



Bioinformatic prediction of the stereoselectivity of modular polyketide synthase: an update of the sequence motifs in ketoreductase domain

Changjun Xiang^{1,2}, Shunyu Yao², Ruoyu Wang² and Lihan Zhang^{*2,3,4}

Full Research Paper

Open Access

Address:

¹Department of Chemistry, Fudan University, Shanghai 200433, China, ²Key Laboratory of Precise Synthesis of Functional Molecules of Zhejiang Province, Department of Chemistry, School of Science and Research Center for Industries of the Future, Westlake University, Hangzhou 310030, China, ³Institute of Natural Sciences, Westlake Institute for Advanced Study, Hangzhou 310024, China and ⁴Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou 310030, China

Email:

Lihan Zhang* - zhanglihan@westlake.edu.cn

* Corresponding author

Keywords:

bioinformatics; conserved motifs; ketoreductase; polyketide synthase; stereocontrol

Beilstein J. Org. Chem. **2024**, *20*, 1476–1485.

<https://doi.org/10.3762/bjoc.20.131>

Received: 18 April 2024

Accepted: 21 June 2024

Published: 02 July 2024

This article is part of the thematic issue "Young investigators in natural products chemistry, biosynthesis, and enzymology".

Guest Editor: L. Barra



© 2024 Xiang et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Polyketides are a major class of natural products, including bioactive medicines such as erythromycin and rapamycin. They are often rich in stereocenters biosynthesized by the ketoreductase (KR) domain within the polyketide synthase (PKS) assembly line. Previous studies have identified conserved motifs in KR sequences that enable the bioinformatic prediction of product stereochemistry. However, the reliability and applicability of these prediction methods have not been thoroughly assessed. In this study, we conducted a comprehensive bioinformatic analysis of 1,762 KR sequences from *cis*-AT PKSs to reevaluate the residues involved in conferring stereoselectivity. Our findings indicate that the previously identified fingerprint motifs remain valid for KRs in β -modules from actinobacteria, but their reliability diminishes for KRs from other module types or taxonomic origins. Additionally, we have identified several new motifs that exhibit a strong correlation with the stereochemical outcomes of KRs. These updated fingerprint motifs for stereochemical prediction not only enhance our understanding of the enzymatic mechanisms governing stereocontrol but also facilitate accurate stereochemical prediction and genome mining of polyketides derived from modular *cis*-AT PKSs.

Introduction

Type I modular polyketide synthases (PKSs) are large enzyme complexes that play a crucial role in the biosynthesis of bacterial polyketides, including many important clinical drugs such as

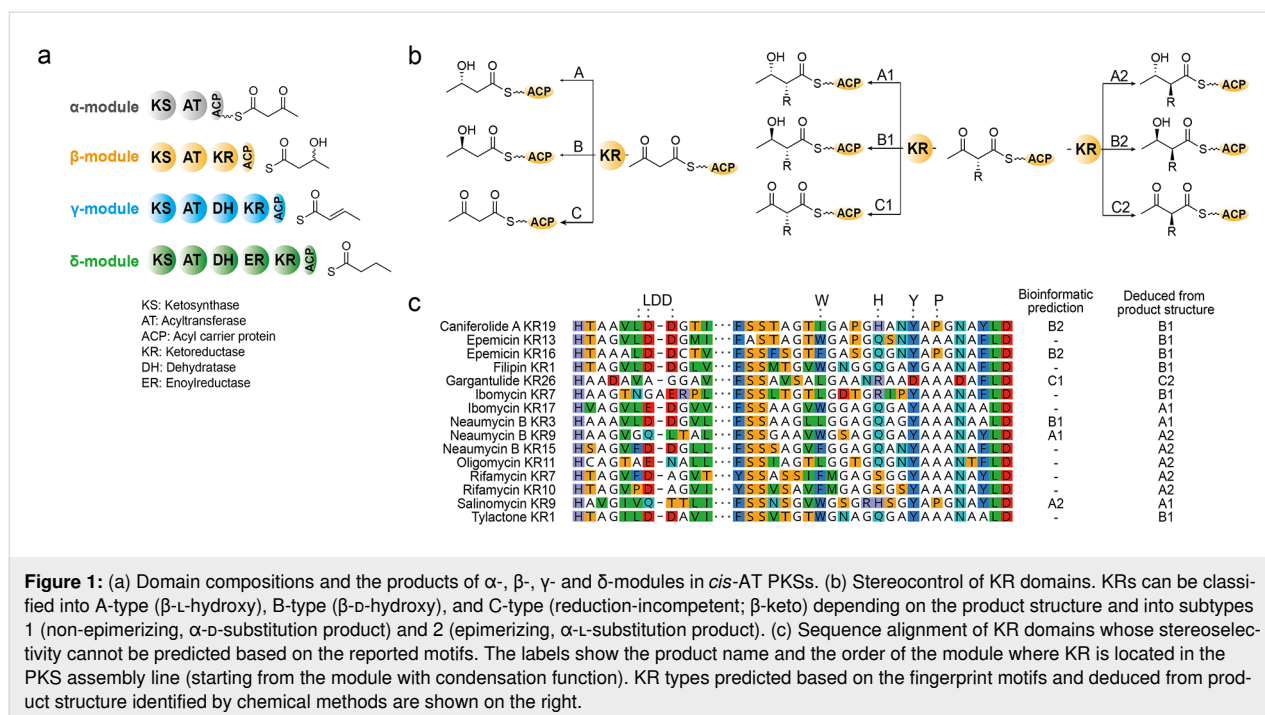
erythromycin (antibiotic), epothilone (anticancer), ivermectin (antiparasitic), and spinosyn (insecticide) [1,2]. Modular PKSs consist of multiple modules that catalyze one round of chain

extension by domains with different functions and are divided into *cis*-acyltransferase (*cis*-AT) and *trans*-AT PKSs depending on whether AT domain is embedded in the assembly line or not. All *cis*-AT PKS modules contain a ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP) to produce β -keto-intermediates, and some modules contain additional β -processing domains such as ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) [3,4]. The building blocks for PKS biosynthesis often include malonyl-CoA or methylmalonyl-CoA, which are loaded onto the ACP by the AT domain. Subsequently, the KS domain catalyzes the decarboxylative Claisen condensation between the ACP-tethered extender unit and the KS-tethered growing chain. The elongated growing chain may undergo further processing by KR, DH, and ER domains, generating β -hydroxy, α,β -alkene, and saturated β -methylene groups, respectively (Figure 1a).

KR domains have garnered significant attention from researchers due to their ability to control the stereochemistry at the α - and β -positions of a polyketide chain [5-7]. KR are classified into three types: A-type and B-type denote KR that form L- and D-configured β -hydroxy products, respectively, while C-type denotes KR incapable of reducing the β -keto group [8,9]. KR are further classified into subtypes 1 and 2 when the substrate contains an α -substitution. Subtype 1 (A1, B1, and C1) KR retain the original α -D-configuration, while subtype 2 (A2, B2, and C2) KR epimerize the α -carbon to yield α -L-configured products (Figure 1b). The stereocontrol of KR has been found to correlate with several conserved sequence motifs identified

through bioinformatic and structural analyses (Figure 1c) [8-11]. A pioneering work by Caffrey identified that A-type KR possess the conserved “W” motif but lack the “LDD” motif, whereas B-type KR possess the “LDD” motif but lack the “W” motif, based on the sequence alignment of 68 KR from 17 modular *cis*-AT PKSs [8]. In addition, Keatinge–Clay reported a conserved “H” motif in the sequence of A2-type KR and a “P” motif in B2-type KR as markers to distinguish them from the non-epimerizing A1/B1-type KR [9]. The presence of the catalytic “Y” motif and the absence of the NADPH binding motif can be used to predict C2-type KR [9]. These conserved motifs have been widely used to predict the stereochemical outcome of modular *cis*-AT PKSs and have facilitated bioinformatics-guided structural determination of complex polyketides [12-17].

However, despite being widely adopted, the prediction accuracy and the applicable range of these conserved motifs remain elusive, and not a few exceptions to these prediction rules have been reported (Figure 1c). For instance, the stereoselectivity of KR7 of ibomycin [12], KR15 of neaumycin B [14], and KR7 and KR10 of rifamycin [18] PKS cannot be accurately predicted due to the absence of both the LDD and W motifs. Moreover, KR3 of neaumycin B [14], KR19 of caniferolide A [15], KR16 of epemicin B [16], and KR26 of gargantulide B [17] show discrepancies between bioinformatic prediction and the determined product configuration. There are also instances where KR contain or lack both of the LDD and W motifs, making it difficult to predict stereochemistry (Figure 1c).



With the increasing availability of whole genome sequencing, a growing number of PKS gene clusters and their products have been characterized [19]. To assess and evaluate the sequence–stereoselectivity relationship of KRs, here we have collected 1,762 KR sequences from modular *cis*-AT PKS gene clusters, whose product structures have been verified using spectroscopic and/or chemical methods. We reveal that the previously identified conserved motifs are best applicable to KRs in the β -module of actinobacterial PKSs but less applicable for other types of KRs. Moreover, we have identified additional fingerprint residues that improve stereochemical prediction. These fingerprint residues also suggest potential mode of interactions among KR, ACP, and DH domains, deepening our understanding of the stereocontrol of PKSs.

Results and Discussion

Preparation of KR sequence dataset

We first curated the amino acid sequences of KR domains from characterized bacterial *cis*-AT PKSs recorded in MIBiG database [20] and by manual literature review. In total, 1,762 KRs whose product structures were experimentally determined, such as by crystallography, nuclear magnetic resonance (NMR) analysis or by chemical synthesis, were obtained for further analysis.

The modules in PKSs were categorized as α -module (containing KS-AT-ACP tridomain), β -module (KS-AT-KR-ACP), γ -module (KS-AT-DH-KR-ACP), and δ -module (KS-AT-DH-ER-KR-ACP) based on their domain composition regardless of the domain activity [2] (Figure 1a). While the stereoselectivity of a KR in a β -module can be directly inferred from the product hydroxy group, the stereochemical outcome of a KR from a γ - and δ -module is obscured by the following dehydration by the DH domain. It is widely believed that an α,β -*trans* (*E*) double bond is generated from a D- β -hydroxy intermediate produced by a B-type KR, whereas a *cis* (*Z*) double bond is from an L- β -hydroxy intermediate generated by an A-type KR via *syn* elimination [21]. However, based on the *syn*-elimination mechanism, the L- α -methyl,L- β -hydroxy substrate produced by A2-type KR can also result in a *trans* (*E*) double bond [18,22], and some DH domains are reported to have epimerase activity on the α -substitution [23]. Thus, the stereochemical outcomes of KRs in γ - and δ -modules cannot be directly inferred from the final product structure, and we discuss these KRs separately based on their module types.

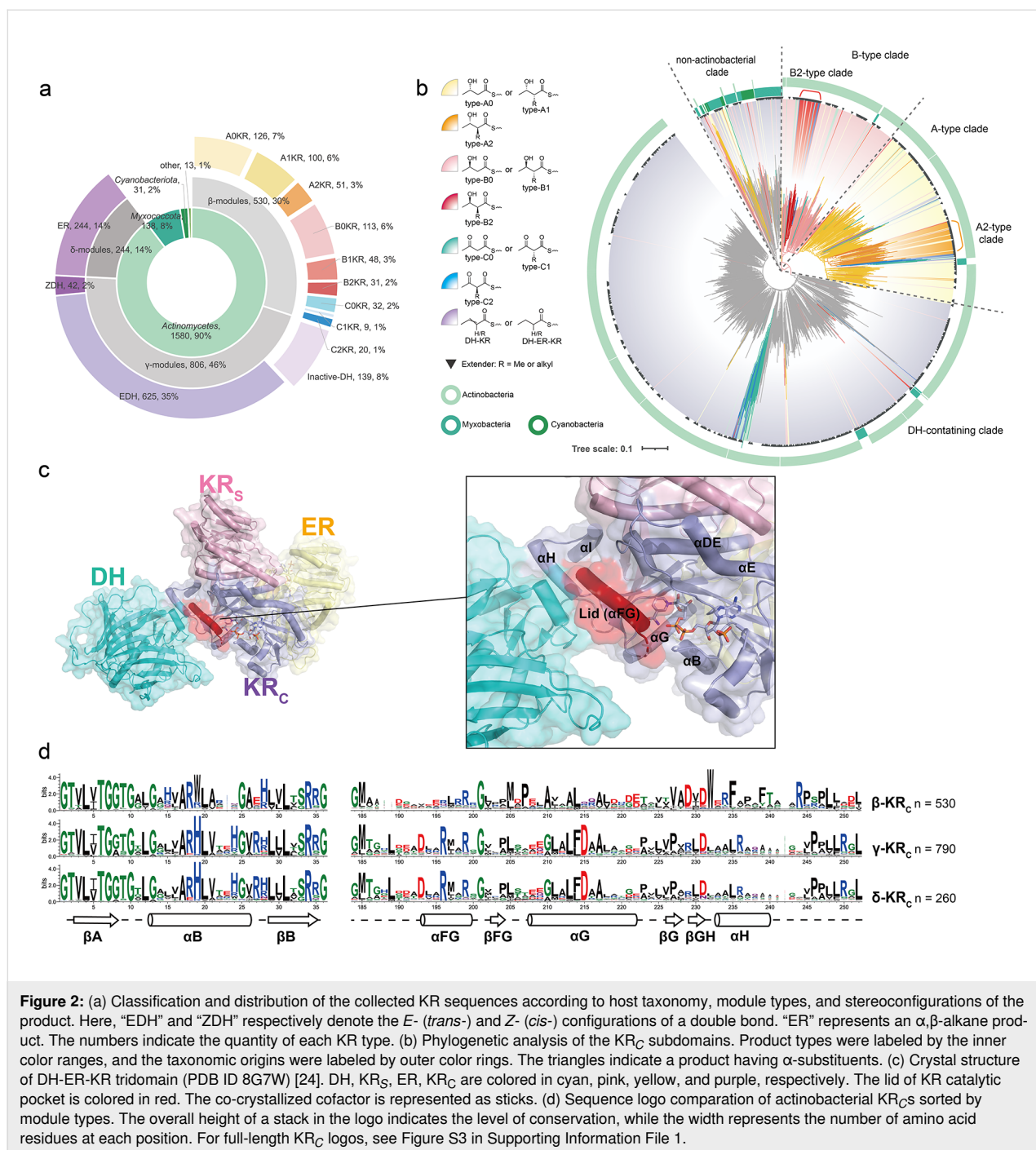
The curated 1,762 KR sequences were grouped into different types based on their product structure and the taxonomy of their host strains (Figure 2a). Among the KR sequences we collected, 90% are from *Actinomycetota*, and less than 10% are from

Myxococcota and *Cyanobacteriota*. Among actinobacterial KRs, γ -modules accounted for more than half, of which approximately 78% resulting in a *trans* (*E*) double bond, 5% *cis* (*Z*) double bond, and the remaining 17% possessing an inactive DH domain. Among actinobacterial KRs from β -modules, A-type KRs (52%) are the most abundant, followed by B-type KRs (36%), and C-type KRs (12%).

Comparison between KR $_{\gamma}$ and KR $_{\delta}$ subdomains

KR domains are structurally divided into two subdomains: a catalytic subdomain (KR $_{\delta}$) with an intact Rossmann fold where the active site resides, and a structural subdomain (KR $_{\gamma}$) with a truncated Rossmann fold that lacks NADPH binding sites and solely provides structural support by forming a heterodimer with KR $_{\delta}$ [24–26]. In δ -modules, an ER domain is inserted between KR $_{\gamma}$ and KR $_{\delta}$. Phylogenetic analyses of KR $_{\delta}$ and KR $_{\gamma}$ revealed that both subdomains exhibited a cladogram pattern dependent on taxonomy, with cyanobacterial and myxobacterial KRs forming an ancestral clade followed by actinobacterial KRs (Figure 2b, and Supporting Information File 1, Figures S1 and S2). Such cladogram pattern was also observed in their KS domains [27], suggesting that they have diverged at an early point during evolution. Furthermore, both KR $_{\gamma}$ and KR $_{\delta}$ trees showed that KRs belonging to γ - and δ -modules formed a separate clade distinct from those in β -modules, indicating that the existence of DH domain significantly influences the sequence of both KR subdomains [13]. This suggests a strong domain–domain interaction between KR and DH, as observed in the structure of KR-DH-ER tridomain (Figure 2c) [24], and indicates that KRs in β -modules and γ/δ -modules may not be generally interchangeable for domain swapping.

However, the phylogenetic cladograms of the two subdomains were significantly different within the clade formed by β -modules. In the KR $_{\delta}$ tree, stereoselectivity-dependent clades were formed, but such clades were not observed in the KR $_{\gamma}$ tree (Figure 2b and Figure S1 in Supporting Information File 1). This indicates that the stereoselectivity of KR is controlled solely by KR $_{\delta}$, and KR $_{\gamma}$ does not influence its catalytic selectivity. Moreover, A2- and B2-type KR $_{\delta}$ s formed a separable clade from A1- and B1-type KR $_{\delta}$ s, respectively, suggesting that phylogenetic analysis can be used for stereochemical prediction. It is noteworthy that A0- (0 denotes as product without α -substitution) and A1-type KRs, and B0- and B1-type KRs cannot be distinguished in the phylogenetic trees of both subdomains. This implies that KRs likely exhibit promiscuity for α -substitution, and the structural differences in the α -position should not affect the original stereoselectivity of a KR. Such promiscuity can also be observed in the MycA KR in mycolactone biosynthesis, which accepts both α -substituted and α -unsubstituted



substrates while retaining the same stereoselectivity for the β -hydroxy group [28].

To further investigate the sequence features of each module type, we performed sequence logo analyses of KR_C subdomains from β -, γ - and δ -modules. Consistent with the phylogenetic cladogram, KR_Cs from γ - and δ -modules showed nearly identical logo features each other but were significantly different from the logo of β -modules. Major differences were ob-

served in two regions of KR_C, the *N*-terminal helix α B and the lid region α FG located at the *C*-terminal of KR_C (Figure 2d). Indeed, this lid region exhibits a direct interaction with the DH and DH-KR linker (Figure 2c). However, such interaction was not observed between the *N*-terminal helix α B and DH in the crystal structure of DH-ER-KR tridomain (Figure 2c). This finding suggests that the helix α B may have an allosteric interaction with DH or potential large conformational changes between DH and KR during catalysis.

was found in B-type KR_Cs, while a XLXR motif was found in A-type KR_Cs. A highly conserved T (**5**) was also identified in B-type KR_Cs, whereas it was replaced by a more hydrophobic L in A-type KR_Cs. On the loop between α E and β E, a conserved D (**6**) was found in A-type KR_Cs.

For A2-type KR_Cs, we identified a highly conserved L (**2**) in the position corresponding to the LDD motif (Figure 3). This residue may participate in interactions with the lid helix to adjust the pocket for epimerization of α -substituted intermediates. The lack of the conserved GM (**11**) motif preceding the α FG helix, which is highly conserved in other types of reductase-competent KR_Cs, was also observed in A2-type KR_Cs. The absence of this region in the electron density map of the A2-type AmpKR11 [30] may indicate a more flexible structure of the A2-type lid helix.

In B2-type KR_Cs, we identified a conserved H (**13**) located at the end of lid helix (α FG) and the previously identified P (**10**) as signature motifs (Figure 3). Slight differences were observed in the loop region between α EF and α F, such as the more conserved P and less conserved Q, preceding the catalytic Y (**9**). Furthermore, some differences were noted in the α DE region downstream of the LDD motif, such as the R to Q mutation (**3**) and a conserved (D/E)RVLR between the R (**3**) and K (**4**) motifs, which may also regulate the structure of the loop forming the catalytic groove.

The sequence logo of C0/C1-type KR_Cs showed that some of them possess the catalytic Y (**9**), but such KR_Cs instead possess mutation in NADPH binding site (Figure S3 in Supporting Information File 1). In C2-type KR_Cs, we found that most C2-type KR_Cs possessed the catalytic Y (**9**), but some possessed Q at this residue. This substitution can be reasonable, as it serves as a hydrogen bond donor to activate the β -keto moiety of a substrate [31,32]. Additionally, the presence of K (**4**), which activates the catalytic tyrosine, was more conserved in C2-type KR_Cs. In general, C2-KR_Cs show similarity to either A2- or B2-KR_Cs, but are likely to contain mutations in NADPH binding site [31].

Sequence logo analysis of KR_C from γ - and δ -modules

In the actinobacterial γ - and δ -modules, the stereochemical outcome of a KR is obscured by further dehydration catalyzed by DH. The dehydration mechanism by DH has been investigated in several DH domains, which all follow the *syn*-elimination mechanism [5,21,33]. Based on the product configuration, we classified DHs into two types, *E*-type DHs (*trans*-double bond) and *Z*-type DHs (*cis*-double bond) and analyzed sequence logos of the DH-associated KR_Cs.

The sequence logo analysis of these KR_Cs revealed that the LDD (**2**) motif representative of B-type KR is highly conserved in KR_Cs associated with *E*-type DHs (Figure 4 and Figure S5 in Supporting Information File 1), which is in agreement with previous proposal that the *trans* (*E*) double bond is formed from the D - β -hydroxyacyl intermediate produced by B-type KR, and a *cis* (*Z*) double bond is formed from the L - β -hydroxyacyl intermediate [5]. However, only 5 out of 42 KR_Cs associated with *Z*-type DHs contain the conserved W (**7**) motif found in A-type KR_Cs. Additionally, some KR_Cs associated with *Z*-type DHs possessed the LDD (**2**) motif. These observations suggest several possibilities: (1) the stereoselectivity prediction for KR_Cs in γ - and δ -modules is inaccurate, (2) a *cis* (*Z*) double bond can also be produced from D - β -hydroxyacyl intermediates due to a different substrate binding mode in DH [7], or (3) a *cis* (*Z*) double bond is formed during post-PKS modification [22,34].

To gain a better understanding of the stereoselectivity of KR_Cs in γ - and δ -modules, we focused on KR_Cs associated with an inactive DH domain (DH⁰) that still produce a β -hydroxy intermediate. These KR_Cs are phylogenetically mixed together with those in γ - and δ -modules with an active DH (Supporting Information File 1, Figure S2), making them good models for analyzing the stereoselectivity of DH-associated KR_Cs. Based on product structures, we found eleven A-type KR_Cs and as many as 102 B-type KR_Cs from γ - and δ -modules, similar to the ratio of *Z*- and *E*-DHs (Figure 4). However, all A-type KR_Cs associated with an inactive DH lacked the diagnostic W (**7**) motif, and other A-type characteristic motifs such as L (**5**) and D (**6**) were also missing or less conserved (Figure 4). Contrary, nearly all B-type KR_Cs associated with an inactive DH possessed the second D in the LDD (**2**) and the T (**5**) motifs. B2-type KR_Cs showed slightly less conserved LD in the LDD motif, which was also reported as a feature of KR_Cs in *trans*-AT PKSs [11]. In γ -A2-type KR, the characteristic L (**2**) and H (**8**) motifs found in β -A2-type KR were absent, and γ - and δ -B2-type KR lacked the characteristic P (**10**) and H (**13**) features. Therefore, the situation for KR_Cs in γ - and δ -modules is complicated: a KR containing the second D in the LDD motif is likely to be B-type, but the possibility of it being A-type cannot be excluded, and the stereochemistry of the α -carbon cannot be predicted.

Towards better bioinformatic prediction of stereochemical outcomes

Based on the above analyses, here we provide an update in the scope and motifs used for bioinformatic prediction of the α -methyl (or -alkyl) and β -hydroxy stereochemistry in modular *cis*-AT PKSs. First, actinobacterial β -module KR_Cs are good targets for bioinformatic stereochemical prediction by the previously reported motifs, and newly identified motifs can further support the prediction. However, other KR_Cs cannot be pre-

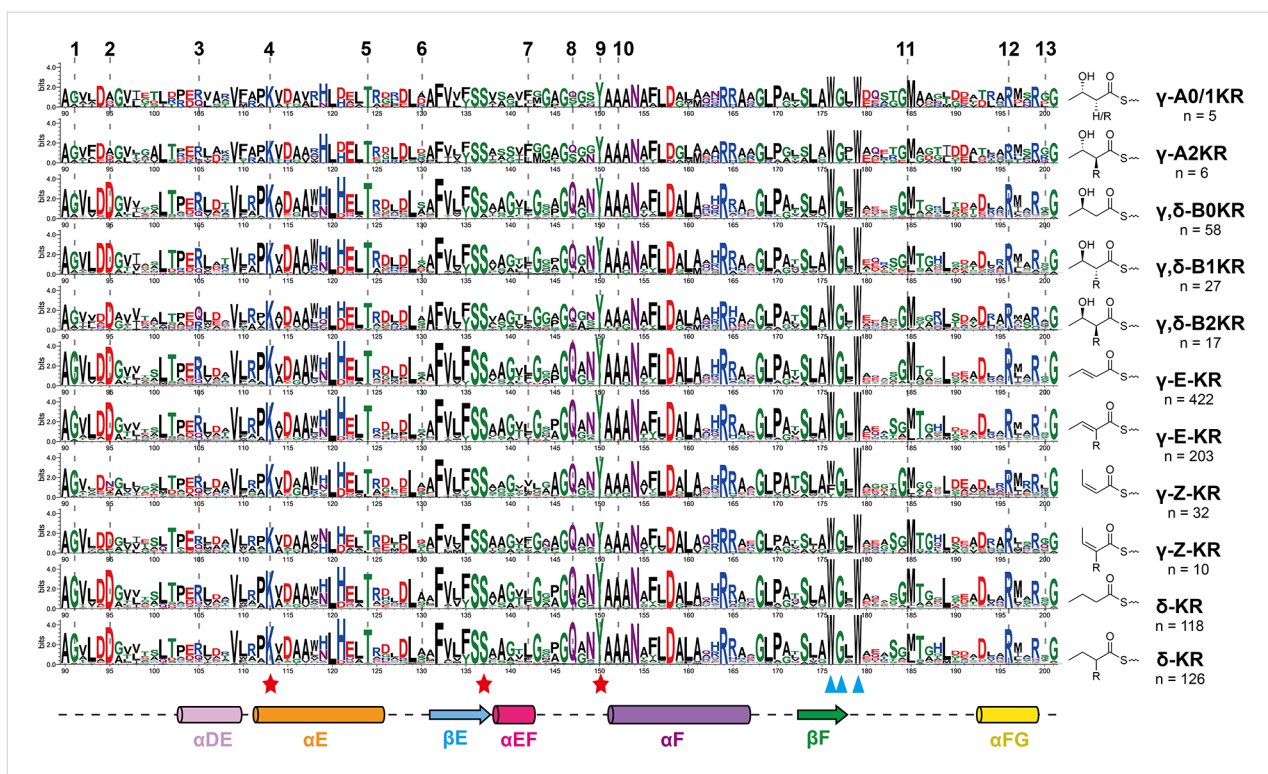


Figure 4: Sequence logo comparison of γ - and δ -module KR_C based on the classification of their products. Top five rows show KR_C associated with an inactive DH that produces the hydroxy products. The key catalytic residues are marked by red stars, and the NADPH-binding residues (partially) are marked by blue triangles. The motif numbers at the top are corresponding to the location of fingerprints in Figure 3. For full-length sequence logo, see Figure S5 in Supporting Information File 1.

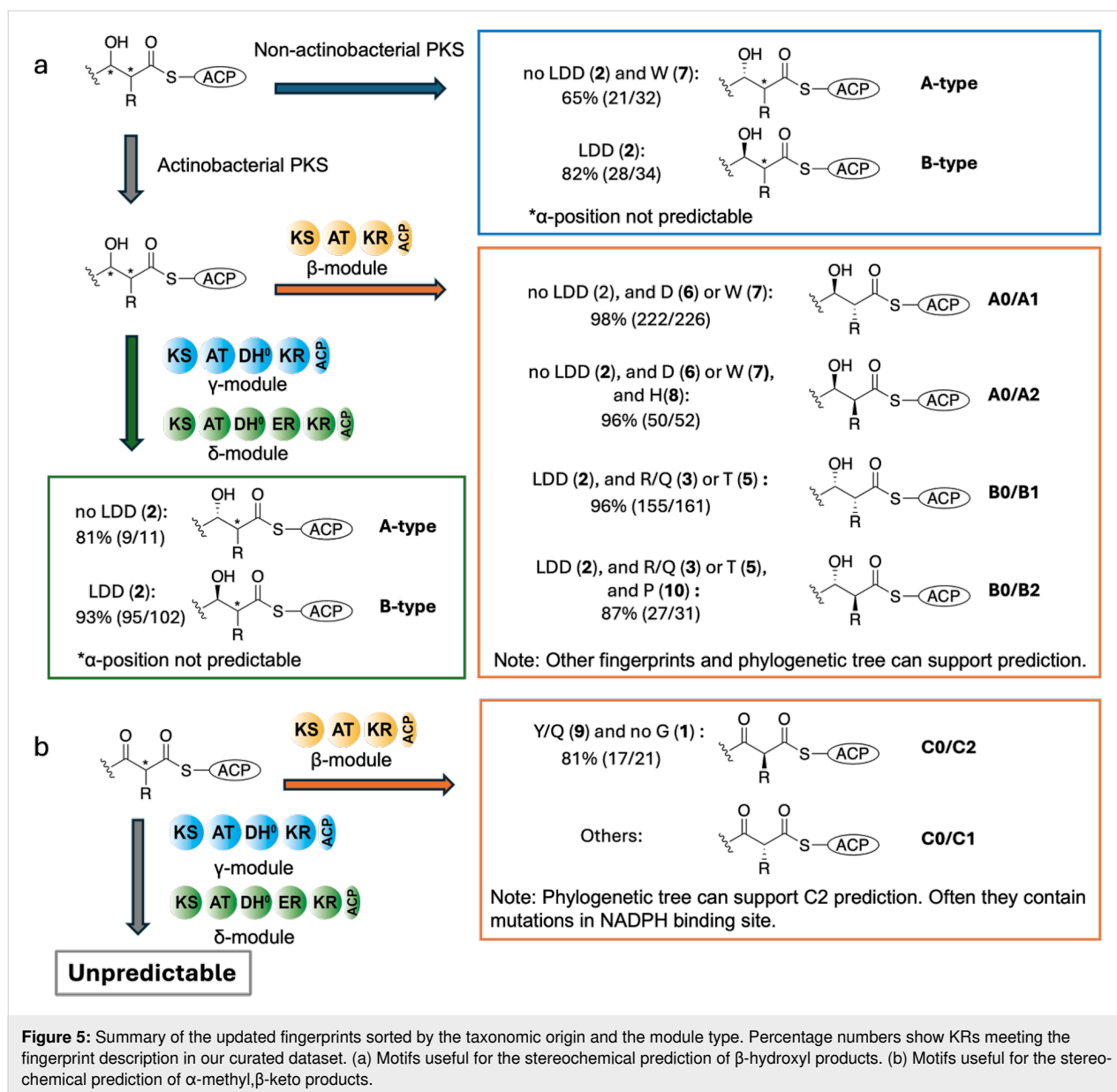
dicted with high accuracy, especially for the stereochemistry at the α -position. Second, phylogenetic analysis of KR_C subdomain is an effective means to predict stereoselectivity of KR_C from actinobacterial β -modules. Specifically, our results indicate that the stereochemistry can be predicted well using the fingerprint motifs if the KR falls within the A- and B-type clades by phylogenetic analysis of KR_C subdomain (and most of actinobacterial β -module KR_C do fall within there). Overall, the motifs having correlation with stereoselectivity are summarized in Table S1 in Supporting Information File 1. Below we provide a detailed guideline for stereochemical prediction of KR_C.

For non-actinobacterial PKSs, stereochemistry can be predicted by the presence of LDD motif for B-type. The absence of the LDD and the presence of the W motif are indicative for A-type, but with moderate accuracy (Figure 5).

For actinobacterial β -module KR_C, stereochemistry can be predicted by using the LDD and W, as well as the newly identified R/Q (3), T (5), and D (6) motifs as summarized in Figure 5. About 87% to 98% of the KR_C belonging to each product type satisfied our prediction criteria, verifying the accuracy of stereochemical prediction within this group of KR_C.

In contrast, for actinobacterial γ - and δ -module KR_C, the newly identified motifs cannot guide the prediction, and only A- and B-type can be roughly predicted by the presence and absence of the LDD motif. However, phylogenetic analysis can support bioinformatic prediction if a KR falls within A1/A2 or B1/B2 clades: for example, fostriecin KR2, which produces a *cis* double bond together with a DH, locates in the A1-type clade, in agreement with its stereoselectivity [35].

Lastly, we applied our prediction criteria for the previously mispredicted or unpredictable β -module KR_C, for example due to having both LDD and W motifs, for validation. Among nine KR_C we analyzed, five can be accurately predicted by either conserved motif analysis or by phylogenetic analysis, illustrating the advantage of using multiple strong predictive motifs (Table S2 in Supporting Information File 1). It is also worth noting that some of the KR_C analyzed here may contain wrong stereochemical assignment of the product, potentially derived from misinterpretation of NMR analyses [14,36,37]. Nevertheless, many of the previously reported unpredictable KR_C are from non-actinobacteria or from γ - and δ -module KR_C, and bioinformatic prediction of these KR_C would require further studies and more characterized sequences.



Conclusion

In summary, we systematically assessed the stereoselectivity-associated sequence motifs from 1,762 KR sequences in bacterial modular *cis*-AT PKSs whose product structures were experimentally determined. Our analyses revealed that different KR_Cs, by taxonomic origin or by module types, have differences in the fingerprint motifs, affecting the prediction accuracy by conserved motif analysis. We identified several additional fingerprint motifs in the KR_C subdomain that can be used to better predict the stereochemistry of KR_Cs from actinobacterial β-modules. The identified motifs reside at the interface of domain–domain or domain–substrate interactions, shedding light on the enzyme mechanism for stereocontrol. Our work provides an overview for current bioinformatic prediction of

stereoselectivity of KR in *cis*-AT PKS, expands the understanding of the stereocontrol of PKS from bioinformatics perspective, and will facilitate accurate stereochemical prediction and genome mining of complex polyketides.

Experimental

Sequence collection and screening

All *cis*-AT PKS and PKS-NRPS amino acid sequences recorded in MIBiG, as well as the PKS sequences reported in literatures we have curated were targeted for further analysis. Sequences with multiple modules on a peptide chain were cut into single-module sequences based on the boundaries annotated of modules. These sequences were then filtered according to the following rules: (a) The PKS sequences that produce the

same final product but originate from different strains of bacteria only retain the sequences from one of the sources. For example, only the erythromycin PKS from *Saccharopolyspora erythraea* were retained. (b) To associate sequence information with reliable absolute configuration of the product, the stereochemical configurations of collected polyketide products were manually checked by literature searches. Absolute configurations determined entirely by chemical methods, such as crystal structures, NMR, chemical degradation and derivatization, were considered reliable. Alternatively, relative configurations determined by NMR methods that corresponded exactly to the results predicted by bioinformatics were also considered reliable. Compounds of which only the relative configurations were elucidated were excluded from the dataset. (c) Sequences for which it was impossible to infer the stereochemistry of KR product were removed, such as M1 to M3 of rifamycin PKS whose product experiences later aromatization. (d) Modules with unconventional module compositions (e.g., modules with two KR domains), modules without KR domains, and loading modules were excluded from the data. After organizing and filtering, the KR_S and KR_C subdomain sequences were extracted for further bioinformatic analysis.

Bioinformatic analysis

The multiple sequence alignment, sequence logo analysis, and tree building were performed on Geneious Prime 2023.2 (<https://www.geneious.com>) with options shown below: multiple sequence alignments were generated by MAFFT 7.450 (E-INS-i option), Neighbor-Joining trees were built by Geneious Tree Builder using Jukes-Cantor distance model. Sequence logo visualization was conducted by WebLogo 3. Protein structural analysis was conducted by PyMOL.

Supporting Information

Supporting Information File 1

Supporting Figures and Tables.

[<https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-20-131-S1.pdf>]

Supporting Information File 2

Supporting dataset.

[<https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-20-131-S2.xlsx>]

Acknowledgements

We thank the Zhang lab members Zilei Huang, Shengling Xie, Zhihan Zhang, Shanchong Chen, Cui Yang, Fanzhong Zhang, for KR sequence retrieval from public database.

Funding

This research was supported by the “Pioneer” and “Leading Goose” R&D Program of Zhejiang (2023SDXHDX0007) and National Natural Science Foundation of China General Program (22177092) to L.Z.

Author Contributions

Changjun Xiang: data curation; formal analysis; investigation; writing – original draft. Shunyu Yao: data curation. Ruoyu Wang: data curation. Lihan Zhang: conceptualization; data curation; formal analysis; funding acquisition; supervision; writing – review & editing.

ORCID® iDs

Ruoyu Wang - <https://orcid.org/0009-0004-3368-0265>

Lihan Zhang - <https://orcid.org/0000-0003-4364-6242>

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information to this article.

References

- Hertweck, C. *Angew. Chem., Int. Ed.* **2009**, *48*, 4688–4716. doi:10.1002/anie.200806121
- Keatinge-Clay, A. T. *Chem. Rev.* **2017**, *117*, 5334–5366. doi:10.1021/acs.chemrev.6b00683
- Keatinge-Clay, A. T. *Nat. Prod. Rep.* **2012**, *29*, 1050–1073. doi:10.1039/c2np20019h
- Grininger, M. *Nat. Chem. Biol.* **2023**, *19*, 401–415. doi:10.1038/s41589-023-01277-7
- Keatinge-Clay, A. T. *Nat. Prod. Rep.* **2016**, *33*, 141–149. doi:10.1039/c5np00092k
- Drufva, E. E.; Spengler, N. R.; Hix, E. G.; Bailey, C. B. *ChemBioChem* **2021**, *22*, 1122–1150. doi:10.1002/cbic.202000613
- Weissman, K. J. *Beilstein J. Org. Chem.* **2017**, *13*, 348–371. doi:10.3762/bjoc.13.39
- Caffrey, P. *ChemBioChem* **2003**, *4*, 654–657. doi:10.1002/cbic.200300581
- Keatinge-Clay, A. T. *Chem. Biol.* **2007**, *14*, 898–908. doi:10.1016/j.chembiol.2007.07.009
- Kitsche, A.; Kalesse, M. *ChemBioChem* **2013**, *14*, 851–861. doi:10.1002/cbic.201300063
- Piasecki, S. K.; Zheng, J.; Axelrod, A. J.; Detelich, E. M.; Keatinge-Clay, A. T. *Proteins: Struct., Funct., Bioinf.* **2014**, *82*, 2067–2077. doi:10.1002/prot.24561
- Robbins, N.; Spitzer, M.; Wang, W.; Waglechner, N.; Patel, D. J.; O'Brien, J. S.; Ejim, L.; Ejim, O.; Tyers, M.; Wright, G. D. *Cell Chem. Biol.* **2016**, *23*, 1383–1394. doi:10.1016/j.chembiol.2016.08.015
- Zhang, L.; Hashimoto, T.; Qin, B.; Hashimoto, J.; Kozono, I.; Kawahara, T.; Okada, M.; Awakawa, T.; Ito, T.; Asakawa, Y.; Ueki, M.; Takahashi, S.; Osada, H.; Wakimoto, T.; Ikeda, H.; Shin-ya, K.; Abe, I. *Angew. Chem., Int. Ed.* **2017**, *56*, 1740–1745. doi:10.1002/anie.201611371

14. Kim, M. C.; Machado, H.; Jang, K. H.; Trzoss, L.; Jensen, P. R.; Fenical, W. *J. Am. Chem. Soc.* **2018**, *140*, 10775–10784. doi:10.1021/jacs.8b04848
15. Pérez-Victoria, I.; Oves-Costales, D.; Lacret, R.; Martín, J.; Sánchez-Hidalgo, M.; Díaz, C.; Cautain, B.; Vicente, F.; Genilloud, O.; Reyes, F. *Org. Biomol. Chem.* **2019**, *17*, 2954–2971. doi:10.1039/c8ob03115k
16. Kontou, E. E.; Gren, T.; Ortiz-López, F. J.; Thomsen, E.; Oves-Costales, D.; Díaz, C.; de la Cruz, M.; Jiang, X.; Jørgensen, T. S.; Blin, K.; Charusanti, P.; Reyes, F.; Genilloud, O.; Weber, T. *ACS Chem. Biol.* **2021**, *16*, 1456–1468. doi:10.1021/acscchembio.1c00318
17. Carretero-Molina, D.; Ortiz-López, F. J.; Gren, T.; Oves-Costales, D.; Martín, J.; Román-Hurtado, F.; Sparholt Jørgensen, T.; de la Cruz, M.; Díaz, C.; Vicente, F.; Blin, K.; Reyes, F.; Weber, T.; Genilloud, O. *Org. Chem. Front.* **2022**, *9*, 462–470. doi:10.1039/d1qo01480c
18. You, Y.-O.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2013**, *135*, 7406–7409. doi:10.1021/ja4014776
19. Kishore, S.; Khosla, C. *Open Biol.* **2023**, *13*, 230096. doi:10.1098/rsob.230096
20. Terlouw, B. R.; Blin, K.; Navarro-Muñoz, J. C.; Avalon, N. E.; Chevrette, M. G.; Egbert, S.; Lee, S.; Meijer, D.; Recchia, M. J. J.; Reitz, Z. L.; van Santen, J. A.; Selem-Mojica, N.; Tørring, T.; Zaroubi, L.; Alanjary, M.; Aleti, G.; Aguilar, C.; Al-Salihi, S. A. A.; Augustijn, H. E.; Avelar-Rivas, J. A.; Avitia-Domínguez, L. A.; Barona-Gómez, F.; Bernaldo-Aguero, J.; Bielinski, V. A.; Biermann, F.; Booth, T. J.; Carrion Bravo, V. J.; Castelo-Branco, R.; Chagas, F. O.; Cruz-Morales, P.; Du, C.; Duncan, K. R.; Gavriilidou, A.; Gayrard, D.; Gutiérrez-García, K.; Haslinger, K.; Helfrich, E. J. N.; van der Hooft, J. J. J.; Jati, A. P.; Kalkreuter, E.; Kalyvas, N.; Kang, K. B.; Kautsar, S.; Kim, W.; Kunjapur, A. M.; Li, Y.-X.; Lin, G.-M.; Loureiro, C.; Louwen, J. J. R.; Louwen, N. L. L.; Lund, G.; Parra, J.; Philmus, B.; Pourmohsenin, B.; Pronk, L. J. U.; Rego, A.; Rex, D. A. B.; Robinson, S.; Rosas-Becerra, L. R.; Roxborough, E. T.; Schorn, M. A.; Scobie, D. J.; Singh, K. S.; Sokolova, N.; Tang, X.; Udway, D.; Vigneshwari, A.; Vind, K.; Vromans, S. P. J. M.; Waschulin, V.; Williams, S. E.; Winter, J. M.; Witte, T. E.; Xie, H.; Yang, D.; Yu, J.; Zdouc, M.; Zhong, Z.; Collemare, J.; Linington, R. G.; Weber, T.; Medema, M. H. *Nucleic Acids Res.* **2023**, *51*, D603–D610. doi:10.1093/nar/gkac1049
21. Yin, Z.; Dickschat, J. S. *Nat. Prod. Rep.* **2021**, *38*, 1445–1468. doi:10.1039/d0np00091d
22. Gay, D.; You, Y.-O.; Keatinge-Clay, A.; Cane, D. E. *Biochemistry* **2013**, *52*, 8916–8928. doi:10.1021/bi400988t
23. Xie, X.; Garg, A.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2017**, *139*, 9507–9510. doi:10.1021/jacs.7b05502
24. McCullough, T. M.; Dhar, A.; Akey, D. L.; Konwerski, J. R.; Sherman, D. H.; Smith, J. L. *Structure* **2023**, *31*, 1109–1120.e3. doi:10.1016/j.str.2023.05.019
25. Keatinge-Clay, A. T.; Stroud, R. M. *Structure* **2006**, *14*, 737–748. doi:10.1016/j.str.2006.01.009
26. Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Clay, A. T. *Nat. Chem. Biol.* **2012**, *8*, 615–621. doi:10.1038/nchembio.964
27. Hirsch, M.; Fitzgerald, B. J.; Keatinge-Clay, A. T. *ACS Chem. Biol.* **2021**, *16*, 2515–2526. doi:10.1021/acscchembio.1c00598
28. Piasecki, S. K.; Taylor, C. A.; Detelich, J. F.; Liu, J.; Zheng, J.; Komsoukaniants, A.; Siegel, D. R.; Keatinge-Clay, A. T. *Chem. Biol.* **2011**, *18*, 1331–1340. doi:10.1016/j.chembiol.2011.07.021
29. Liu, C.; Yuan, M.; Xu, X.; Wang, L.; Keatinge-Clay, A. T.; Deng, Z.; Lin, S.; Zheng, J. *J. Struct. Biol.* **2018**, *203*, 135–141. doi:10.1016/j.jsb.2018.04.001
30. Zheng, J.; Piasecki, S. K.; Keatinge-Clay, A. T. *ACS Chem. Biol.* **2013**, *8*, 1964–1971. doi:10.1021/cb400161g
31. Zheng, J.; Keatinge-Clay, A. T. *J. Mol. Biol.* **2011**, *410*, 105–117. doi:10.1016/j.jmb.2011.04.065
32. Xie, X.; Garg, A.; Keatinge-Clay, A. T.; Khosla, C.; Cane, D. E. *Biochemistry* **2016**, *55*, 1179–1186. doi:10.1021/acs.biochem.6b00024
33. Keatinge-Clay, A. *J. Mol. Biol.* **2008**, *384*, 941–953. doi:10.1016/j.jmb.2008.09.084
34. Kandziora, N.; Andexer, J. N.; Moss, S. J.; Wilkinson, B.; Leadlay, P. F.; Hahn, F. *Chem. Sci.* **2014**, *5*, 3563–3567. doi:10.1039/c4sc00883a
35. Shah, D. D.; You, Y.-O.; Cane, D. E. *J. Am. Chem. Soc.* **2017**, *139*, 14322–14330. doi:10.1021/jacs.7b08896
36. Zhang, S.; Zhang, S.; Fan, Y.; Zhang, X.; Chen, J.; Jin, C.; Chen, S.; Wang, L.; Zhang, Q.; Chen, Y. *Angew. Chem., Int. Ed.* **2023**, *62*, e202313186. doi:10.1002/anie.202313186
37. Ding, J.; Smith, A. B., III. *J. Am. Chem. Soc.* **2023**, *145*, 18240–18246. doi:10.1021/jacs.3c06573

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (<https://www.beilstein-journals.org/bjoc/terms>), which is identical to the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at: <https://doi.org/10.3762/bjoc.20.131>