [b]Buffers for affinity chromatography listed in **Tab. 4.16**, [c]Buffers for size-exclusion chromatography listed in **Tab. 4.17**, [d]purification modified from Zhou et al.[26]

The combined fractions were concentrated to typically 10-50 $mg\,mL^{-1}$ using 3 kDa, 10 kDa, or 30 kDa molecular weight cut-off (MWCO) Amicon Ultra-15 Centrifugal Filters and flash frozen in liquid nitrogen for storage at -80°C. Successful cleavage of the tag and protein's purity were monitored by SDS-PAGE during all purification steps. Purifications of selenomethionine-labeled protein constructs (*Af*yg1p and *Xd*ApeF:K) were performed identically compared to the purification of the native proteins.

4.2.3 Protein crystallization

The sitting drop vapor diffusion method was used to screen for initial crystallization conditions at 20°C. The 96-well Intelli-Plates (Art Robbins Instruments, Sunnyvale, US) with NeXtal Tubes Suite screens (Qiagen, Hilden, DE) were prepared with a Crystal Liquid Handling System Phoenix or Gryphon (Art Robbins Instruments, Sunnyvale, US). Hereafter, protein and reservoir solutions were mixed (0.2 µL+0.2 µL, 0.2 µL+0.1 µL, and 0.3 µL+0.1 µL) with the Crystal Phoenix or Gryphon or the Oryx4 Protein Crystallization Robot (Douglas Instruments, East Garston, UK). Identified crystallization parameters were further optimized by applying the sitting drop and hanging drop vapor diffusion methods. After the addition of

cryoprotectants, the crystals were vitrified in liquid nitrogen and used for the collection of diffraction data at the SLS, Villigen, CH (beamline X06SA at Paul Scherrer Institute). The optimized crystallization conditions of all constructs are summarized in **Tab. 4.25**.

Table 4.25 Crystallization conditions

Construct	ρ_{protein} (mg mL^{-1})	reservoir solution	ratio[a]	cryoprotectant 15-30 % (v/v)
Hanging drop vapor diffusion:				
***Wd*yg1p**	20	0.1 M tris/HCl, pH 8.0-9.0 22-25 % (w/v) PEG3350	1:1	ethylene glycol
***Af*yg1p**	15	0.1-0.2 M K/Na tartrate 15-26 % (w/v) PEG3350	1:1	ethylene glycol
***Pl*AntI**	20	0.1 M hepes, pH 7.0-8.0 0.1 M NaOAc, pH 6.6 20-25 % (w/v) PEG3350	1:1	ethylene glycol
***Xd*ApeE:K**[b]	15	0.1 M mes, pH 6.0 0.1 M $MgCl_2$, 1 % (w/v) PEG6000	1:1	ethylene glycol
Sitting drop vapor diffusion:				
***Xd*ApeF:K**$^{\mathrm{SeMET,\ c}}$	16.7	0.1 M hepes, pH 7.5 25 % (w/v) PEG3350	2:1	glycerol
***Xd*ApeR**	14.5	0.1 M hepes, pH 7.0 1 M LiCl, 30 % (w/v) PEG6000	1:1	glycerol
***Ec*ApeQ**	+2.5 mM NADPH	0.1 M bistris, pH 6.5 0.1 M NH_4OAc, 49.5 % (v/v) MPD	1:1	none
Xd*ApeK** (Xd*ApeE:K** was used)[d]	15	0.1 M citric acid, pH 4.0 30 % (v/v) MPD	1:1	none
***Xd*ApeI:P**	20	0.1 M hepes, pH 7.0 30 % (w/v) PEG6000	1:1	glycerol
***Xd*ApeO:C**	18.7	1 M Na/K phosphate, pH 8.0	4:1	glycerol
***Af*yg1p**$^{\mathrm{SeMET}}$	31.5	0.1 M bistris, pH 5.5 25 % (w/v) PEG3350	2:1	ethylene glycol

[a]Ratio of protein to reservoir sulution in µl [b]**Several rounds of seeding were necessary** [c]Best resolution was achieved with selenomethionine-labeled *Xd*ApeF:K [d]Under these conditions the crystallization of *Xd*ApeE:K did yield *Xd*ApeK crystals.

Incorporation of ligands

Crystals of all described protein:ligand complex structures were grown under conditions that were used for the native crystals as well. The applied methods were soaking or co-crystallization experiments. The main parameters screened in hanging-drop vapor diffusion experiments were pH, solvent, the concentration of ligand or protein, and duration of the soaking. The soaking and co-crystallization conditions of all protein:ligand structures are summarized in **Tab. 4.26**. Notably, *Wd*yg1p did predominantly crystallize as small needles that were hardly reproducible. However, in the presence of ligands (e.g., DHNs or THN), reproducible growth of crystals was accomplished even without the ligand resolved in the crystal structures.

Table 4.26 Crystallization of protein:ligand

protein:ligand	ρ_{protein} (mg mL^{-1})	reservoir solution	$c_{\text{ligand}}^{\text{final}}$ (mM)	$V_{\text{ligand}}^{\text{stock}}$ (µL)	$c_{\text{ligand}}^{\text{stock}}$ (mM)
Soaking:					
***Af*yg1p**$^{\text{PMS}}$	25	0.15 M K/Na tartrate 25 % (w/v) PEG3350	4	2.5	7[a]
***Wd*yg1p**$^{\text{Dha,*}}$	31	0.1 M tris/HCl, pH 9.2 22 % (w/v) PEG3350	4	2.5	7[a,c]
***Wd*yg1p**$^{\text{Dha,\#}}$	29	0.1 M tris/HCl, pH 9.0 23 % (w/v) PEG3350	5	0.3	33.3[b,d]
***Pl*AntI**$^{\text{Dha,*}}$	35	0.1 M hepes, pH 7.1 0.1 M NaOAc, pH 6.6 24 % (w/v) PEG3350	5	0.1	100
***Pl*AntI**$^{\text{Dha,\#}}$	25	0.1 M hepes, pH 7.1 0.1 M NaOAc, pH 6.6 24 % (w/v) PEG3350	5	0.1	100
***Pl*AntI:1-N**	15	0.1 M hepes, pH 7.9 0.1 M NaOAc, pH 6.6 21 % (w/v) PEG3350	10	0.1	200
***Af*yg1p:1,3-DHN**	50	0.1 M K/Na tartrate 25 % (w/v) PEG3350	10	0.1	200
Co-crystallization:					
***Pl*AntI:PAC**	20	0.1 M hepes, pH 7.5 0.1 M NaOAc, pH 6.6	0.5		
***Wd*yg1p:1,3-DHN**	20	0.1 M tris/HCl, pH 8.7 24 % (w/v) PEG3350	0.5		
***Af*yg1p:THN**	20	0.1 M K/Na tartrate 25.5 % (w/v) PEG3350	1		
***Wd*yg1p:THN**	31	0.1 M tris/HCl, pH 8.5 24 % (w/v) PEG3350 21 % (w/v) PEG3350	0.5		
***Pl*AntI:CA**	15	0.1 M hepes, pH 7.1 0.1 M NaOAc, pH 6.6 25 % (w/v) PEG3350	5		

Sulfonyl fluorides used were PMSF (*) and PSF ($^{\#}$). Stock solution of ligands (100 mM) were diluted 1:14 ($^{\text{a}}$) and 1:3 ($^{\text{b}}$) with reservoir solution. Soaking was performed overnight at 20°C. Exceptions were $^{\text{c}}$ and $^{\text{d}}$ with 2 and 6 hours, respectively.

4.2.4 Structure determination and data analysis

The diffraction images were recorded using synchrotron radiation (λ = 1.0 Å) at the beamline X06SA of the Swiss Light Source (SLS, Paul Scherrer Institute, Villigen, CH). The obtained reflection intensities were evaluated with the XDS suite[117,118], and data reductions were performed using XSCALE[117,118]. If necessary, the topology and geometric restraints of the ligands were calculated using ProDRG[148] or AceDRG[149]. Solvent content and the number of protomers were estimated based on the calculation of the Matthews coefficients[150] with consideration of molecular weight and oligomeric state of the proteins and the crystal parameters (**Tab. A.3** to **A.10**). This information was combined with solved or modeled (us-

ing Alphafold2[97]) coordinates that possess high structural and/or sequence similarity) and used for the generation of initial phases by Patterson Search calculations[120]. Experimental phases for *Af*yg1p were determined by single-wavelength anomalous dispersion (SAD[151]) methods at the peak absorption wavelength of selenium incorporated into a single crystal of *Af*yg1p$^{\mathrm{SeMet}}$ (λ = 0.97914 Å, f' = -7.58, f" = 5.56). CRANK2[126,127] was used for the automatic X-ray structure solution of the anomalous dataset (2.2 Å resolution, **Tab. A.5**). The hereby obtained model was sufficient for the phase determination of *Af*yg1p$^{\mathrm{native}}$ and *Wd*yg1p$^{\mathrm{native}}$. The structures were optimized by restrained refinements in iterative rounds (REFMAC5[116] or PHENIX[124,125]), and model building was carried out with Coot[113], Main[119], or PHENIX Autobuild[124,125]. The structures were finally refined in several cases with Translation/Libration/Screw. Water molecules were automatically placed using ARP/wARP solvent[121,122]. The obtained X-ray structures were validated using the online tools MolProbity[139] and wwPDB Validation System[140], and the statistics are summarized in **Tables A.3** to **A.10**.

Evaluation of homologous proteins and analysis of obtained homo- or heterooligomeric complexes

DALI[90] protein structure comparison server was applied to identify homologous proteins based on their 3D structures. Structural similarities between proteins were analyzed by superpositions performed by the topological and atomic comparison (TOPP[129]). Multiple sequence alignments were performed based on their amino acid sequences or on their structural information. For this, the T-Coffee[135] Simple Alignment or the Espresso[135] Structural Alignment, respectively, was applied, whereas the graphical illustration of the alignments was performed with the online tool ESPript[133] (small helices, sheets, and turns are not indicated for simplicity). Sequence identities and similarities were calculated with the Sequence Manipulation Suite: Ident and Sim tool[136]. Stable complexes were evaluated by calculating the interface areas, the complexation significance score (CSS, the value of 1 is the most stable), and the gain in solvation free energy (Δ^{I}G in $\mathrm{kcal\,mol^{-1}}$) with the online tool PDBePISA[71]. For validation of the putative interaction patches *in silico* alanine scanning was applied using the online tool DrugScore$^{\mathrm{PPI,\ 85}}$. To visualize the overall regions that participate in complex formation, the online tool bioCOmplexes COntact MAPS (COCOMAPS[70], distance threshold of 6 Å) was used. The electrostatic surface potentials of proteins were calculated by the Adaptive Poisson-Boltzmann Solver (APBS[152]) electrostatics plugin for PyMOL.

Preparation of figures

Protein structure figures were prepared using the PyMOL Molecular Graphics System (Schrödinger, New York, US), while *Xd*ApeO:C and ApeI:P figures were prepared with MolScript[128]. The data obtained in size exclusion experiments were analyzed with GraphPad Prism 5 (GraphPad Software, San Diego, US). All chemical drawings were generated with the software ChemDraw Professional (PerkinElmer, Waltham, US), and all figures were finalized with CorelDRAW 2019 (Corel Corporation, Ottawa, CA).

Additional information to X-ray crystal structures

Ketoreductase *Xd*ApeQ$_4$ To achieve the incorporation of the cofactor into the ketoreductase, high amounts of NADPH (BioCatalytics, Pasadena, US) were used. However, spectral analysis of the protein samples showed a mixture of NADPH and NADP$^+$ (**Fig. 4.1a**).

Ketosynthase *Xd*ApeR$_2$ In our obtained *Xd*ApeR$_2$ structure, some parts are disordered and, therefore, not resolved in the electron density map. Despite extensive screening and several trials, ApeR crystals used for the measurement of the dataset were not reproducible so far.

ACP:K complexes The electron density maps of some of the ACP chains contained locally disordered regions. A possible reason might include the high flexibility of the ACPs, as seen in the temperature factor (B-factor) distribution (**Fig. 4.1b**). In addition, the complexes seem partly unstable, as discussed above (**Fig. 2.8b**, **c**). Furthermore, the asymmetric unit of *Xd*ApeK:E contains 8 ApeK but only 6 ApeE subunits. Notably, the E chains are fully defined in their electron density maps. Interestingly, the Ppant arms fit the electron density maps in each chain due to their rigidity and intensive coordination with protein residues (**Fig. 4.1c**). To obtain the crystal structure of the quaternary ApeK$_4$:F$_4{}^{holo}$ complex, combinations of the native ApeK:F, selenomethionine-incorporated ApeK:F, and mutated ApeK:F$^{\mathrm{pro60leu}}$ were generated and purified. The best quality and resolution were achieved in a native dataset of selenomethionine-incorporated ApeK:F$^{\mathrm{pro60leu}}$, which adopts an identical fold, as seen in the native structure (overall complex: rmsd = 0.6 Å over 805 Cα-atoms, ApeF: rmsd = 0.5 Å over 75 Cα-atoms).

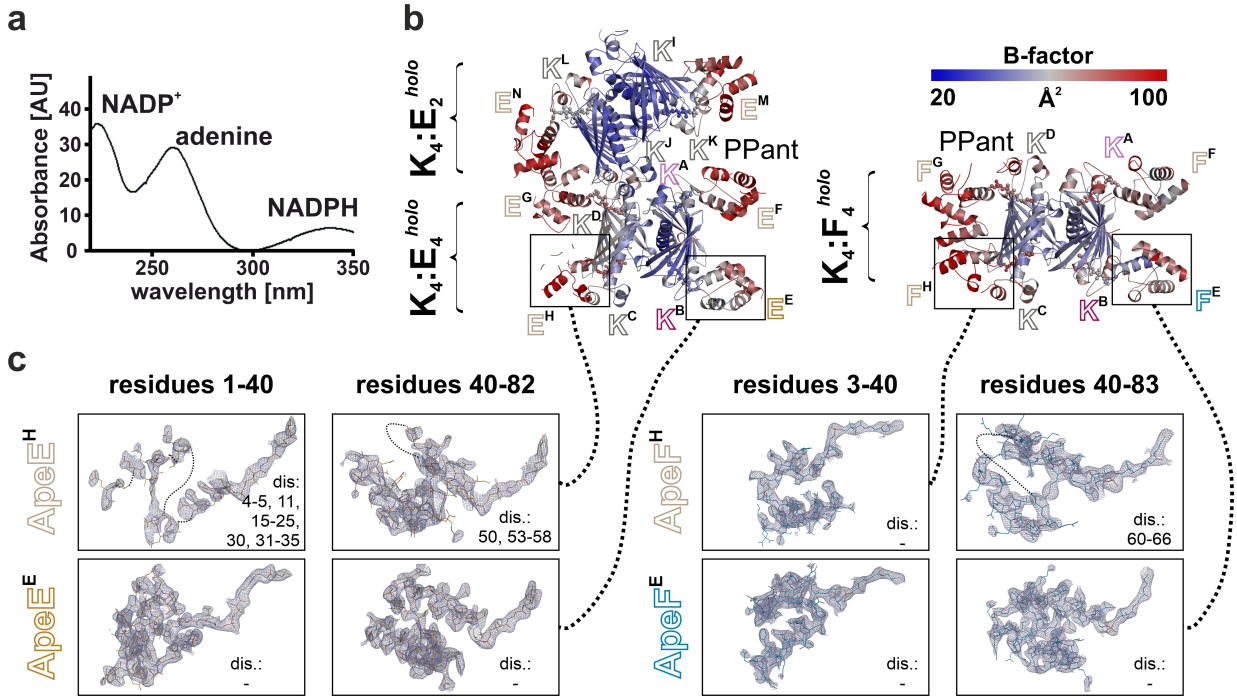

Figure 4.1 Quality of electron densities of *Xd*ApeK$_4$:E$_4{}^{\mathrm{holo}}$ and *Xd*ApeK$_4$:F$_4{}^{\mathrm{holo}}$. (a) UV-VIS spectrum of *Ec*ApeQ used for crystallization. The peaks for adenine, NADP$^+$, and NADPH are characteristic. **(b)** Representation of the contents of the asymmetric unit cells of XdApeK:E (left) and XdApeK:F (right). Residues are colored according to their B-factor values. **(c)** 2F$_O$-F$_C$ electron density maps (blue mesh, contoured to 1.0σ) are shown for the ACP Chains E (best quality) and H (poor quality). Disordered (dis) regions are listed.

A Appendix

Table A.1 Optimized genes used for molecular cloning of ApeQ, ApeE, ApeF, and ApeK constructs from *Escherichia coli* O157:H7 (strain Sakai)

Construct (Gene ID)	DNA sequence
***Ec*ApeQ (915807)**	TCT CGT TCT GTT CTG GTT ACC GGT GCT TCT AAA GGT ATC GGT CGT GCT ATC GCT TGC CAG CTG GCT GCT GAC GGT TTC AAC ATC GGT GTT CAC TAC CAC CGT GAC GCT ACC GGT GCT CAG GAA ACC CTG AAC GCT ATC GTT GCT AAC GGT GGT AAC GGT CGT CTG CTG TCT TTC GAC GTT GCT AAC CGT GAA CAG TGC CGT GAA GTT CTG GAA CAC GAA ATC GCT CAG CAC GGT GCT TGG TAC GGT GTT GTT TCT AAC GCT GGT ATC GCT CGT GAC GCT GCT TTC CCG GCT CTG TCT GAC GAC GAC TGG GAC GCT GTT ATC CAC ACC AAC CTG GAC TCT TTC TAC AAC GTT ATC CAG CCG TGC ATC ATG CCG ATG ATC GGT GCT CGT CAG GGT GGT CGT ATC ATC ACC CTG TCT TCT GTT TCT GGT GTT ATG GGT AAC CGT GGT CAG GTT AAC TAC TCT GCT GCT AAA GCT GGT ATC ATC GGT GCT ACC AAA GCT CTG GCT ATC GAA CTG GCT AAA CGT AAA ATC ACC GTT AAC TGC ATC GCT CCG GGT CTG ATC GAC ACC GGT ATG ATC GAA ATG GAA GAA TCT GCT CTG AAA GAA GCT ATG TCT ATG ATC CCG ATG AAA CGT ATG GGT CAG GCT GAA GAA GTT GCT GGT CTG GCT TCT TAC CTG ATG TCT GAC ATC GCT GGT TAC GTT ACC CGT CAG GTT ATC TCT ATC AAC GGT GGT ATG CTG TAA
***Ec*ApeE (916270)**	CAG GCT CTG TAC CTG GAA ATC AAA AAC CTG ATC ATC TCT ACC CTG AAC CTG GAC GAA CTG ACC CCG GAC GAC ATC GAC ACC GAC GCT CCG CTG TTC GGT GAC GGT CTG GGT CTG GAC TCT ATC GAC GCT CTG GAA CTG GGT CTG GCT GTT AAA AAC GAA TAC GGT ATC GTT CTG TCT GCT GAA TCT GAA GAA ATG CGT CAG CAC TTC TTC TCT GTT GCT ACC CTC GCG AGC TTC ATC GCT GCT CAG CGT GCT TAA
***Ec*ApeF (916293)**	ACC GAA CAG CAG ACC GTT TAC CAG GAA GTT TCT GCT CTG CTG GTT AAA CTG TTC GAA ATC GAC CCG CAG GAC ATC AAA CCG GAA ACC CGT CTG TAC GAA GAC CTG GAA CTG GAC TCT ATC GAC GCT GTT GAC ATG ATC GTT CAC CTG CAA AAA AAA ACC GGT AAA AAA ATC AAA CCG GAA GAA TTC AAA GCT GTT CGT ACC GTT CAG GAC GTT GTT GAA GCT GTT GAA CGT CTG CTG CAA GAA GCG TAA
***Ec*ApeK (916364)**	GTT CTG AAC GAC CCG CGT TTC ACC GCT GAA GTT GAA CTG ACC ATC CCG TTC CAC GAC GTT GAC ATG ATG GGT GTT GCT TGG CAC GGT AAC TAC TTC CGT TAC TTC GAA GTT GCT CGT GAA GCT CTG CTG AAC CAG TTC AAC TAC GGT TAC CGT CAG ATG AAA GAA TCT GGT TAC CTG TGG CCG GTT GTT GAC GCT CGT GTT AAA TAC CGT CAC GCT CTG ACC TTC GAA CAG CGT ATC CGT GTT CGT GCT CAC ATC GAA GAA TTC GAA AAC CGT CTG CGT ATC GGT TAC CAG ATC TTC GAC GCT GAA ACC GGT AAA CGT GCT ACC ACC GGT TAC ACC ATC CAG GTT GCT GTT GAC GAA CAG TCT CGT GAA CTG TGC TTC GTT TCT CCG GAC ATC CTG TTC GAA CGT ATG GGT GTT AAA CCG TAA

Table A.2 Optimized genes used for molecular cloning of *Af*yg1p from *Aspergillus fumigatus* and *Wd*yg1p constructs from *Exophiala dermatitidis*

Construct (Gene ID)	DNA sequence
***Af*yg1p (GenBank: AF116902.1)**	CCG CGT TGG ATC TTA GGT GAT AAA TTT GAC ACC GTA TTC CCG CAC AAA GGC TCT CTG AAA GTT CTG TGG GAA TCT CGT TGG AAA TTC GCT TGC TCT AAA TCT GTT TAC CCG TTC CAC GAC GGT TCT ATC GAA GAC TTC GAA CCG ATC TTC AAC CAC CTG ATC TCT AAA AAC ATC AAC GAC GCT GCT TCT GAC GAA TAC ACC CAG GCT TTC CTG CCG ACC GCT TCT GCT CTG GAA GAA AAA GCT GCT CAG GCT CTG CAA GCT GGT AAA CAC GAA GAA GCT AGT AAC CTG CTG TGC CGT GCT GCT GTT GTT TAC CGT ATC TCT CGT TTC CCG TAC GTT GAC ATC ACC AAA CCG AAC TCT ATC AAA CGT GTT GCT TTC GAA CGT CAG AAA CAG GCT TAC CTG AAA GCT ACC TCT CTG TGG ACC CAG CCG ATC CGT GAA GTT ACC GTT CCG CAC ACC TAC CGT ACC GGT AAC GAC GGT GCT CAC ATC CCG ATC TAC ATC CGT ACC CCG GCT GGT GCT GAC CAG TCT AAC CCG GTT CCG ATC GTT CTG ATC ATG ACC GGT CTG GAC GGT TAC CGT CCG GAC AAC TCT CAG CGT ACC CAC GAA ATC CTG GCT CGT GGT TGG GCT GCT GTT GTT GCT GAA ATC CCG GGT ACC GCT GAC TGC CCG GCT GAC CCG GCT GAC CCG GCT TCT CCG GAC CGT CTG TGG GAC TCT GTT CTG TCT TAC CTG GAC CAG CGT CCG GAA CTG AAC ACC GCT AAA ATG GTT GTT TGG GGT CTG TCT GCT GGT GGT TAC TAC GCT ATC CGT GCT GCT CAC ACC CAC CGT GAC CGT CTG CTG GGT GCT ATC GCT CAC GGT CCG GGT TGC CAC TAC TAC CTG GAC CCG GAA TGG CTG GCT AAA GTT AAC GAC CAC GAA TAC CCG TTC GAA ATC ACC GCT GCT TGG GCT ACC AAA CAC GGT TAC AAA ACC GTT GAA GAA TTC GTT GCT GGT GCT CAG AAA AAA TTC TCT CTG GTT GAA ACC GGT ATC GTT GAC CAG CCG TCT TGC CGT CTG CTG CTG CTG AAC GGT GTT GAC GAC GGT GTT GTT CCG ATC GAA GAC TGC CTG GTT CTG TTC GAA CAC GGT TCT CCG AAA GAA GGT CGT TTC TAC AAA GGT CTG CCG CAC ATG GGT TAC CCG AAC TCT CTG CCG GTT TCT TAC GAA TGG CTG GAA CAG GTT CTG GCT TCT CCG TCT AAA ACC AAA AAC TAA
***Wd*yg1p (20306951)**	GCT ACC GAA AAA TAC TAC ATC CGT GAC GCT ATC ACC AAA CCG GCT GTT CAC CAC GAA TCT TAC CAG AAA CTG TGG GAA ACC AAA TGG AAA AAA CCG TGC GAA ATG GGT GTT TAC CCG TTC ATG TTC GGT TCT ATC AAA GAC TTC GAA CCG GTT GCT CAG GAA ATC ATC AAA AAA GGT CTG AAA GAA CCG TAC GAC TGG GAC GAA TAC GCT CAG ATG TAC TTC CCG AAA GCT GAA GAA CTG GCT AAA ATC GCT GAA GAA GCT GAA GCT GCT GGT GAA AAA GAA AAA GCT AGT GAA TAC TAC CTG CGT TCT TCT GCT GTT TAC CGT ATC TCT CGT TTC CCG ACC CCG CGT TCT GAA AAA CAG AAA TAC GCT TGG CGT AAA GGC TGC GAA GTC TTC TAC AAA GGG GCT GCT CTG ATG GAA TAC CCG ATC AAA GAA GTT CGT ATC CCG CAC AAA CAC GGT ATC GAA GGT GAA GGT GAC GTT GTT CCG GTT AAC TTC CTG CTG CCG CCG AAC GCT TCT GAA ACC TCT CCG GTT CCG TGC GTT CTG ATC ATC ACC GGT CTG GAC GGT TAC AGG ACC GAG CTG GCT GTG TGG CAG CAG GGG TGG CGT TCT AAA GGT GTT GCT ACC GTT ATC GCT GAA ATC CCG GGT ACC GGT GAC TCT CCG GCT CTG CGT CAG GAC CCG ACC TCT CCG GAC CGT CAG TGG TCT TCT GTT CTG GAC TGG ATC GAA TCT CAG AAA GCT GTT GAC TCT AAA AAA ATC GTT GCT TGG GGT TTC TCT ACC GGT GGT TAC TAC GCT CTG CGT ATG GCT CAC ACC CAC AAA GAC CGT CTG CTG GCT ACC ATC TCT CTG GGT GGT GGT GCT CAC CAC ATG TTC GAC CGT GAA TGG CTG GAA CAC GCT AAC AAA CTG GAA TAC CCG TTC GAC CTG TCT AAC ACC CTG GCT TAC AAA TTC GGT TAC CCG GAC CTC GAA TCG TTC ATC GAA GAA TCG TCT AAA TTC TCT CTG CTG AAC GAC GGT ACC CTG CAA AAA CCG TGC ACC AAA GTT CTG CTG GTT AAC GGT AAC GAC GAC GAA ATC TTC CCG ATC GAC GAC ATG TTC GTT TCT CTG GAA AAC GGT CAG CCG AAA CTG GCT CGT ATG GTT AAA GGT AAA AAA CAC ATG GGT GAA CCG GAA TCT TTC TCT ATC ATC CTG GAA TGG ATA CAC AAA CTG CTG GGT CTG GAC GGT AAA ATC AAA GAA CAG CTG GCT ATG ATC CCG TCT CGT ACC AAA TAA

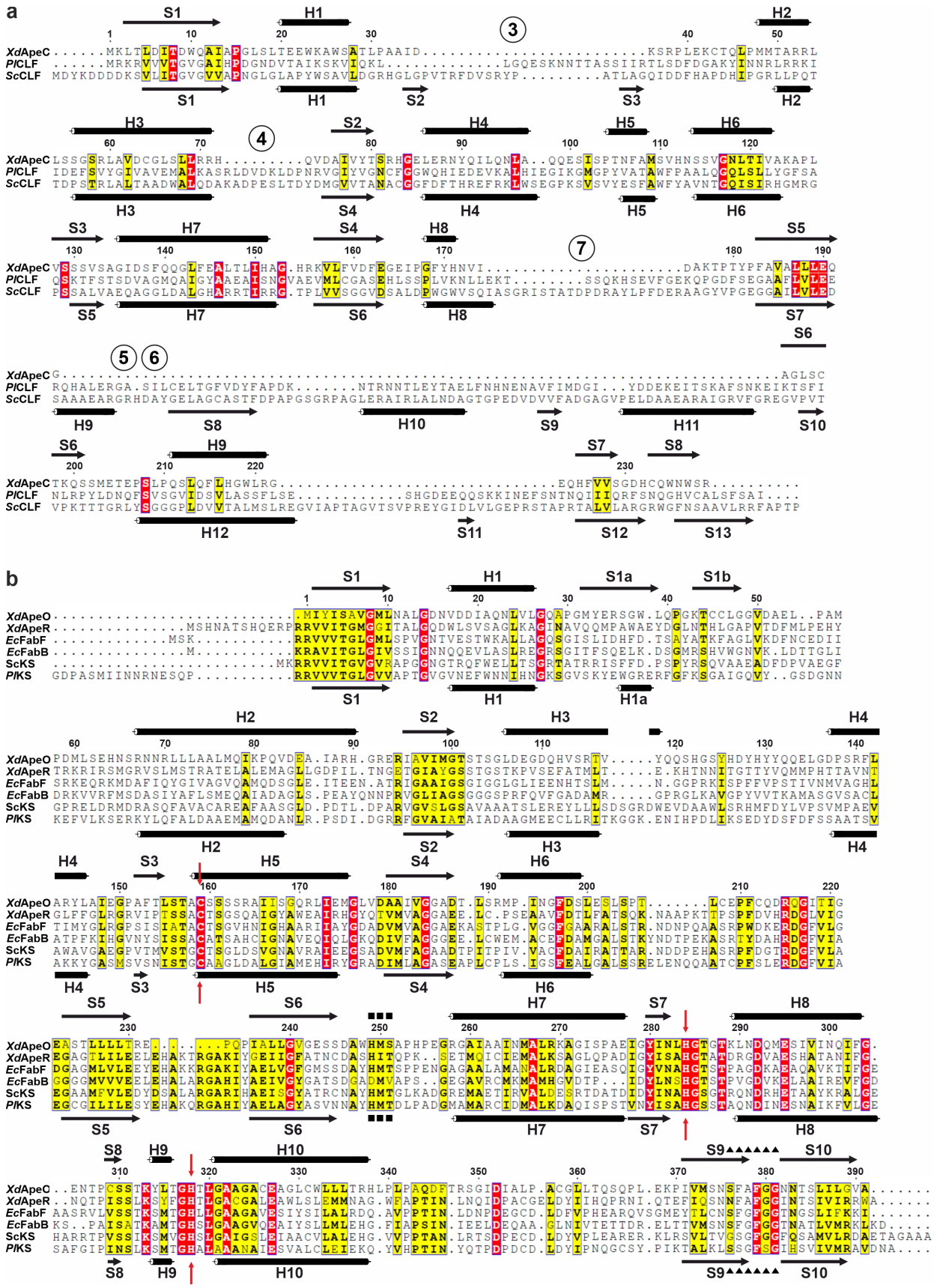

Figure A.1 Sequence alignments of *Xd*ApeO:C and *Xd*ApeR with KSs from aromatic type II PKS and type II FAS based on structural data. (a) Shown is the sequence alignment of *Xd*ApeC[35] with CLFs involved in BS of AQ-256 (*Pl*CLF)[27] and actinorhodin (*Sc*CLF).[74] **(b)** Shown is the sequence alignment of *Xd*ApeR and *Xd*ApeC[35] with KSs involved in BS of AQ-256 (*Pl*CLF),[27] actinorhodin (*Sc*CLF),[74] and FabF and FabB from *E. coli*.[72] Red arrows, black triangles, and black squares indicate active sites, GFGG-motif (loop 1), and H(M/I)T-motif (loop 2), respectively.

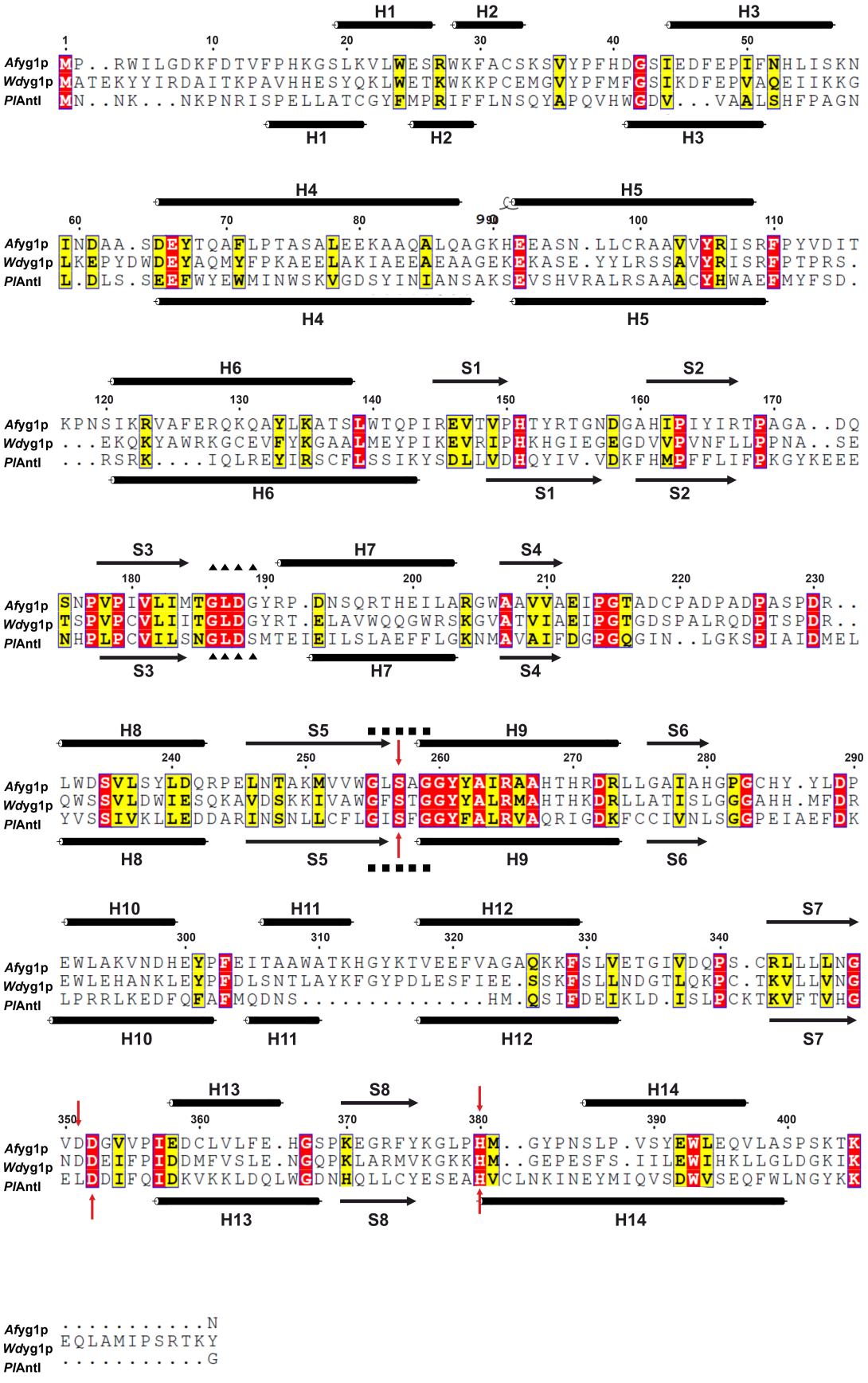

Figure A.2 Sequence alignments of *Af*yg1p and *Wd*yg1p with *Pl*Antl based on structural data.
Red arrows, black squares, and black triangles indicate the active site residues, the motif GXSXG (S is active site nucleophile), and the motif GLD(G/S) (β-turn), respectively.

Table A.3 Crystallographic data collection and refinement statistics I

	***Xd*ApeI:P**	***Xd*ApeO:C**	***Xd*ApeR**	***Ec*ApeQ**
Crystal parameters				
Space group	C2	$P2_1$	P1	$P6_5$
Cell constants	a = 144.7 Å b = 47.8 Å c = 42.7 Å	a = 57.0 Å b = 88.9 Å c = 63.0 Å	a = 61.5 Å b = 91.3 Å c = 196.3 Å α = 96.3° β = 90.2° γ = 109.6°	a = b = 125.7 Å c = 113.5 Å
Subunits / AU[a]	1	1	10	4
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-1.85 (1.95-1.85)	30-1.45 (1.55-1.45)	30-2.8 (2.9-2.8)	30-2.6 (2.7-2.6)
No. observed reflections	68,219	358,419	169,678	128,724
No. unique reflections[c]	23,313	103,673	88,990	31,143
Completeness (%)[b]	95.9 (97.6)	98.3 (96.9)	90.5 (92.7)	99.3 (99.9)
R_{merge} (%)[b,d]	6.9 (47.0)	4.3 (65.7)	12.3 (71.0)	9.3 (59.2)
I/σ (I)[b]	11.0 (2.7)	12.9 (2.0)	7.1 (2.1)	9.5 (2.4)
Refinement (REFMAC5)				
Resolution range (Å)	30-1.85	30-1.45	30-2.8	30-2.6
No. refl. working set	22,103	98,482	84,486	29,567
No. refl. test set	1,163	5,183	4,455	1,557
No. non hydrogen	2,244	5,186	27,573	7,219
No. of ligands	-	-	-	3
Solvent	106	328	549	109
R_{work} / R_{free} (%)[e]	19.8 / 23.6	12.8 / 16.2	23.3 / 26.2	18.1 / 24.3
rmsd bond (Å) / (°)[f]	0.007 / 1.29	0.006 / 1.34	0.001 / 1.1	0.003 / 1.3
Average B-factor (Å^2)	30.7	20.6	54.5	57.2
Ramachandran Plot (%)[g]	98.4 / 1.6 / 0	97.6 / 2.4 / 0	95.7 / 4.2 / 0.1	95.8 / 4.2 / 0
PDB accession code	6QSR	6QSP		

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.4 Crystallographic data collection and refinement statistics II

	***Xd*ApeK**	***Xd*ApeK:E**	***Xd*ApeK:F**
Crystal parameters			
Space group	I222	P1	C2
Cell constants	a = b = 53.2 Å c = 89.2 Å	a = 48.1 Å b = 92.0 Å c = 108.7 Å α = 85.4° β = 89.9° γ = 75.6°	a = 201.3 Å b = 60.1 Å c = 92.6 Å β = 111.7°
Subunits / AU[a]	1	14	8
Data collection			
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0
Resolution range (Å)[b]	30-2.6 (2.7-2.6)	30-2.3 (2.4-2.3)	30-2.45 (2.55-2.45)
No. observed reflections	16,022	228,574	115,016
No. unique reflections[c]	4,061	75,666	37,041
Completeness (%)[b]	98.1 (99.5)	94.7 (97.9)	96.9 (97.0)
R_{merge} (%)[b,d]	4.3 (54.7)	9.8 (68.8)	5.6 (58.1)
I/σ (I)[b]	15.3 (2.2)	7.7 (2.4)	11.4 (2.0)
Refinement (REFMAC5)			
Resolution range (Å)	30-2.6	30-2.3	30-2.45
No. refl. working set	3,856	71,838	35,171
No. refl. test set	202	3,798	1,851
No. non hydrogen	1,047	12,384	7,151
No. of ligands	-	6	4
Solvent	5	515	83
R_{work} / R_{free} (%)[e]	21.0 / 25.7	26.8 / 29.5	19.9 / 24.3
rmsd bond (Å) / (°)[f]	0.008 / 1.3	0.002 / 1.2	0.002 / 1.2
Average B-factor (Å^2)	78.8	53.8	63.8
Ramachandran Plot (%)[g]	96.9 / 3.1 / 0	96.0 / 4 / 0	96.7 / 3.3 / 0
PDB accession code			

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.5 Crystallographic data collection and refinement statistics III

	***Af*yg1p^{apo}**	***Wd*yg1p^{apo}**	***Af*yg1p:PMS**	***Af*yg1p$^{\mathrm{SeMet}}$**
Crystal parameters				
Space group	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$
Cell constants	a = 85.8 Å	a = 77.0 Å	a = 85.6 Å	a = 85.3 Å
	b = 108.5 Å	b = 86.6 Å	b = 107.6 Å	b = 92.1 Å
	c = 91.9 Å	c = 115.7 Å	c = 92.5 Å	c = 93.3 Å
	β = 90.03°		β = 90.01°	
Subunits / AU[a]	4	2	4	2
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	0.97914
Resolution range (Å)[b]	30-1.7 (1.8-1.7)	30-1.85 (1.95-1.85)	30-1.8 (1.9-1.8)	30-2.2 (2.3-2.2)
No. observed reflections	556,406	295,307	413,528	503,398
No. unique reflections[c]	179,315	65,754	148,528	71,952
Completeness (%)[b]	97.2 (97.9)	98.6 (98.4)	95.9 (97.4)	99.9 (99.9)
R_{merge} (%)[b,d]	6.5 (52.4)	6.0 (54.3)	6.6 (54.6)	11.4 (54.6)
I/σ (I)[b]	9.6 (2.4)	16.1 (3.2)	9.4 (1.9)	11.6 (4.9)
Refinement (REFMAC5)				
Resolution range (Å)	30-1.7	30-1.85	30-1.8	30-2.2
No. refl. working set	170,348	62,450	141,275	36,011
No. refl. test set	8,937	3,287	7,396	1,895
No. non hydrogen	14,233	7,183	12,868	6,229
No. of ligands	-	-	-	-
Solvent	1,743	602	434	-
R_{work} / R_{free} (%)[e]	16.5 / 19.8	15.6 / 19.7	23.5 / 24.6	21.6 / 28.3
rmsd bond (Å) / (°)[f]	0.003 / 1.1	0.003 / 1.1	0.003 / 1.1	0.007 / 1.3
Average B-factor (Å^2)	23.7	26.7	26.8	33.4
Ramachandran Plot (%)[g]	96.1 / 2.6 / 1.3	95.8 / 2.8 / 1.4	96.2 / 3.2 / 0.6	96.1 / 2.4 / 1.5
PDB accession code				

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{\mathrm{merge}}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.6 Crystallographic data collection and refinement statistics IV

	***Wd*yg1p**$^{Dha,\#}$	***Pl*Antl**^{Dha,*}	***Af*yg1p**^{old}	***Wd*yg1p**^{old}
Crystal parameters				
Space group	$P2_12_12_1$	$C2_1$	$P2_1$	$P2_12_12_1$
Cell constants	a = 85.5 Å b = 88.3 Å c = 122.0 Å	a = 54.7 Å b = 155.0 Å c = 91.4 Å	a = 85.6 Å b = 107.4 Å c = 92.5 Å β = 90.1 °	a = 85.8 Å b = 88.4 Å c = 122.4 Å
Subunits / AU[a]	2	1	4	2
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-1.85 (1.95-1.85)	30-1.7 (1.8-1.7)	30-1.8 (1.9-1.8)	30-1.75 (1.85-1.75)
No. observed reflections	416,769	199,233	437,720	421,966
No. unique reflections[c]	78,363	42,427	150,156	93,327
Completeness (%)[b]	98.6 (95.5)	98.3 (98.5)	97.2 (97.2)	98.8 (98.0)
R_{merge} (%)[b,d]	5.9 (63.1)	4.8 (62.4)	7.7 (57.1)	5.5 (64.5)
I/σ (I)[b]	16.0 (2.9)	15.6 (2.5)	9.6 (2.8)	15.5 (2.4)
Refinement (REFMAC5)				
Resolution range (Å)	30-1.85	30-1.7	30-1.8	30-1.75
No. refl. working set	74,434	40,299	142,699	88,567
No. refl. test set	3,918	2,121	7,428	4,662
No. non hydrogen	7,328	3,252	12,572	7,442
No. of ligands	-	-	-	-
Solvent	739	243	-	828
R_{work} / R_{free} (%)[e]	15.1 / 18.9	16.6 / 18.8	22.3 / 25.1	15.0 / 18.0
rmsd bond (Å) / (°)[f]	0.003 / 1.1	0.003 / 1.1	0.01 / 1.3	0.003 / 1.1
Average B-factor (Å^2)	30.1	32.4	22.5	27.8
Ramachandran Plot (%)[g]	96.1 / 3.1 / 0.8	98.1 / 1.9 / 0	95.8 / 2.9 / 1.3	96.0 / 2.7 / 1.3
PDB accession code				

Sulfonyl fluorides used were PMSF (*) and PSF (#)

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.7 Crystallographic data collection and refinement statistics V

	***Wd*yg1p**^{Dha,*}	***Pl*Antl**$^{Dha,\#}$
Crystal parameters		
Space group	$P2_12_12_1$	$C2_1$
Cell constants	a = 85.6 Å	a = 54.6 Å
	b = 88.5 Å	b = 154.6 Å
	c = 122.5 Å	c = 91.5 Å
Subunits / AU[a]	2	1
Data collection		
Beam line	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0
Resolution range (Å)[b]	30-1.85 (1.95-1.85)	30-1.8 (1.9-1.8)
No. observed reflections	301,708	168,300
No. unique reflections[c]	79,470	35,748
Completeness (%)[b]	99.3 (99.7)	98.4 (99.4)
R_{merge} (%)[b,d]	5.2 (53.8)	4.6 (64.3)
I/σ (I)[b]	15.9 (2.7)	18.2 (2.3)
Refinement (REFMAC5)		
Resolution range (Å)	30-1.85	30-1.8
No. refl. working set	75,482	33,957
No. refl. test set	3,973	1,787
No. non hydrogen	7,279	3,338
No. of ligands	-	-
Solvent	706	217
R_{work} / R_{free} (%)[e]	14.6 / 17.9	16.1 / 19.5
rmsd bond (Å) / (°)[f]	0.003 / 1.1	0.003 / 1.1
Average B-factor (Å^2)	27.9	32.1
Ramachandran Plot (%)[g]	96.2 / 2.4 / 1.4	98.4 / 1.6 / 0
PDB accession code		

Sulfonyl fluorides used were PMSF (*) and PSF (#)

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.8 Crystallographic data collection and refinement statistics VI

	***Pl*AntI:PAC**	***Pl*AntI:1-N**
Crystal parameters		
Space group	$C2_1$	$C2_1$
Cell constants	a = 55.4 Å b = 154.7 Å c = 92.1 Å	a = 55.9 Å b = 154.7 Å c = 92.1 Å
Subunits / AU[a]	1	1
Data collection		
Beam line	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0
Resolution range (Å)[b]	30-1.65 (1.75-1.65)	30-2.05 (2.15-2.05)
No. observed reflections	212,226	110,560
No. unique reflections[c]	46,968	24,440
Completeness (%)[b]	97.9 (99.4)	95.8 (96.8)
R_{merge} (%)[b,d]	3.5 (54.7)	4.5 (67.0)
I/σ (I)[b]	19.6 (2.5)	17.3 (3.8)
Refinement (REFMAC5)		
Resolution range (Å)	30-1.65	30-2.05
No. refl. working set	44,614	23,215
No. refl. test set	2,348	1,221
No. non hydrogen	3,170	3,113
No. of ligands	1	1
Solvent	172	99
R_{work} / R_{free} (%)[e]	17.2 / 20.3	18.9 / 20.6
rmsd bond (Å) / (°)[f]	0.003 / 1.1	0.002 / 1.1
Average B-factor (Å^2)	36.7	51.6
Ramachandran Plot (%)[g]	98.4 / 1.6 / 0	97.6 / 1.9 / 0.5
PDB accession code		

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.9 Crystallographic data collection and refinement statistics VII

	***Af*yg1p:1,3-DHN**	***Wd*yg1p:1,3-DHN**
Crystal parameters		
Space group	$P2_1$	$P2_12_12_1$
Cell constants	a = 85.5 Å	a = 85.8 Å
	b = 107.5 Å	b = 88.9 Å
	c = 92.5 Å	c = 123.1 Å
	β = 90.1 °	
Subunits / AU[a]	4	2
Data collection		
Beam line	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0
Resolution range (Å)[b]	30-1.75 (1.85-1.75)	30-1.65 (1.75-1.65)
No. observed reflections	496,678	493,745
No. unique reflections[c]	159,406	111,963
Completeness (%)[b]	94.8 (96.2)	98.7 (98.4)
R_{merge} (%)[b,d]	10.9 (61.4)	5.8 (65.2)
I/σ (I)[b]	6.5 (2.1)	11.8 (2.1)
Refinement (REFMAC5)		
Resolution range (Å)	30-1.75	30-1.65
No. refl. working set	151,446	106,353
No. refl. test set	7,932	5,598
No. non hydrogen	13,725	7,439
No. of ligands	4	2
Solvent	1,198	757
R_{work} / R_{free} (%)[e]	22.9 / 23.8	16.3 / 18.8
rmsd bond (Å) / (°)[f]	0.003 / 1.2	0.003 / 1.1
Average B-factor (Å^2)	20.8	29.5
Ramachandran Plot (%)[g]	96.2 / 2.5 / 1.3	95.8 / 2.9 / 1.3
PDB accession code		

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

Table A.10 Crystallographic data collection and refinement statistics VIII

	***Af*yg1p^{ser257ala}**	***Af*yg1p:THN**	***Wd*yg1p:THN**	***Pl*Antl:CA**
Crystal parameters				
Space group	$P2_1$	$P2_1$	$P2_12_12_1$	$C2_1$
Cell constants	a = 85.6 Å	a = 85.5 Å	a = 86.2 Å	a = 54.4 Å
	b = 108.0 Å	b = 107.4 Å	b = 88.5 Å	b = 154.5 Å
	c = 92.5 Å	c = 92.7 Å	c = 122.4 Å	c = 91.1 Å
	β = 89.9°	β = 90.2°		
Subunits / AU[a]	4	4	2	1
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-1.45 (1.55-1.45)	30-1.7 (1.8-1.7)	30-1.95 (2.05-1.95)	30-1.4 (1.5-1.4)
No. observed reflections	897,522	556,281	306,939	335,698
No. unique reflections[c]	291,071	173,708	67,832	74,543
Completeness (%)[b]	98.1 (99.4)	94.6 (95.7)	98.5 (96.4)	98.4 (98.6)
R_{merge} (%)[b,d]	6.7 (47.8)	9.4 (60.0)	6.2 (63.8)	3.0 (66.2)
I/σ (I)[b]	8.9 (2.1)	6.7 (1.9)	14.1 (2.2)	20.2 (2.0)
Refinement (REFMAC5)				
Resolution range (Å)	30-1.45	30-1.7	30-1.95	30-1.4
No. refl. working set	276,418	164,908	64,426	70,809
No. refl. test set	14,624	8,775	3,391	3,727
No. non hydrogen	14,787	13,719	7,069	3,387
No. of ligands	-	12	2	1
Solvent	2,231	1,028	511	295
R_{work} / R_{free} (%)[e]	16.2 / 18.3	19.6 / 22.1	15.8 / 20.3	14.7 / 18.1
rmsd bond (Å) / (°)[f]	0.002 / 1.1	0.002 / 1.2	0.003 / 1.2	0.006 / 1.4
Average B-factor (Å^2)	18.7	20.7	36.1	25.7
Ramachandran Plot (%)[g]	96.5 / 2.4 / 1.1	95.9 / 2.7 / 1.4	95.8 / 3.0 / 1.2	98.1 / 1.9 / 0
PDB accession code				

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge}, and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

[d] $R_{merge}(I) = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the jth measurement of the intensity of reflection hkl and $\langle I_{hkl}\rangle$ is the average intensity

[e] $R = \frac{\sum_{hkl} ||F^{obs}_{hkl}| - |F^{calc}_{hkl}||}{\sum_{hkl} |F^{obs}_{hkl}|}$, where R_{free} is calculated without a σ-cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths / angles

[g] Number of residues in favored region / allowed region / outlier region

B Bibliography

[1] J. M. Berg, J. L. Tymoczko, G. J. Gatto, and L. Stryer. In *Stryer Biochemie*, pages 499–525. Springer Berlin Heidelberg, 2018.

[2] C.-I. Lin, R. M. McCarty, and H.-W. Liu. The enzymology of organic transformations: A survey of name reactions in biological systems. *Angewandte Chemie International Edition*, 56(13):3446–3489, 2017.

[3] A. Kossel. Ueber die chemische zusammensetzung der zelle. *Archiv für Physiologie*, 181:181–186, 1891.

[4] O. Tyc, C. Song, J. S. Dickschat, M. Vos, and P. Garbeva. The ecological role of volatile and soluble secondary metabolites produced by soil bacteria. *Trends in Microbiology*, 25(4):280–292, 2017.

[5] P. B. Schiff, J. Fant, and S. B. Horwitz. Promotion of microtubule assembly *in vitro* by taxol. *Nature*, 277(5698):665–667, 1979.

[6] M. C. Wani, H. L. Taylor, M. E. Wall, P. Coggon, and A. T. McPhail. Plant antitumor agents. VI. isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *Journal of the American Chemical Society*, 93(9):2325–2327, 1971.

[7] F. Yang, S. Hanon, P. Lam, and P. Schweitzer. Quinidine revisited. *The American Journal of Medicine*, 122(4):317–321, 2009.

[8] A. Fleming. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzæ*. *Br J Exp Pathol.*, 10(3):226–236, 1929.

[9] R. Y. Calne. Immunosuppression for organ grafting observations on cyclosporin A. *Immunological Reviews*, 46(1):113–124, 1979.

[10] J. F. Borel, C. Feurer, H. U. Gubler, and H. Stähelin. Biological effects of cyclosporin A: A new antilymphocytic agent. *Agents and Actions*, 6(4):468–475, 1976.

[11] H. A. Kirst, K. H. Michel, J. S. Mynderase, E. H. Chio, R. C. Yao, W. M. Nakasukasa, L. D. Boeck, J. L. Occlowitz, J. W. Paschal, J. B. Deeter, and G. D. Thompson. *Discovery, Isolation, and Structure Elucidation of a Family of Structurally Unique, Fermentation-Derived Tetracyclic Macrolides*, chapter 20, pages 214–225. 1992.

[12] G. D. Thompson, R. Dutton, and T. C. Sparks. Spinosad – a case study: an example from a natural products discovery programme. *Pest Management Science*, 56(8):696–702, 2000.

[13] B. Salehi, P. V. T. Fokou, M. Sharifi-Rad, P. Zucca, R. Pezzani, N. Martins, and J. Sharifi-Rad. The therapeutic potential of naringenin: A review of clinical trials. *Pharmaceuticals*, 12(1), 2019.

[14] L. F. Wright and D. A. Hopwood. Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor*A3(2). *Microbiology*, 96(2):289–297, 1976.

[15] E. L. Hazen and R. Brown. Fungicidin, an antibiotic produced by a soil *Actinomycete*. *Proceedings of the Society for Experimental Biology and Medicine*, 76(1):93–97, 1951.

[16] H. Umezawa, K. Maeda, T. Takeuchi, and Y. Okami. New antibiotics, bleomycin A and B. *The Journal of antibiotics*, 19 5:200–9, 1966.

[17] J. Staunton and K. J. Weissman. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.*, 18:380–416, 2001.

[18] A. R. Fernie, F. Carrari, and L. J. Sweetlove. Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Current Opinion in Plant Biology*, 7(3):254–261, 2004.

[19] R. C. Hider and X. Kong. Chemistry and biology of siderophores. *Nat. Prod. Rep.*, 27:637–657, 2010.

[20] D. J. Newman and G. M. Cragg. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of Natural Products*, 83(3):770–803, 2020.

[21] A. G. Atanasov, S. B. Zotchev, V. M. Dirsch, I. E. Orhan, M. Banach, J. M. Rollinger, D. Barreca, W. Weckwerth, R. Bauer, E. A. Bayer, M. Majeed, A. Bishayee, V. Bochkov, G. K. Bonn, N. Braidy, F. Bucar, A. Cifuentes, G. D'Onofrio, M. Bodkin, M. Diederich, A. T. Dinkova-Kostova, T. Efferth, K. El Bairi, N. Arkells, T.-P. Fan, B. L. Fiebich, M. Freissmuth, M. I. Georgiev, S. Gibbons, K. M. Godfrey, C. W. Gruber, J. Heer, L. A. Huber, E. Ibanez, A. Kijjoa, A. K. Kiss, A. Lu, F. A. Macias, M. J. S. Miller, A. Mocan, R. Müller, F. Nicoletti, G. Perry, V. Pittalà, L. Rastrelli, M. Ristow, G. L. Russo, A. S. Silva, D. Schuster, H. Sheridan, K. Skalicka-Woźniak, L. Skaltsounis, E. Sobarzo-Sánchez, D. S. Bredt, H. Stuppner, A. Sureda, N. T. Tzvetkov, R. A. Vacca, B. B. Aggarwal, M. Battino, F. Giampieri, M. Wink, J.-L. Wolfender, J. Xiao, A. W. K. Yeung, G. Lizard, M. A. Popp, M. Heinrich, I. Berindan-Neagoe, M. Stadler, M. Daglia, R. Verpoorte, C. T. Supuran, and the International Natural Product Sciences Taskforce. Natural products in drug discovery: advances and opportunities. *Nature Reviews Drug Discovery*, 20(3):200–216, 2021.

[22] A. Chen, R. N. Re, and M. D. Burkart. Type II fatty acid and polyketide synthases: deciphering protein–protein and protein–substrate interactions. *Natural Product Reports*, 35(10):1029–1045, 2018.

[23] C. Hertweck. The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition*, 48(26):4688–4716, 2009.

[24] A. Craney, S. Ahmed, and J. Nodwell. Towards a new science of secondary metabolism. *The Journal of Antibiotics*, 66(7):387–400, 2013.

[25] S. Okamoto, T. Taguchi, K. Ochi, and K. Ichinose. Biosynthesis of actinorhodin and related antibiotics: Discovery of alternative routes for quinone formation encoded in the act gene cluster. *Chemistry Biology*, 16(2):226–236, 2009.

[26] Q. Zhou, A. Bräuer, H. Adihou, M. Schmalhofer, P. Saura, G. L. C. Grammbitter, V. R. I. Kaila, M. Groll, and H. B. Bode. Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photorhabdus luminescens*. *Chem. Sci.*, 10:6341–6349, 2019.

[27] A. Bräuer, Q. Zhou, G. L. C. Grammbitter, M. Schmalhofer, M. Rühl, V. R. I. Kaila, H. B. Bode, and M. Groll. Structural snapshots of the minimal PKS system responsible for octaketide biosynthesis. *Nature Chemistry*, 12(8):755–763, 2020.

[28] F. Malpartida and D. A. Hopwood. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature*, 309(5967):462–464, 1984.

[29] M. Cummings, R. Breitling, and E. Takano. Steps towards the synthetic biology of polyketide biosynthesis. *FEMS Microbiology Letters*, 351(2):116–125, 2014.

[30] T. Itoh, T. Taguchi, M. R. Kimberley, K. I. Booker-Milburn, G. R. Stephenson, Y. Ebizuka, and K. Ichinose. Actinorhodin biosynthesis: Structural requirements for Post-PKS tailoring intermediates revealed by functional analysis of ActVI-ORF1 reductase. *Biochemistry*, 46(27):8181–8188, 2007.

[31] M. Hashimoto, S. Watari, T. Taguchi, K. Ishikawa, T. Kumamoto, S. Okamoto, and K. Ichinose. Actinorhodin biosynthesis terminates with an unprecedented biaryl coupling reaction. *Angewandte Chemie International Edition*, 62(5):e202214400, 2023.

[32] D. Du, Y. Katsuyama, H. Onaka, M. Fujie, N. Satoh, K. Shin-Ya, and Y. Ohnishi. Production of a novel amide-containing polyene by activating a cryptic biosynthetic gene cluster in Streptomyces sp. MSC090213JE08. *ChemBioChem*, 17(15):1464–1471, 2016.

[33] D. Du, Y. Katsuyama, K. Shin-Ya, and Y. Ohnishi. Reconstitution of a type II polyketide synthase that catalyzes polyene formation. *Angewandte Chemie International Edition*, 57(7):1954–1957, 2018.

[34] P. Cimermancic, M. H. Medema, J. Claesen, K. Kurita, L. C. Wieland Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy, B. W. Birren, E. Takano, A. Sali, R. G. Linington, and M. A. Fischbach. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell*, 158(2):412–421, 2014.

[35] G. L. C. Grammbitter, M. Schmalhofer, K. Karimi, Y.-M. Shi, T. A. Schöner, N. J. Tobias, N. Morgner, M. Groll, and H. B. Bode. An uncommon type II PKS catalyzes biosynthesis

of aryl polyene pigments. *Journal of the American Chemical Society*, 141(42):16615–16623, 2019.

[36] S. Pohle, C. Appelt, M. Roux, H.-P. Fiedler, and R. D. Süssmuth. Biosynthetic gene cluster of the non-ribosomally synthesized cyclodepsipeptide skyllamycin: Deciphering unprecedented ways of unusual hydroxylation reactions. *Journal of the American Chemical Society*, 133(16):6194–6205, 2011.

[37] Z. Deng, J. Liu, T. Li, H. Li, Z. Liu, Y. Dong, and W. Li. An unusual type II polyketide synthase system involved in cinnamoyl lipid biosynthesis. *Angewandte Chemie International Edition*, 60(1):153–158, 2021.

[38] G. L. C. Grammbitter, Y.-M. Shi, Y.-N. Shi, S. P. B. Vemulapalli, C. Richter, H. Schwalbe, M. Alanjary, A. Schüffler, M. Witt, C. Griesinger, and H. B. Bode. The chemical structure of widespread microbial aryl polyene lipids. *bioRxiv*, 2020.

[39] A. G. Andrewes, C. L. Jenkins, M. P Starr, J. Shepherd, and H. Hope. Structure of xanthomonadin I, a novel dibrominated aryl-polyene pigment produced by the bacterium *Xanthomonas juglandis*. *Tetrahedron Letters*, 17(45):4023–4024, 1976.

[40] A. G. Andrewes, S. Hertzberg, S. Liaaen-Jensen, and M. P Starr. *Xanthomonas* pigments. 2. the *Xanthomonas* "carotenoids" — non-carotenoid brominated aryl-polyene esters. *Acta Chemica Scandinavica*, 27(7):2383–2395, 1973.

[41] H. Achenbach, W. Kohl, H. Reichenbach, and H. Kleinig. Zur struktur des flexirubins. *Tetrahedron Letters*, 15(30):2555–2556, 1974.

[42] T. A. Schöner, S. Gassel, A. Osawa, N. J. Tobias, Y. Okuno, Y. Sakakibara, K. Shindo, G. Sandmann, and H. B. Bode. Aryl polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. *ChemBioChem*, 17(3):247–253, 2016.

[43] L. Rajagopal, C. S. Sundari, D. Balasubramanian, and R. V. Sonti. The bacterial pigment xanthomonadin offers protection against photodamage. *FEBS Letters*, 415(2):125–128, 1997.

[44] I. Johnston, L. J. Osborn, R. L. Markley, E. A. McManus, A. Kadam, K. B. Schultz, N. Nagajothi, P. P. Ahern, J. M. Brown, and J. Claesen. Identification of essential genes for *Escherichia coli* aryl polyene biosynthesis and function in biofilm formation. *NPJ biofilms and microbiomes*, 7(1):56, 2021.

[45] K. Langfelder, M. Streibel, B. Jahn, G. Haase, and A. A. Brakhage. Biosynthesis of fungal melanins and their importance for human pathogenic fungi. *Fungal Genetics and Biology*, 38(2):143–158, 2003.

[46] J.-P. Latgé and G. Chamilos. *Aspergillus fumigatus* and aspergillosis in 2019. *Clinical Microbiology Reviews*, 33(1):e00140–18, 2019.

[47] A. A. Bell, R. D. Stipanovic, and J. E. Puhalla. Pentaketide metabolites of *verticillium dahliae*: Identification of (+)-scytalone as a natural precursor to melanin. *Tetrahedron*, 32(12):1353–1356, 1976.

[48] A. A. Bell, J. E. Puhalla, W. J. Tolmsoff, and R. D. Stipanovic. Use of mutants to establish (+)-scytalone as an intermediate in melanin biosynthesis by *Verticillium dahliae*. *Canadian Journal of Microbiology*, 22(6):787–799, 1976.

[49] M. H. T. Stappers, A. E. Clark, V. Aimanianda, S. Bidula, D. M. Reid, P. Asamaphan, S. E. Hardison, I. M. Dambuza, I. Valsecchi, B. Kerscher, A. Plato, C. A. Wallace, R. Yuecel, B. Hebecker, M. da Glória Teixeira Sousa, C. Cunha, Y. Liu, T. Feizi, A. A. Brakhage, K. J. Kwon-Chung, N. A. R. Gow, M. Zanda, M. Piras, C. Zanato, M. Jaeger, M. G. Netea, F. L. van de Veerdonk, J. F. Lacerda, A. Campos, A. Carvalho, J. A. Willment, J.-P. Latgé, and G. D. Brown. Recognition of DHN-melanin by a C-type lectin receptor is required for immunity to *Aspergillus*. *Nature*, 555(7696):382–386, 2018.

[50] J. Bayry, A. Beaussart, Y. F. Dufrêne, M. Sharma, K. Bansal, O. Kniemeyer, V. Aimanianda, A. A. Brakhage, S. V. Kaveri, K. J. Kwon-Chung, J.-P. Latgé, and A. Beauvais. Surface structure characterization of *Aspergillus fumigatus* conidia mutated in the melanin synthesis pathway and their human cellular immune response. *Infection and Immunity*, 82(8):3141–3153, 2014.

[51] F. Mech, A. Thywißen, R. Guthke, A. A. Brakhage, and M. T. Figge. Automated image analysis of the host-pathogen interaction between phagocytes and *Aspergillus fumigatus*. *PLOS ONE*, 6(5):1–10, 2011.

[52] A. Thywißen, T. Heinekamp, H.-M. Dahse, J. Schmaler-Ripcke, S. Nietsche, P. Zipfel, and A. Brakhage. Conidial dihydroxynaphthalene melanin of the human pathogenic fungus *Aspergillus fumigatus* interferes with the host endocytosis pathway. *Frontiers in Microbiology*, 2, 2011.

[53] I. Fujii, Y. Mori, A. Watanabe, Y. Kubo, G. Tsuji, and Y. Ebizuka. Heterologous Expression and Product Identification of *Colletotrichum lagenarium* Polyketide Synthase Encoded by the PKS1 Gene Involved in Melanin Biosynthesis. *Bioscience, Biotechnology, and Biochemistry*, 63(8):1445–1452, 1999.

[54] I. Fujii, Y. Mori, A. Watanabe, Y. Kubo, G. Tsuji, and Y. Ebizuka. Enzymatic synthesis of 1,3,6,8-tetrahydroxynaphthalene solely from malonyl coenzyme a by a fungal iterative type I polyketide synthase PKS1. *Biochemistry*, 39(30):8853–8858, 2000.

[55] A. L. Vagstad, E. A. Hill, J. W. Labonte, and C. A. Townsend. Characterization of a fungal thioesterase having claisen cyclase and deacetylase activities in melanin biosynthesis. *Chemistry Biology*, 19(12):1525–1534, 2012.

[56] A. Watanabe, I. Fujii, H.-F. Tsai, Y. C. Chang, K. J. Kwon-Chung, and Y. Ebizuka. *Aspergillus fumigatus* alb1 encodes naphthopyrone synthase when expressed in *Aspergillus oryzae*. *FEMS Microbiology Letters*, 192(1):39–44, 2000.

[57] H.-F. Tsai, I. Fujii, A. Watanabe, M. H. Wheeler, Y. C. Chang, Y. Yasuoka, Y. Ebizuka, and K. J. Kwon-Chung. Pentaketide melanin biosynthesis in *Aspergillus fumigatus* requires chain-length shortening of a heptaketide precursor. *Journal of Biological Chemistry*, 276(31):29292–29298, 2001.

[58] Isao Fujii, Y. Yasuoka, H.-F. Tsai, Y. C. Chang, K. J. Kwon-Chung, and Y. Ebizuka. Hydrolytic polyketide shortening by ayg1p, a novel enzyme involved in fungal melanin biosynthesis. *Journal of Biological Chemistry*, 279(43):44613–44620, 2004.

[59] H.-F. Tsai, M. H. Wheeler, Y. C. Chang, and K. J. Kwon-Chung. A developmentally regulated gene cluster involved in conidial pigment biosynthesis in *Aspergillus fumigatus*. *Journal of Bacteriology*, 181(20):6469–6477, 1999.

[60] A. Watanabe, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake, and Y. Ebizuka. Re-identification of *Aspergillus nidulans wA* gene to code for a polyketide synthase of naphthopyrone. *Tetrahedron Letters*, 40(1):91–94, 1999.

[61] Y.-M. Chiang, K. M. Meyer, M. Praseuth, S. E. Baker, K. S. Bruno, and C. C.C. Wang. Characterization of a polyketide synthase in *Aspergillus niger* whose product is a precursor for both dihydroxynaphthalene (DHN) melanin and naphtho-γ-pyrone. *Fungal Genetics and Biology*, 48(4):430–437, 2011.

[62] T. R. Jørgensen, J. Park, M. Arentshorst, A. M. van Welzen, G. Lamers, P. A. vanKuyk, R. A. Damveld, C. A. M. van den Hondel, K. F. Nielsen, J. C. Frisvad, and A. F. J. Ram. The molecular and genetic basis of conidial pigmentation in *Aspergillus niger*. *Fungal Genetics and Biology*, 48(5):544–553, 2011.

[63] R. J.N. Frandsen, C. Schütt, B. W. Lund, D. Staerk, J. Nielsen, S. Olsson, and H. Giese. Two novel classes of enzymes are required for the biosynthesis of aurofusarin in *Fusarium graminearum*. *Journal of Biological Chemistry*, 286(12):10419–10428, 2011.

[64] P. A. Geis, M. H. Wheeler, and P. J. Szaniszlo. Pentaketide metabolites of melanin synthesis in the dematiaceous fungus *Wangiella dermatitidis*. *Archives of Microbiology*, 137(4):324–328, 1984.

[65] M. H. Wheeler, D. Abramczyk, L. S. Puckhaber, M. Naruse, Y. Ebizuka, I. Fujii, and P. J. Szaniszlo. New biosynthetic step in the melanin pathway of *Wangiella* (*Exophiala*) *dermatitidis*: Evidence for 2-acetyl-1,3,6,8-tetrahydroxynaphthalene as a novel precursor. *Eukaryotic Cell*, 7(10):1699–1711, 2008.

[66] T. P. Korman, J. M. Crawford, J. W. Labonte, A. G. Newman, J. Wong, C. A. Townsend, and S.-C. Tsai. Structure and function of an iterative polyketide synthase thioesterase domain catalyzing claisen cyclization in aflatoxin biosynthesis. *Proceedings of the National Academy of Sciences*, 107(14):6246–6251, 2010.

[67] L. Baugh, L. A. Gallagher, R. Patrapuvich, M. C. Clifton, A. S. Gardberg, T. E. Edwards, B. Armour, D. W. Begley, S. H. Dieterich, D. M. Dranow, J. Abendroth, J. W. Fairman,

D. Fox, III, B. L. Staker, I. Phan, A. Gillespie, R. Choi, S. Nakazawa-Hewitt, M. T. Nguyen, A. Napuli, L. Barrett, G. W. Buchko, R. Stacy, P. J. Myler, L. J. Stewart, C. Manoil, and W. C. Van Voorhis. Combining functional and structural genomics to sample the essential *Burkholderia* structome. *PLOS ONE*, 8(1):1–10, 2013.

[68] C. E. Christensen, B. B. Kragelund, P. von Wettstein-Knowles, and A. Henriksen. Structure of the human β-ketoacyl (ACP) synthase from the mitochondrial type II fatty acid synthase. *Protein Science*, 16(2):261–272, 2007.

[69] M. Mathieu, J. P Zeelen, R. A. Pauptit, R. Erdmann, W.-H. Kunau, and R. K. Wierenga. The 2.8 Å crystal structure of peroxisomal 3-ketoacyl-coa thiolase of *Saccharomyces cerevisiae*: a five-layered α-β-α-β-α structure constructed from two core domains of identical topology. *Structure*, 2(9):797–808, 1994.

[70] A. Vangone, R. Spinelli, V. Scarano, L. Cavallo, and R. Oliva. COCOMAPS: a web application to analyze and visualize contacts at the interface of biomolecular complexes. *Bioinformatics*, 27(20):2915–2916, 2011.

[71] E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, 372(3):774–797, 2007.

[72] J. T. Mindrebo, A. Patel, W. E. Kim, T. D. Davis, A. Chen, T. G. Bartholow, J. J. La Clair, J. A. McCammon, J. P. Noel, and M. D. Burkart. Gating mechanism of elongating β-ketoacyl-ACP synthases. *Nature Communications*, 11(1):1727, 2020.

[73] W. C. Lee, S. Choi, A. Jang, J. Yeon, E. Hwang, and Y. Kim. Structural basis of the complementary activity of two ketosynthases in aryl polyene biosynthesis. *Scientific Reports*, 11(1):16340, 2021.

[74] A. T. Keatinge-Clay, D. A. Maltby, K. F. Medzihradszky, C. Khosla, and R. M. Stroud. An antibiotic factory caught in action. *Nature Structural & Molecular Biology*, 11(9):888–893, 2004.

[75] J. T. Mindrebo, A. Chen, W. E. Kim, R. N. Re, T. D. Davis, J. P. Noel, and M. D. Burkart. Structure and mechanistic analyses of the gating mechanism of elongating ketosynthases. *ACS Catalysis*, 11(12):6787–6799, 2021.

[76] M. G. Rossmann, D. Moras, and K. W. Olsen. Chemical and biological evolution of a nucleotide-binding protein. *Nature*, 250(5463):194–199, 1974.

[77] H. Jörnvall, B. Persson, M. Krook, S. Atrian, R. Gonzalez-Duarte, J. Jeffery, and D. Ghosh. Short-chain dehydrogenases/reductases (SDR). *Biochemistry*, 34(18):6003–6013, 1995.

[78] A. C. Price, Y.-M. Zhang, C. O. Rock, and S. W. White. Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG. *Structure*, 12(3):417–428, 2004.

[79] T. P. Korman, J. A. Hill, T. N. Vu, and S.-C. Tsai. Structural analysis of actinorhodin polyketide ketoreductase: cofactor binding and substrate specificity. *Biochemistry*, 43(46):14529–14538, 2004.

[80] W. C. Lee, S. Choi, A. Jang, K. Son, and Y. Kim. Structural comparison of *Acinetobacter baumannii* β-ketoacyl-acyl carrier protein reductases in fatty acid and aryl polyene biosynthesis. *Scientific Reports*, 11(1):7945, 2021.

[81] J. Hou, T. Osinski, H. Zheng, I. Shumilin, I. Shabalin, S. Shatsman, W. F. Anderson, and W.; Center for Structural Genomics of Infectious Diseases (CSGID) Minor. Crystal structure of a putative 3-oxoacyl-[acyl-carrier protein]reductase from *Escherichia coli* strain CFT073 complexed with $NADP^+$ at 2.1 Å resolution. *PDB DOI: 10.2210/pdb4IIU/pdb*, 2013.

[82] K. Blin, S. Shaw, A. M. Kloosterman, Z. Charlop-Powers, G. P. van Wezel, M. H Medema, and T. Weber. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Research*, 49(W1):W29–W35, 2021.

[83] M. Leesong, B. S. Henderson, J. R. Gillig, J. M. Schwab, and J. L. Smith. Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure*, 4(3):253–264, 1996.

[84] G. J. Dodge, A. Patel, K. L. Jaremko, J. A. McCammon, J. L. Smith, and M. D. Burkart. Structural and dynamical rationale for fatty acid unsaturation in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 116(14):6775–6783, 2019.

[85] D. M. Krüger and H. Gohlke. DrugScorePPI webserver: fast and accurate *in silico* alanine scanning for scoring protein–protein interactions. *Nucleic Acids Research*, 38(2):W480–W486, 2010.

[86] L. Zhang, J. Xiao, J. Xu, T. Fu, Z. Cao, L. Zhu, H.-Z. Chen, X. Shen, H. Jiang, and L. Zhang. Crystal structure of FabZ-ACP complex reveals a dynamic seesaw-like catalytic mechanism of dehydratase in fatty acid biosynthesis. *Cell Research*, 26(12):1330–1344, 2016.

[87] C. Nguyen, R. W. Haushalter, D. J. Lee, P. R. L. Markwick, J. Bruegger, G. Caldara-Festin, K. Finzel, D. R. Jackson, F. Ishikawa, B. O'Dowd, J. A. McCammon, S. J. Opella, S.-C. Tsai, and M. D. Burkart. Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature*, 505(7483):427–431, 2014.

[88] M. S. Kimber, F. Martin, Y. Lu, S. Houston, M. Vedadi, A. Dharamsi, K. M. Fiebig, M. Schmid, and C. O. Rock. The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from *Pseudomonas aeruginosa*. *Journal of Biological Chemistry*, 279(50):52593–52602, 2004.

[89] R. R. Annand, J. F. Kozlowski, V. J. Davisson, and J. M. Schwab. Mechanism-based inactivation of *Escherichia coli* β-hydroxydecanoyl thiol ester dehydrase: assignment of

the imidazole ^{15}N NMR resonances and determination of the structure of the alkylated histidine. *Journal of the American Chemical Society*, 115(3):1088–1094, 1993.

[90] L. Holm. Dali server: structural unification of protein families. *Nucleic Acids Research*, 50(W1):W210–W215, 2022.

[91] L. Zhang, W. Liu, T. Hu, L. Du, C. Luo, K. Chen, X. Shen, and H. Jiang. Structural basis for catalytic and inhibitory mechanisms of β-hydroxyacyl-acyl carrier protein dehydratase (FabZ). *Journal of Biological Chemistry*, 283(9):5370–5379, 2008.

[92] D. Kostrewa, F. K. Winkler, G. Folkers, L. Scapozza, and R. Perozzo. The crystal structure of PfFabZ, the unique β-hydroxyacyl-ACP dehydratase involved in fatty acid biosynthesis of *Plasmodium falciparum*. *Protein Science*, 14(6):1570–1580, 2005.

[93] Lew S. Wang H. Janjua H. Foote E. L. Xiao R. Nair R. Everett J. K. Acton T.B. Rost B. Montelione G. T. Hunt J. F. Tong L. Northeast Structural Genomics Consortium (NESG). Seetharaman, J. Crystal structure of an uncharacterized protein from *Erwinia carotovora* subsp. atroseptica. NESG target EwR179. *PDB DOI: 10.2210/pdb3ESI/pdb*, 2008.

[94] T. G. Bartholow, T. Sztain, A. Patel, D. J. Lee, M. A. Young, R. Abagyan, and M. D. Burkart. Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis. *Communications Biology*, 4(1):340, 2021.

[95] A. M. Murad, H. Brognaro, S. Falke, J. Lindner, M. Perbandt, C. Mudogo, R. Schubert, C. Wrenger, and C. Betzel. Structure and activity of the DHNA coenzyme-a thioesterase from *Staphylococcus aureus* providing insights for innovative drug development. *Scientific Reports*, 12(1):4313, 2022.

[96] Kunishima N. Mizutani, H. Crystal structure of the acyl carrier protein from *Thermus thermophilus* HB8. *PDB DOI: 10.2210/pdb1X3O/pdb*, 2005.

[97] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger. Colabfold: making protein folding accessible to all. *Nature Methods*, 19(6):679–682, 2022.

[98] Y. Kim and J. H. Prestegard. Refinement of the NMR structures for acyl carrier protein with scalar coupling data. *Proteins: Structure, Function, and Bioinformatics*, 8(4):377–385, 1990.

[99] Y.-M. Zhang, M. S. Rao, R. J. Heath, A. C. Price, A. J. Olson, C. O. Rock, and S. W. White. Identification and analysis of the acyl carrier protein (ACP) docking site on β-ketoacyl-ACP synthase III. *Journal of Biological Chemistry*, 276(11):8231–8238, 2001.

[100] Y.-M. Zhang, B. Wu, J. Zheng, and C. O. Rock. Key residues responsible for acyl carrier protein and β-ketoacyl-acyl carrier protein reductase (FabG) interaction. *Journal of Biological Chemistry*, 278(52):52935–52943, 2003.

[101] Chong Wai Liew, Andrew Sharff, Masayo Kotaka, Rong Kong, Huihua Sun, Insaf Qureshi, Gérard Bricogne, Zhao-Xun Liang, and Julien Lescar. Induced-fit upon ligand binding revealed by crystal structures of the hot-dog fold thioesterase in dynemicin biosynthesis. *Journal of Molecular Biology*, 404(2):291–306, 2010.

[102] C. Schleberger, P. Sachelaru, R. Brandsch, and G. E. Schulz. Structure and action of a CC bond cleaving α/β-hydrolase involved in nicotine degration. *Journal of Molecular Biology*, 367(2):409–418, 2007.

[103] D. L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman. The α-/β hydrolase fold. *Protein Engineering, Design and Selection*, 5(3):197–211, 1992.

[104] D. M. Blow, J. J. Birktoft, and B. S. Hartley. Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature*, 221(5178):337–340, 1969.

[105] A. A. Kossiakoff and S. A. Spencer. Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: neutron structure of trypsin. *Biochemistry*, 20(22):6462–6474, 1981.

[106] D. E. Fahrney and A. M. Gold. Sulfonyl fluorides as inhibitors of esterases. I. rates of reaction with acetylcholinesterase, α-chymotrypsin, and trypsin. *Journal of the American Chemical Society*, 85(7):997–1000, 1963.

[107] A. M. Gold and D. E. Fahrney. Sulfonyl fluorides as inhibitors of esterases. II. formation and reactions of phenylmethanesulfonyl α-chymotrypsin. *Biochemistry*, 3(6):783–791, 1964.

[108] G. Schoellmann and E. Shaw. Direct evidence for the presence of histidine in the active center of chymotrypsin. *Biochemistry*, 2(2):252–255, 1963.

[109] C. L. Brooks, D. A. Ostrov, N. C. Schumann, S. Kakkad, D. Li, K. Peña, B. P. Williams, and N. E. Goldfarb. 2.1 Å crystal structure of the *Mycobacterium tuberculosis* serine hydrolase, Hip1, in its anhydro-form (Anhydrohip1). *Biochemical and Biophysical Research Communications*, 630:57–63, 2022.

[110] M. Gersch, R. Kolb, F. Alte, M. Groll, and S. A. Sieber. Disruption of oligomerization and dehydroalanine formation as mechanisms for ClpP protease inhibition. *Journal of the American Chemical Society*, 136(4):1360–1366, 2014.

[111] M. P. Malakhov, M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks, and T. R. Butt. Sumo fusions and sumo-specific protease for efficient expression and purification of proteins. *Journal of Structural and Functional Genomics*, 5(1):75–86, 2004.

[112] R. B. Kapust, J. Tözsér, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland, and D. S. Waugh. Tobacco etch virus protease: mechanism of autolysis and rational design of

stable mutants with wild-type catalytic proficiency. *Protein Engineering, Design and Selection*, 14(12):993–1000, 2001.

[113] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of *Coot*. *Acta Crystallographica Section D*, 66(4):486–501, 2010.

[114] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson. Overview of the *CCP*4 suite and current developments. *Acta Crystallographica Section D*, 67(4):235–242, 2011.

[115] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, and M. Wojdyr. *CCP*4*i*2: the new graphical user interface to the *CCP*4 program suite. *Acta Crystallographica Section D*, 74(2):68–84, 2018.

[116] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, and A. A. Vagin. *REFMAC*5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D*, 67(4):355–367, 2011.

[117] W. Kabsch. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, 26(6):795–800, 1993.

[118] W. Kabsch. *XDS*. *Acta Crystallographica Section D*, 66(2):125–132, 2010.

[119] D. Turk. *MAIN* software for density averaging, model building, structure refinement and validation. *Acta Crystallographica Section D*, 69(8):1342–1357, 2013.

[120] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. *Phaser* crystallographic software. *Journal of Applied Crystallography*, 40(4):658–674, 2007.

[121] G. G. Langer, S. Hazledine, T. Wiegels, C. Carolan, and V. S. Lamzin. Visual automated macromolecular model building. *Acta Crystallographica Section D*, 69(4):635–641, 2013.

[122] V. S. Lamzin and K. S. Wilson. Automated refinement of protein models. *Acta Crystallographica Section D*, 49(1):129–147, 1993.

[123] M. W. Davis and E. M. Jorgensen. Ape, a plasmid editor: A freely available DNA manipulation and visualization program. *Frontiers in Bioinformatics*, 2, 2022.

[124] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner,

R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D*, 66(2):213–221, 2010.

[125] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, and P. D. Adams. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallographica Section D*, 75(10):861–877, 2019.

[126] N. S. Pannu, W.-J. Waterreus, P. Skubák, I. Sikharulidze, J. P. Abrahams, and R. A. G. de Graaff. Recent advances in the *CRANK* software suite for experimental phasing. *Acta Crystallographica Section D*, 67(4):331–337, 2011.

[127] P. Skubák and N. S. Pannu. Automatic protein structure solution from weak X-ray data. *Nature Communications*, 4(1):2777, 2013.

[128] P. J. Kraulis. *MOLSCRIPT*: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography*, 24(5):946–950, 1991.

[129] G. Lu. *TOP*: a new method for protein structure comparisons and similarity searches. *Journal of Applied Crystallography*, 33(1):176–183, 2000.

[130] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, 1990.

[131] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, R Wilkins, D Appel, and A Bairoch. *Protein Identification and Analysis Tools on the Expasy Server. Walker, M (ed): The proteomics protocols handbook*. Humana Press, 571-607, 2005, USA, 2005.

[132] A. Drozdetskiy, C. Cole, J. Procter, and G. J. Barton. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research*, 43(W1):W389–W394, 2015.

[133] X. Robert and P. Gouet. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1):W320–W324, 2014.

[134] The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1):D506–D515, 2018.

[135] C. Notredame, D. G. Higgins, and J. Heringa. T-coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, 302(1):205–217, 2000.

[136] P. Stothard. The sequence manipulation suite: Javascript programs for analyzing and formatting protein and DNA sequences. *BioTechniques*, 28(6):1102–1104, 2000.

[137] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The Protein Data Bank. *Nucleic Acids Research*, 28(1):235–242, 2000.

[138] H. T. Allawi and J. SantaLucia. Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry*, 36(34):10581–10594, 1997.

[139] C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall III, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, and D. C. Richardson. Molprobity: More and better reference data for improved all-atom structure validation. *Protein Science*, 27(1):293–315, 2018.

[140] H. Berman, K. Henrick, and H. Nakamura. Announcing the worldwide protein data bank. *Nature Structural & Molecular Biology*, 10(12):980–980, 2003.

[141] A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D. C. Hempel, and D. Jahn. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research*, 33(2):W526–W531, 2005.

[142] K. Mullis, S. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harb Symp Quant Biol, 51: 263-273, 1986*, pages 51: 263–273, 1986.

[143] E. Neumann, M. Schaefer-Ridder, Y. Wang, and P. H. Hofschneider. Gene transfer into mouse lyoma cells by electroporation in high electric fields. *The EMBO Journal*, 1(7):841–845, 1982.

[144] W. J. Dower, J. F. Miller, and C. W. Ragsdale. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Research*, 16(13):6127–6145, 1988.

[145] U. K. Laemmli. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259):680–685, 1970.

[146] H. Schägger. Tricine–SDS-PAGE. *Nature Protocols*, 1(1):16–22, 2006.

[147] G. D. Van Duyne, R. F. Standaert, P. A. Karplus, S. L. Schreiber, and J. Clardy. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *Journal of Molecular Biology*, 229(1):105–124, 1993.

[148] A. W. Schüttelkopf and D. M. F. van Aalten. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr.*, D60:1355–1363, 2004.

[149] F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus, and G. N. Murshudov. *AceDRG*: a stereochemical description generator for ligands. *Acta Crystallographica Section D*, 73(2):112–122, 2017.

[150] B. W. Matthews. Solvent content of protein crystals. *Journal of Molecular Biology*, 33(2):491–497, 1968.

[151] W. A. Hendrickson and M. M. Teeter. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature*, 290(5802):107–113, 1981.

[152] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J. A. McCammon, and N. A. Baker. Improvements to the APBS biomolecular solvation software suite. *Protein Science*, 27(1):112–128, 2018.

C Abbreviations

Å	Ångstrom
aa	amino acid
Ac	acetyl
ACP	acyl carrier protein
Act	actinorhodin
Amp	ampicillin
APE	aryl polyene
APEL	aryl polyene lipid
APS	ammonium persulfate
AQ	anthraquinone
Aro	aromatase
ARO	aromatic
ATHN	2-acetyl-1,3,6,8-tetrahydroxynaphthalene
AU	arbitrary unit
BGC	biosynthetic gene cluster
BS	biosynthesis
CA	cinnamic acid
C-C	carbon-carbon
C-domain	C-terminal domain
cFA	conjugated fatty acid
Cit	citric acid
CL	cinnamoyl lipid
Cm	chloramphenicol
CMK	chloromethylketone
CoA	coenzyme A
CPA	caproic acid
Cyc	cyclase
3D	three-dimensional
Da	dalton [g/mol]
ddH_2O	double-distilled water
DH, dehydratase	β-hydroxyacyl [ACP] dehydratase
Dha	dehydroalanine
DHNA	1,4-dihydroxy-2-naphthoyl
DHN-melanin	1,8-dihydroxynaphthalene-melanin
X,Y-DHN	X,Y-dihydroxynaphthalene

DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
ds	double-stranded
DTT	dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
et al.	et aliae
Eqn.	equation
ER	enoyl [ACP] reductase
Et	ethyl
FA	fatty acid
FAS	fatty acid synthase
Fig.	figure
4-HB	4-hydroxybenzoyl
H-bond	hydrogen bond
hepes	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HF	High-Fidelity
His_6	hexahistidine
HR	highly reducing
HX^Y	α-helix X in chain Y
IPTG	isopropyl-β-D-1-thiogalactopyranoside
Kan	kanamycin A
KR, ketoreductase	β-ketoacyl [ACP] reductase
KS, ketosynthase	β-ketoacyl [ACP] synthase
KS_α	ketosynthase, active subunit
KS_β	ketosynthase/chain-length factor
LB	lysogeny broth
LX^Y	loop after β-strand or α-helix X in chain Y
Mal	malonyl
MCAT	malonyl-CoA [ACP] transacylase
Me	methyl
min	minutes
MPD	2-methyl-2,4-pentanediol
MR	molecular replacement
MS	mass spectrometry
MW	molecular weight
MWCO	molecular weight cut-off
X-N	X-naphthol
$NADP^+$	nicotinamide adenine dinucleotide phosphate (ox.)
NADPH	nicotinamide adenine dinucleotide phosphate (red.)
N-domain	N-terminal domain
nESI	native electrospray ionization

NMR	nuclear magnetic resonance
NP	natural product
NRP	nonribosomal peptide
NRPS	nonribosomal peptide synthase
OAc	acetate
ox.	oxidized
PAC	phenacyl chloride
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PE	polyene
PEG	polyethylene glycol
PK	polyketide
PKS	polyketide synthase
PMSA	phenylmethanesulfonic acid
PMSF	phenylmethanesulfonyl fluoride
Ppant	4'-phosphopantetheine
red.	reduced
rmsd	root-mean-square deviation
rpm	revolutions per minute
RT	room temperature
s	seconds
SAD	single-wavelength anomalous dispersion
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
SF	sulfonyl fluorides
SI	sequence identity
SOC	super optimal broth with glucose
sp.	species
Spec	spectinomycin
SS	sequence similarity
SX^{Y}	β-strand X in chain Y
t	time
T	temperature
T_{A}	annealing temperature
Tab.	table
TAE	tris base, acetic acid, EDTA
t_{E}	time of elongation
TE	thioesterase
TE/CLC	thioesterase/claisen cyclase
TEMED	tetramethylethylenediamine

TEV	Tobacco Etch Virus
THN	tetrahydroxynaphthalene
tris	2-amino-2-(hydroxymethyl)propane-1,3-diol
UV/VIS	ultraviolet–visible
vdW	van der Waals
V_r	retention volume
v/v	volume per volume
WT	wild type
w/v	weight per volume

D List of Publications

(Co-)Authored publications in peer-reviewed journals or manuscripts in preparation that are part of this thesis:

- G. L. C. Grammbitter*, M. Schmalhofer*, Y.-M. Shi*, Y.-N. Shi*, S. P. B. Vemulapalli*, C. Richter, H. Schwalbe, M. Alanjary, A. Schüffler, M. Witt, C. Griesinger, M. Groll, and H. B. Bode. **The chemical structure of widespread microbial aryl polyene lipids.** *Equal contributing first author. *Manuscript in preparation.* First draft without structures was published in *bioRxiv*, 2020[38].
- M. Schmalhofer, A. L. Vagstad, H. B. Bode, and M. Groll. **Lyases drive polyketide trimming and cyclization in the biosynthesis of dihydroxynaphthalene-melanins and anthraquinones.** *Manuscript in preparation.*
- G. L. C. Grammbitter*, M. Schmalhofer*, H. Vural, N. J. Tobias, M. Groll, and H. B. Bode. **Anthraquione production is influenced by cinnamic acid.** *Equal contributing first author, *Manuscript in preparation.*
- G. L. C. Grammbitter, M. Schmalhofer, K. Karimi, Y.-M. Shi, T. A. Schöner, N. J. Tobias, N. Morgner, M. Groll, and H. B. Bode. **An uncommon type II PKS catalyzes biosynthesis of aryl polyene pigments.** *Journal of the American Chemical Society*, 141(42):16615–16623, 2019.
- Q. Zhou, A. Bräuer, H. Adihou, M. Schmalhofer, P. Saura, G. L. C. Grammbitter, V. R. I. Kaila, M. Groll, and H. B. Bode. **Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photorhabdus luminescens*.** *Chem. Sci.*, 10:6341–6349, 2019.

(Co-)Authored publications in peer-reviewed journals that are not part of this thesis:

- A. E. Fraley, M. Dell, M. Schmalhofer, R. A. Meoded, C. Bergande, M. Groll, and J. Piel. **Heterocomplex structure of a polyketide synthase component involved in modular backbone halogenation**. *Structure*, 31(5):565–572.e4, 2023
- A. Bräuer, Q. Zhou, G. L. C. Grammbitter, M. Schmalhofer, M. Rühl, V. R. I. Kaila, H. B. Bode, and M. Groll. **Structural snapshots of the minimal PKS system responsible for octaketide biosynthesis.** *Nature Chemistry*, 12(8):755–763, 2020.
- K. Rajeeve, N. Vollmuth, S. Janaki-Raman, T. F. Wulff, A. Baluapuri, F. R. Dejure, C. Huber, J. Fink, M. Schmalhofer, W. Schmitz, R. Sivadasan, M. Eilers, E. Wolf, W. Eisenreich, A. Schulze, J. Seibel, and T. Rudel. **Reprogramming of host glutamine metabolism during *Chlamydia trachomatis* infection and its key role in peptidoglycan synthesis.** *Nature Microbiology*, 5(11):1390–1402, 2020.
- J. Sobotta, M. Schmalhofer, T. M. Steiner, W. Eisenreich, G. Wächtershäuser, and C. Huber. **One-pot formation of 2,4-di- or 2,4,6-tri-olefinic monocarboxylic acids by straight chain C4-extension.** *Heliyon,* 3(7):e00368, 2017.

E Acknowledgements

Meine Promotion ist auf diese Art und Weise nur möglich gewesen durch die eine oder andere Unterstützung. Daher will ich mich an dieser Stelle bei allen bedanken, die mich über diese aufregende Zeit begleitet haben.

Mein hauptsächlicher Dank gilt Prof. Dr. Michael Groll für die Gelegenheit meine Forschung an seinem Lehrstuhl für Biochemie durchzuführen. Seine Begeisterung für die Wissenschaft und die Erforschung von Proteinstrukturen war sehr motivierend und hat stark abgefärbt. Seine Ratschläge und fast unendliche Zahl an Ideen haben mich bei vielen Projekten inspiriert. Ich bin sehr dankbar, dass ich unter diesen tollen Rahmenbedingungen forschen durfte.

Für drei erfolgreiche Kollaborationen bedanke ich mich nachdrücklich bei Prof. Dr. Helge Bode und Dr. Gina Grammbitter von dem MPI Marburg bzw. von der Goethe Universität Frankfurt, bei Dr. Anna Vagstad von der ETH Zürich sowie bei Prof. Dr. Piel und Dr. Amy Fraley von der ETH Zürich. Mein besonderer Dank gilt hier Gina und Helge für die wunderbare Kooperation, für die vielen Diskussionen und konstruktiven Gespräche.

Vielen Dank an alle meine aktuellen und ehemaligen Kollegen und Kolleginnen für die schöne Zeit in und außerhalb des Labors: Wolfgang H., Eva, Felix, Thomas B., Julia, Wolfgang K., Rachel, Annkathrin und Wolfgang E., Claudia, Jessica, Thomas G., Thomas S., Fan, Leonhard und Christopher.

Ich bedanke mich ganz herzlich bei Astrid, Raphaela, Katrin und Ute, denn ihr habt viele Dinge sehr vereinfacht.

Ebenso bedanke ich mich bei allen Studierenden für die Unterstützung im Labor.

Ganz besonders bedanke ich mich bei meiner Familie, Freunden und Freundinnen für euer echtes Interesse. Ihr habt immer an mich geglaubt und bei jeder Messung mitgefiebert.