JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

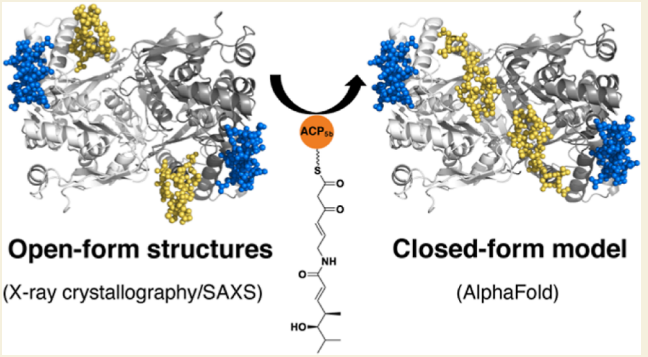
pubs.acs.org/jacsau
Article

Structural Plasticity within 3-Hydroxy-3-Methylglutaryl Synthases Catalyzing the First Step of β-Branching in Polyketide Biosynthesis Underpins a Dynamic Mechanism of Substrate Accommodation

Sabrina Collin,* Kira J. Weissman,* and Arnaud Gruez*

Cite This: JACS Au 2024, 4, 3833−3847
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Understanding how enzymes have been repurposed by evolution to carry out new functions is a key goal of mechanistic enzymology. In this study we aimed to identify the adaptations required to allow the 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase (HMGCS) enzymes of primary isoprenoid assembly to function in specialized polyketide biosynthetic pathways, where they initiate β-branching. This role notably necessitates that the HMG synthases (HMGSs) act on substrates tethered to non-catalytic acyl carrier protein (ACP) domains instead of coenzyme A, and accommodation of substantially larger chains within the active sites. Here, we show using a combination of X-ray crystallography and small-angle X-ray scattering, that a model HMGS from the virginiamycin system exhibits markedly increased flexibility relative to its characterized HMGCS counterparts. This mobility encompasses multiple secondary structural elements that define the dimensions and chemical nature of the active site, as well the catalytic residues themselves. This result was unexpected given the well-ordered character of the HMGS within the context of an HMGS/ACP complex, but analysis by synchrotron radiation circular dichroism demonstrates that this interaction leads to increased HMGS folding. This flexible to more rigid transition is notably not accounted for by AlphaFold2, which yielded a structural model incompatible with binding of the native substrates. Taken together, these results illustrate the continued necessity of an integrative structural biology approach combining crystallographic and solution-phase data for elucidating the mechanisms underlying enzyme remodeling, information which can inform strategies to replicate such evolution effectively in the laboratory.



KEYWORDS: *3-hydroxy-3-methylglutaryl-CoA synthase, crystal structure, conformational flexibility, substrate specificity, polyketide biosynthesis*

INTRODUCTION

A central aim of modern enzymology is to elucidate how enzymes have evolved to acquire new specificities and/or activities, as these insights can inform experimental attempts to rationally expand and improve function.[1] The basis for such evolution is the latent promiscuity of enzymes, which has been argued to result from their intrinsic conformational diversity.[2] Via repeated cycles of mutagenesis coupled with selection, secondary specificities and activities can be enhanced, generating enzymes with wholly new functions. These mechanisms notably act when enzymes from primary metabolism are integrated into specialized metabolic pathways.[3] An apparent microbial example of this phenomenon is the presence of 3-hydroxy-3-methylglutaryl (HMG)-coenzyme A synthase (HMGCS)-like enzymes in a number of polyketide biosynthetic pathways.[4,5] HMGCSs, which belong to the thiolase superfamily of acyl-condensing enzymes,[6] catalyze the second step in the mevalonate-dependent (MVA) pathway of isoprenoid assembly in certain bacteria,[7] in which acetyl-CoA and acetoacetyl-CoA are condensed to yield HMG-CoA. In this reaction, acetylation of the HMGCS active site Cys is followed by base-catalyzed deprotonation to generate an enolate nucleophile. The enolate then participates in an aldol addition with acetoacetyl-CoA, followed by selective hydrolysis of the HMGCS thioester.[5,8]

In polyketide assembly,[9] HMG synthases (HMGSs) carry out the first of a series of reactions by which β-branching is introduced into the growing intermediates at specific positions (Figure 1).[4,5] Such enzymes have been identified principally in *trans*-acyltransferase (AT) polyketide synthase (PKS) pathways,[10] but also in *cis*-AT[11] and type II PKS systems.[12] While the

Received: June 4, 2024
Revised: July 11, 2024
Accepted: July 16, 2024
Published: September 23, 2024

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

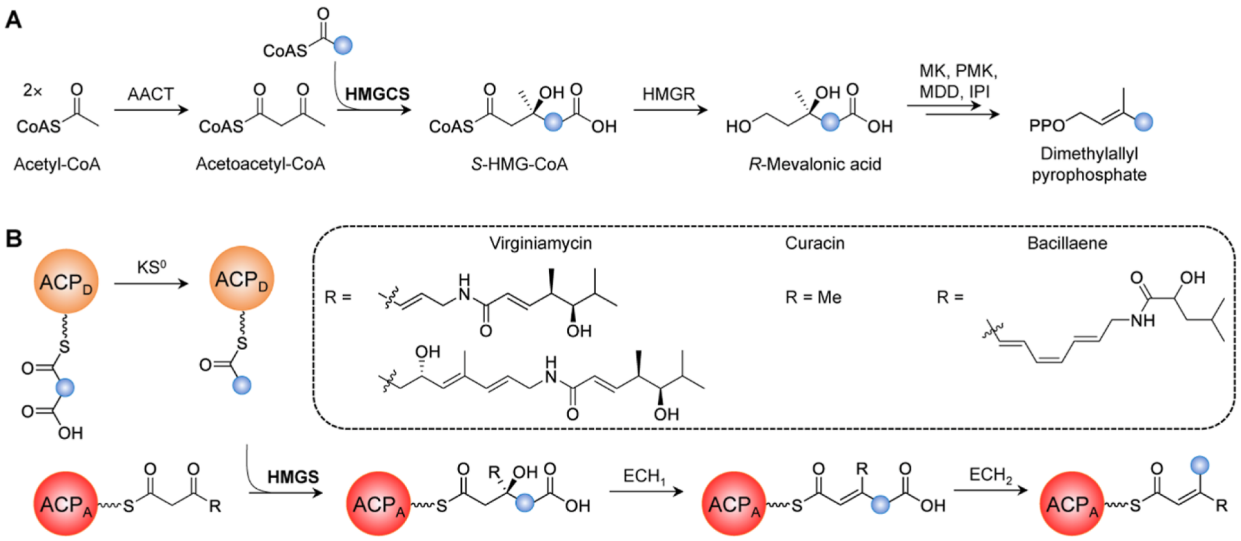


Figure 1. Comparison between assembly of dimethylallyl pyrophosphate, a primary precursor of isoprenoids, and the β-methylation reaction series of polyketide biosynthesis, both of which involve an HMG(C)S homologue.[4,5] (A) The isoprenoid pathway begins with synthesis of acetoacetyl-CoA by acetoacetyl-CoA thiolase (AACT) from two equivalents of acetyl-CoA. The acetoacetyl-CoA is then condensed with a third unit of acetate by a HMG-CoA synthase (HMGCS) to yield (*S*)-HMG-CoA, followed by reduction to (*R*)-mevalonic acid by HMG-CoA reductase (HMGR). Finally, the combined action of mevalonate-5-kinase (MK), phosphomevalonate kinase (PMK), mevalonate diphosphate decarboxylase (MDD) and isopentyl pyrophosphate isomerase (IPI), results in dimethylallyl pyrophosphate. (B) β-Methylation is initiated by decarboxylation of malonate to acetate catalyzed by a condensation-incompetent stand-alone ketosynthase (KS^0). This reaction takes place with the substrate tethered to an ACP. The ACP then becomes the donor (ACP$_D$) of the acetate nucleophile which is used by the HMGS to attack the β-keto group of a polyketide chain attached to an acceptor ACP (ACP$_A$). The resulting HMG-like intermediate is successively dehydrated and decarboxylated by two enoyl-CoA homologues to afford the β-methyl product (the blue dot indicates the origin of the methyl group). The inset shows the specific substrates recognized by the virginiamycin M, Cur and bacillaene HMGS homologues.

basic aldol chemistry is unchanged, the presumed assimilation of HMGCS into polyketide biosynthesis would have required modifications to several key aspects of the reaction. (Note: on the basis that the number of identified HMGCSs in the NCBI far exceeds that of HMGSs, the HMGCSs or their ancestors are considered to be the evolutionary precursors of the HMGSs). Notably, unlike HMGCSs, HMGSs do not act on substrates activated as their CoAs,[13,14] but tethered as thioesters to the phosphopantetheine (Ppant) prosthetic group of acyl carrier proteins (ACPs). More specifically, the nucleophile acetate is delivered by a donor ACP (ACP$_D$), and attacks an electrophilic substrate attached to either a discrete or a multienzyme-integrated acceptor ACP (ACP$_A$) (Figure 1). HMGSs are able to distinguish between ACP$_D$ and ACP$_A$, and, albeit imperfectly,[15] among multiple potential ACP acceptors within a given pathway—a critical feature, as each chain extension cycle gives rise to a potential ACP-linked substrate.[8] A second adaptation concerns the specificity for the acyl chains. HMGCSs only accept acetyl-CoA and acetoacetyl-CoA as substrates, while HMGSs can employ both acetate and propionate-derived nucleophiles in the aldol reaction,[5] as well as recognize acceptor chains of highly diverse functionality and chain length, including polyketides and polyketide−polypeptide hybrids.[10]

The basis for ACP recognition by HMGS was previously addressed for the curacin (Cur) *cis*-AT PKS of the marine cyanobacterium *Lyngbya majuscula*.[8] Initially, analysis of HMGS (CurD) activity in vitro demonstrated that an ACP$_A$ excised from subunit CurA could not replace cognate ACP$_D$ (CurB) as the acetyl donor. In contrast, acetoacetyl-ACP$_D$ could serve as a surrogate for acetoacetyl-ACP$_A$, but with 3-fold lower efficiency. Although no structure of a complex between CurD and ACP$_A$ was obtained, crystal structures were successfully solved of isolated homodimeric CurD (PDB 5KP5) and the wild type or Cys114Ser mutant bound to multiple forms of CurB (*apo*, *holo* and acetyl) (PDB IDs: 5KP6, 5KP7 and 5KP8). These structures in combination with site-directed mutagenesis, revealed that binding between CurB and CurD arises principally from shape complementarity at the contact surfaces, with minimal contribution from charge/charge interactions. Notably, CurB (ACP$_D$) exhibits distinctive surface/structural features relative to Cur ACP$_A$ and other PKS ACPs, explaining the interaction specificity. These differences include an atypical positioning of helix α3 that results in a hydrophobic cleft complementary to a hydrophobic ridge on CurD formed by helix α8. Notably, in typical HMGCSs, helix α8 is polar, suggesting that the evolutionary trajectory from HMGCS to HMGS encompassed modifications to this region. The CurD−CurB structure also accounts for the inability of Cur ACP$_A$ to substitute for CurB, as Cur ACP$_A$ lacks an equivalent hydrophobic crevice.

Correspondingly, Cur ACP$_A$ was hypothesized to dock against a binding site on the HMGS surface distinct from that for CurB, with the specificity of interaction arising from the divergent surface features of ACP$_A$ (negative/neutral/positive surface potential) relative to other ACPs of the same system. However, small-angle X-ray scattering (SAXS) data acquired by us on a HMGS/ACP$_A$ complex from the virginiamycin M (Vir) system[15] fit closely with those calculated from the crystal structure of the CurD/acetyl-ACP$_D$ complex,[8] arguing instead for a shared docking site on the HMGS for the two ACPs. This proposal is also in line with the single binding site identified within HMGCSs for both acetyl-CoA and acetoacetyl-CoA.[16] In any case, resolving this question will require direct comparison of an HMGS bound to both its cognate ACP$_D$ and ACP$_A$.

Concerning acyl chain specificity, the active sites and catalytic residues of CurD and HMGCS are essentially identical (RMSD

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

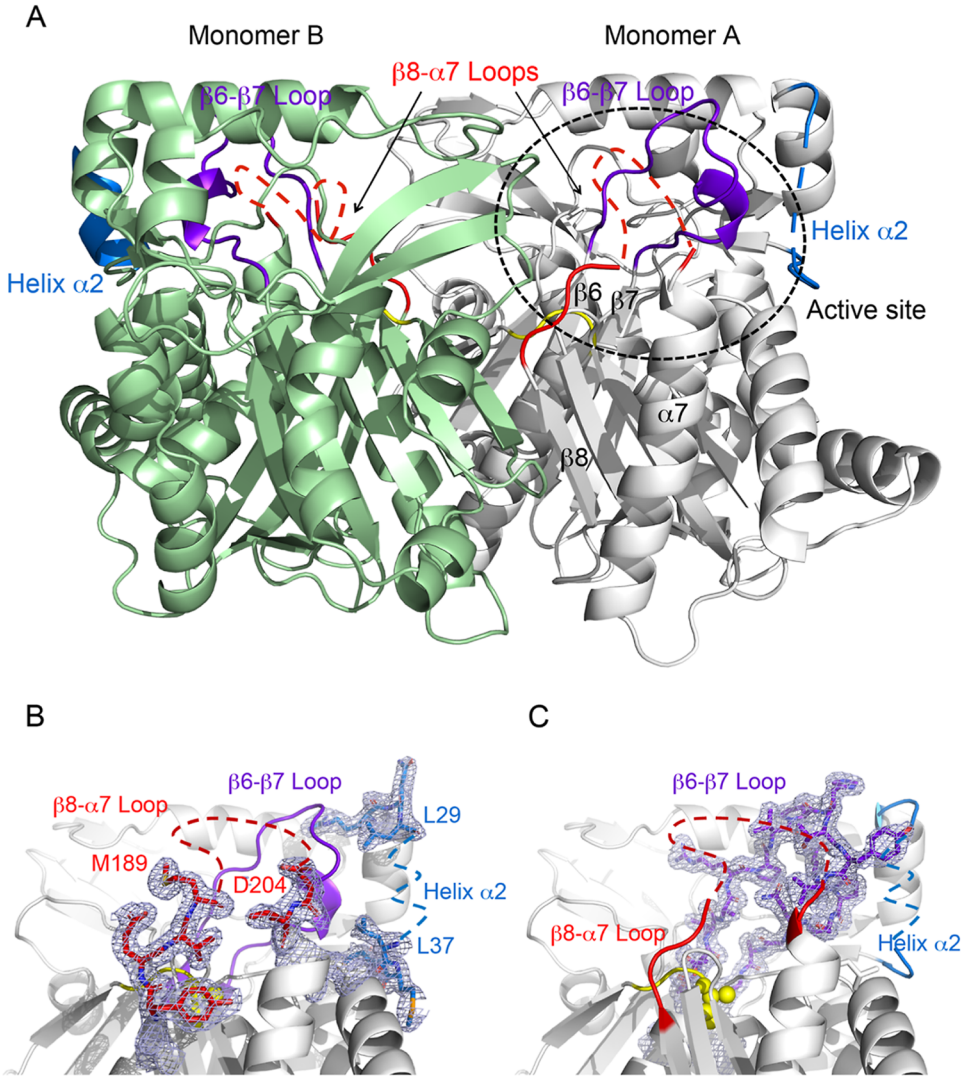


Figure 2. Crystal structure of the VirC^{4A} mutant (Cys114Ala/Gln334Ala/Arg335Ala/Arg338Ala) (PDB ID: 8S81). (A) Side view of the homodimer structure. The two VirC^{4A} monomers are illustrated in cartoon representation, with the two polypeptide chains colored in white (monomer A) and pale green (monomer B), respectively. For clarity, the location of the active site is indicated by a black ellipse in only one monomer. The disordered residues of the β8−α7 loops are represented by dashed red lines, and the β6−β7 loop which forms the other side of the active site is colored in purple. The residues corresponding to helices α2 (which are disordered in monomer A (dashes)) are colored in marine blue. The loop α5−β6 containing the catalytic residue is colored in yellow. (B) Zoom into the active site of monomer A showing the 2Fo−Fc map contoured at 1σ around the disordered β8−α7 loop and helix α2 regions. Residues Asp190−Val203 of the β8−α7 loop (in red) are missing from the 2Fo−Fc map contoured at 1σ. To illustrate this region, the residues directly upstream and downstream the loop are represented in red sticks. In monomer A, the disordered helix α2 (residues Asp30−Asn36) is represented with blue dashed lines, with the 2Fo−Fc map contoured at 1σ in the region of Arg27−Ala40 (residues represented as blue sticks). The catalytic residue is shown as a yellow ball. (C) Zoom into the active site showing the ordered β6−β7 loop (in purple sticks) with the 2Fo-Fc map contoured at 1σ around the residues Val140−Met165. For clarity, residues upstream and downstream of the β8−α7 loop and helix α2 are indicated in cartoon representation.

of 2.03 Å for 368 Cα atoms),[8] consistent with the fact that the enzymes act on the same two acyl substrates (acetate nucleophile and acetoacetate acceptor). Thus, the data obtained to date while providing insights into how HMGS interacts with its ACP partners, fail to illuminate the architectural modifications necessary to accommodate long, complex acceptor substrates.

In this work, we have addressed the molecular basis for this expanded substrate specificity using the HMGS VirC from the virginiamycin M pathway of *Streptomyces virginiae* as a model.[15] In this system, β-methylation occurs principally during the fifth chain extension cycle on an ACP-linked 18-membered linear chain incorporating both polyketide and amino acid building blocks, but also takes place to a lesser extent on the four-carbon longer intermediate generated by module 7 (Figure 1B).[15] In the case of module 5 which incorporates a tandem of ACP domains (ACP_{5a} and ACP_{5b}), only ACP_{5b} is recognized by the β-methylation cassette and thus serves as ACP_A.[15]

Using SAXS, we provide evidence for the high-flexibility of unliganded VirC relative to CurD and homologous HMGCS enzymes. This mobility, which encompasses multiple secondary structure motifs including α-helical and loop elements of the active site, fundamentally remodels the active site to accept the large assembly line substrates. Certain regions of the active site as well as ACP interface elements remain flexible in the VirC crystal structure, although the overall enzyme is more rigid than

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

in its solution state. Indeed, we show directly using synchrotron radiation circular dichroism (SRCD) that the interaction with its ACP_A partner leads to increased structuration of VirC, in accord with the SAXS data acquired on the equivalent complex. This fundamental change in plasticity is notably not accounted for in AlphaFold2[17]-based modeling, which leads to prediction of an HMGCS-like structure for VirC which is incompatible with the size of its native substrates. Taken together, these data highlight the continued need for multidisciplinary approaches for establishing detailed enzyme structure−function relationships.

RESULTS

Structural Characterization of VirC by X-ray Crystallography

We initially aimed to solve the crystal structure of native VirC expressed recombinantly in *Escherichia coli*, but were unable to obtain diffraction-quality crystals. We were also unable to crystallize VirC in the presence of ACP_A in either its *holo* or acetoacetyl forms. We therefore mutated VirC at three residue positions (Gln334Ala/Arg335Ala/Arg338Ala, all located on helix α12) previously described for the homologous HMGS CurD to render it more amenable to crystallization due to a presumed reduction in surface entropy.[8] We additionally introduced an active site Cys to Ala mutation (Cys114Ala) to disable the catalytic activity (the resulting quadruple mutant will hereafter be referred to as $VirC^{4A}$),[8] as we anticipated carrying out cocrystallization studies with analogs of both the donor substrate and the acylated ACP_A acceptor. We ultimately obtained five crystals of Se-$VirC^{4A}$ after several months of crystallization.

The $VirC^{4A}$ structure (Figure 2) was solved by molecular replacement at a resolution of 1.99 Å (PDB ID: 8S81) with MOLREP,[18] using a computed AlphaFold2[17] model of VirC as a search model. The asymmetric unit contains a dimer, with RMSD between monomers of 0.127 Å (312 Cα atoms) (Figure 2A). As expected, the protein fold belongs to the HMGCS family (EC:2.3.3.10), with each monomer composed of two structural regions. The upper region comprises a five-layered $\alpha\beta\alpha\beta\alpha$ core (each α corresponds to two α-helices, and β to a mixed β-sheet), while the lower region includes a three-stranded β-sheet, two two-stranded β-sheets, and three α-helices.

Despite their overall common fold and the shared absence from the electron density maps of loop β8−α7 (residues 190−203), the two monomers exhibit moderate differences in terms of flexibility. Notably, residues 30−36 that comprise helix α2 are absent from the electron density 2Fo−Fc map contoured at 1σ in monomer A, while the same α-helix is well-defined in monomer B. As revealed by the CurD−acetyl-ACP_D structure (PDB ID: 5KP8), the HMGS helix α2 plays several important roles in the active complexes. Notably, Arg33 forms a salt-bridge with the phosphate moiety of the Ppant cofactor, an interaction which is conserved with HMGCS-CoA complexes,[8] while the remainder of the α-helix contributes to the CurD/ACP_D interprotein interface. Second, the main chain NH of the well-conserved Asp30 interacts with the hydroxyl side chain of Tyr153 that is located in loop β6−β7 (residues 145−161), contributing to the loop positioning, and the Tyr in turn helps to correctly orient Arg33 via a cation−π interaction. In contrast, given the disordered character of helix α2 in monomer A of $VirC^{4A}$, it is unlikely to contribute to maintaining the conformation of the β6−β7 loop (Figure 2B). Despite this apparent difference in function, the B-factor of helix α2 of CurD is not substantially lower in the ACP complexes than in the isolated enzyme (PDB ID: 5KP5), and thus it too possesses intrinsic flexibility.

Analysis of the VirC Active Site

Catalysis by HMGCS and HMGS classically depends on a triad of active site residues, Glu82, Cys114 and His250 (VirC numbering; GenBank: BAF50725.1), as well as an oxyanion hole (hydroxyl group of Ser318 and the backbone amide of Gly319 in VirC).[16] In the $VirC^{4A}$ structure, the Ala introduced in place of the catalytic Cys114 is located near the N-terminus of helix α6 in a sterically constrained region of the structure. Indeed, helix α6 is sandwiched between two β-sheets (β8−β10−β11 and β5−β4−β6−β7−β1), and blocked on one side by helix α11 and by the second monomer on the other. The steric constraints at the active site also result from the cis conformation of Cys230 located in the β10−β11 loop.

The structure further reveals that the two loop regions β6−β7 (residues 145−161) and β8−α7 (residues 190−203) help to define the architecture and accessibility of the active site. The β6−β7 loop forms one wall of the site, with its conformation stabilized by a complex network of salt bridges and hydrogen bonds. Specifically, the main chain NH group of Ala146 interacts with the carbonyl group of Gly159 via a water-mediated hydrogen bond network. Glu155 forms a salt bridge to Arg27 that is located at the C-terminus of helix α1. In addition, the Oϵ2 atom of Glu155 forms a hydrogen bond with the hydroxyl group of Tyr357 and with the main chain NH group of His151 via a water molecule, while the hydroxyl group of Ser157 interacts with the side chain of Asn286 via another water molecule. Finally, the oxygen atoms of the amide groups of Glu155 and Gln158 interact with the amino group of Lys42, while the side chain of Gln158 hydrogen bonds to the hydroxyl function of Tyr288. Together, these interactions serve to constrain the main chain conformation of loop β6−β7 (Figure 3), which confers a unique shape to the active site relative to the homologues.

In terms of the β8−α7 loop, it contains several residues which are conserved within the protein family, but is disordered in both monomers of VirC. This substantial flexibility may explain the difficulties we experienced in obtaining diffracting crystals of $VirC^{4A}$. As the active site is exposed to solvent, the observed conformation could correspond to that capable of substrate binding.

Comparison with Homologues Reveals Structural Divergences

The program SuperPose in CCP4[19] was used to compare the $VirC^{4A}$ crystal structure with its nearest homologues identified using the DALI server.[20] This analysis revealed a shared overall fold, as expected from their mutual sequence identity (25−53%). The calculated RMSD values (Table 1) show that $VirC^{4A}$ most closely resembles the two PKS-associated HMGSs CurD[8] and PksG (unpublished), followed by the HMGCS enzymes of bacteria.

The prokaryotic enzymes are structurally conserved, but a difference is nonetheless apparent in terms of helix α2. In the majority of VirC structural homologues including CurD, helix α2 is systematically well-defined in the electron density maps, although as alluded to previously, the B-factors of the backbone atoms (residues 30−36) are higher than the mean B-factors of the overall structures. This rigidification relative to VirC is likely due to the presence of substrate or a substrate analog in the active sites, or that of the ACP in the case of CurD.[8] By contrast, as mentioned previously, this region is flexible in isolated VirC, a

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

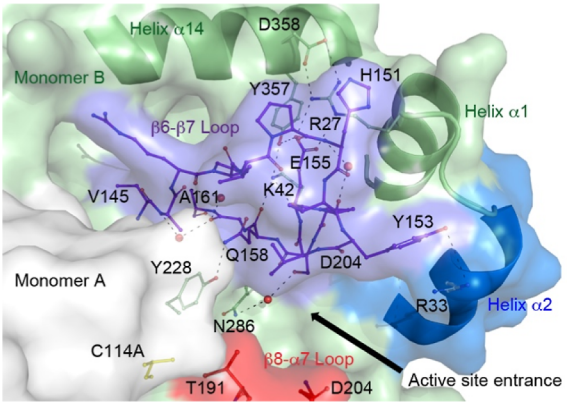


Figure 3. Active site of the VirC^{4A}. The active site is located at the interface between the two monomers. The mutated catalytic Cys114Ala is shown in stick form (yellow). The entrance to the active site is located near the C-terminal end of helix α2. Monomer A is shown in white surface representation, while the surface of monomer B is in green, with the regions of helix α2, and the β6−β7 and β8−α7 loops shown in blue, purple and red, respectively. The disordered β8−α7 loop (residues 190−203, red sticks) renders the active site accessible to solvent. The residues of the β6−β7 loop (Val145−Ala161) that form one side of the active site are shown in stick representation, and are well-ordered in the crystal structure. The β6−β7 loop is less conserved in term of sequence within the protein family, and therefore is likely to participate in substrate recognition. The loop is stabilized along its length via a network of hydrogen bonds (dashed lines) that includes water molecules (red spheres), and together with helix α1 and α14, contributes to structuring the active site. In addition, a cation−π interaction between Tyr153 and Arg33 helps to maintain the structure of helix α2 of monomer B, which likely contributes to the interface with both the ACP and the Ppant phosphate moiety of the acyl-ACP substrate.

Table 1. Comparison of the Structure of VirC^{4A} With Homologs Reveals Structural Divergence

PDB ID	Calc'd RMSD (number C_α atoms)	identity of the homologue
5KP5	0.914 Å (378)	HMGS CurD[8]
4YXT	1.128 Å (383)	HMGS PksG (unpublished)
1XPM	1.440 Å (351)	*Staphylococcus aureus*HMGCS[16]
5HWO	1.478 Å (357)	*Myxococcus xanthus* HMGCS[21]
2FA0	1.507 Å (365)	*Brassica juncea* HMGCS[22]
3SQZ	1.528 Å (344)	*Streptococcus mutans* HMGCS (unpublished)
2WYA	1.686 Å (368)	Human mitochondrial HMGCS[23]
2P8U	1.723 Å (367)	Human HMGCS[23]
6ET9	1.841 Å (195)	*Methanothermococcus thermolithotrophicus* HMGCS[24]

property it shares in common with uncomplexed PksG (PDB ID: 4YXT), an HMGS from the bacillaene PKS-nonribosomal peptide synthetase (NRPS). Notably, the PksG substrate, like that of VirC, is a polyketide-polypeptide hybrid of substantially higher functional complexity and size than acetoacetate (Figure 1B).[25] Thus, intrinsic mobility of this element appears to be a shared property of HMGSs which recognize large, non-acetoacetate substrates.

We next turned our attention to the β6−β7 loop (residues 145−161) which plays an important role in shaping the active site. Among the analyzed HMGCS homologues, the loop sequence is well conserved, and the main chain adopts almost identical conformations (Figure S1). In contrast, comparison of the region between VirC, PksG and CurD reveals visible structural differences in the organization of the main chain, as well as in side chain orientations (Figure S1). Furthermore, multiple sequence alignment of homologous HMGSs from a large number of modular PKS systems, shows this region to be variable and to include a number of insertions (Figure S2). The overall effect of these additions is to introduce structural plasticity into the β6−β7 loop, even when certain residue positions are well-conserved, consistent with a role in determining substrate specificity.

Concerning the β8−α7 loop (residues 190−203), in all of the structures with the exception of VirC^{4A}, the loop caps the active site, shielding it from solvent. Despite its strong sequence conservation among the homologues, it does not adopt the same structure. Specifically, while it is disordered in both monomers of VirC and shows elevated B-factors and weak electron density in the 2Fo−Fc map for CurD,[8] it is well-ordered in all of the other structures, with the main-chain of the loop adopting an essentially identical conformation (PDB ID: 5KP5, 4YXT, 1XPM, 5HWO, 2FA0, 3SQZ, 2WYA, 2P8U and 6ET9). In the case of PksG which is the most relevant homologue, three structures have been deposited in the PDB (4YXQ, 4YXT, 4YXV) which diffract at 2.1−2.75 Å resolution, and which belong to different space groups and unit cells. Comparison among the structures shows that the B-factors for the β8−α7 loop are higher than the average B-factors for each structure, with the differences between the B-factors arising from variable stabilization of the loop by the crystal packing. In any case, this analysis shows that the flexibility of the β8−α7 loop is likely to be a common feature of all PKS-integrated HMGSs.

Evidence of Induced Folding of VirC^{4A} by Specific Interaction with Partner

The strong agreement between the scattering curve calculated from the crystal structure of the CurD/ACP$_D$[8] complex and the SAXS data obtained on the VirC/ACP$_A$ complex,[15] provides evidence that the two HMGSs exhibit a similar, elevated degree of order. Thus, the crystallographic indicators for the intrinsic flexibility of several VirC elements were unexpected. In this context, we were motivated to probe directly whether VirC undergoes structural modification upon interaction with ACP$_A$, adopting a structure closer to that of well-folded CurD. Indeed, it was previously suggested that CurD itself undergoes induced folding upon interaction with ACP$_A$.[8] For this, we investigated the secondary structures of wild type VirC, VirC^{4A}, *holo*-ACP$_A$, *holo*-ACP$_{5a}$, VirC^{4A}/*holo*-ACP$_A$ and VirC^{4A}/*holo*-ACP$_{5a}$ by SRCD in the far-UV regions (180−260 nm).[26]

The SRCD spectra of wild type and VirC^{4A} both exhibit a maximum at 194 nm and two minima at 209 and 222 nm, while the ellipticity profiles are closely similar in the 180−260 nm range (Figures 4 and S3). These results confirmed that the four introduced mutations did not fundamentally alter the VirC structure, and therefore VirC^{4A} was used for the binding experiments.

The SRCD spectrum of *holo*-ACP$_A$ reveals a maximum at 193 nm and two minima at 208 and 222 nm (Figure 4A). The mathematical sum of the VirC^{4A} and ACP$_A$ spectra, which can be taken to represent a theoretical VirC/ACP$_A$ complex in which there is no change in the secondary structure of the partners upon interaction, exhibited a maximum at 194 nm and two minima at 210 and 222 nm. We then compared these data to the SRCD spectrum recorded on the VirC^{4A}/*holo*-ACP$_A$ complex, which revealed several important differences. Notably, the two

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

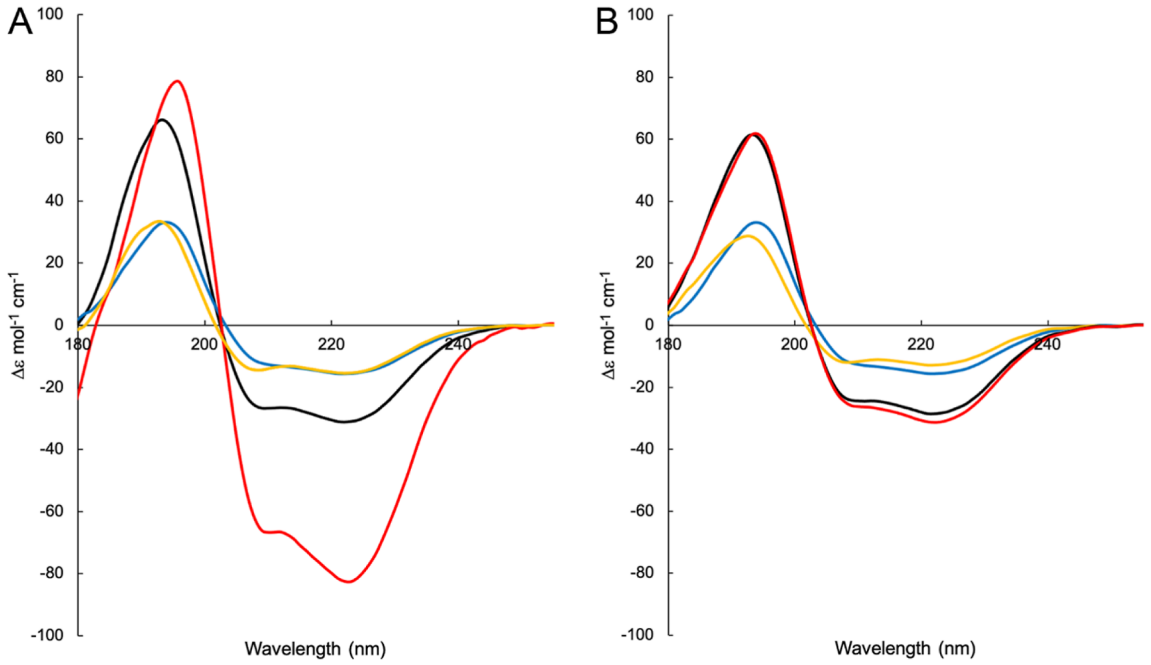


Figure 4. SRCD spectra of VirC^{4A} in the presence of potential ACP partners. (A) Spectra of proteins VirC^{4A} (in blue), *holo*-ACP$_A$ (in yellow) and VirC^{4A}/*holo*-ACP$_A$ complex (in red). The theoretical summation of the VirC^{4A} and *holo*-ACP$_A$ spectra is shown in black. The spectrum of the VirC^{4A}/*holo*-ACP$_A$ complex does not superimpose on the theoretical spectrum, revealing a substantial increase in secondary structure upon interaction between the two partners. (B) Spectra of proteins VirC^{4A} (in blue), *holo*-ACP$_{5a}$ (in yellow) and VirC^{4A}/*holo*-ACP$_{5a}$ complex (in red). The theoretical summation of the VirC^{4A} and *holo*-ACP$_{5a}$ spectra is shown in black. The spectrum of the combined VirC^{4A} and *holo*-ACP$_{5a}$ superimposes well on the theoretical summation, revealing no substantial change in secondary structure when VirC^{4A} is mixed with noninteracting *holo*-ACP$_{5a}$.[15]

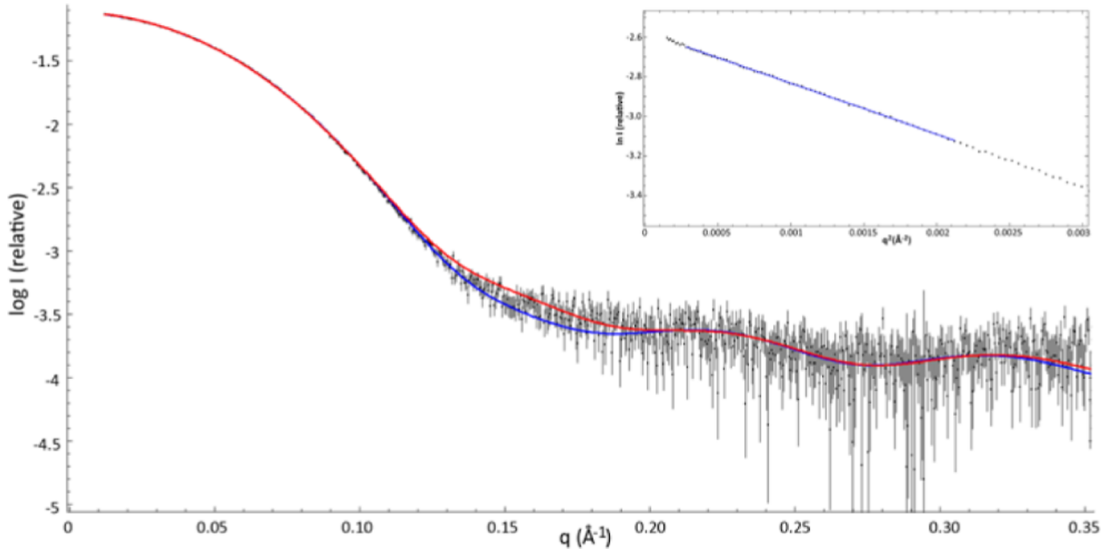


Figure 5. Fit between the experimental SAXS data, the VirC^{4A} crystal structure and the VirC^{4A} AlphaFold2[17] model. The experimental SAXS curve for VirC^{4A} is represented in black, with the Guinier plot inset. The scattering curve calculated from the crystal structure (PDB ID: 8S81) using CRYSOL[29] (in red) agrees well with the experimental SAXS curve (χ^2 = 1.88), but a discrepancy is nonetheless apparent in the range of 0.12−0.2 Å^{-1}. The theoretical SAXS curve (in blue) calculated using CRYSOL from an AlphaFold2 model of dimeric VirC^{4A}, yielded a χ^2 of 1.42 relative to the experimental data, and showed a major divergence from the curve calculated from the crystal structure in the range of 0.12−0.2 Å^{-1}.

spectra do not superimpose, as the maximum shifts from 194 to 196 nm, one of the two minima shifts from 222 to 223 nm, and the ellipticity in the spectrum of the VirC^{4A}/*holo*-ACP$_A$ complex is substantially increased. As a control for the observed spectral changes, we acquired the spectrum of VirC^{4A} in the presence of *holo*-ACP$_{5a}$, as we have shown previously that this ACP is not recognized by VirC.[15] Indeed, in this case, the signal from a mixture of VirC^{4A} and *holo*-ACP$_{5a}$ was essentially identical to that generated by summing the ellipticity data acquired on separate VirC^{4A} and *holo*-ACP$_{5a}$ (Figure 4B). Thus, the spectral changes observed upon combining VirC^{4A} and ACP$_A$ reflect an increase in secondary structure, which is likely limited to VirC given the defined character of the ACP fold.[15] Globally these data show that even in the absence of substrate attached to ACP$_A$, induced structuration occurs upon interaction of VirC^{4A}

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

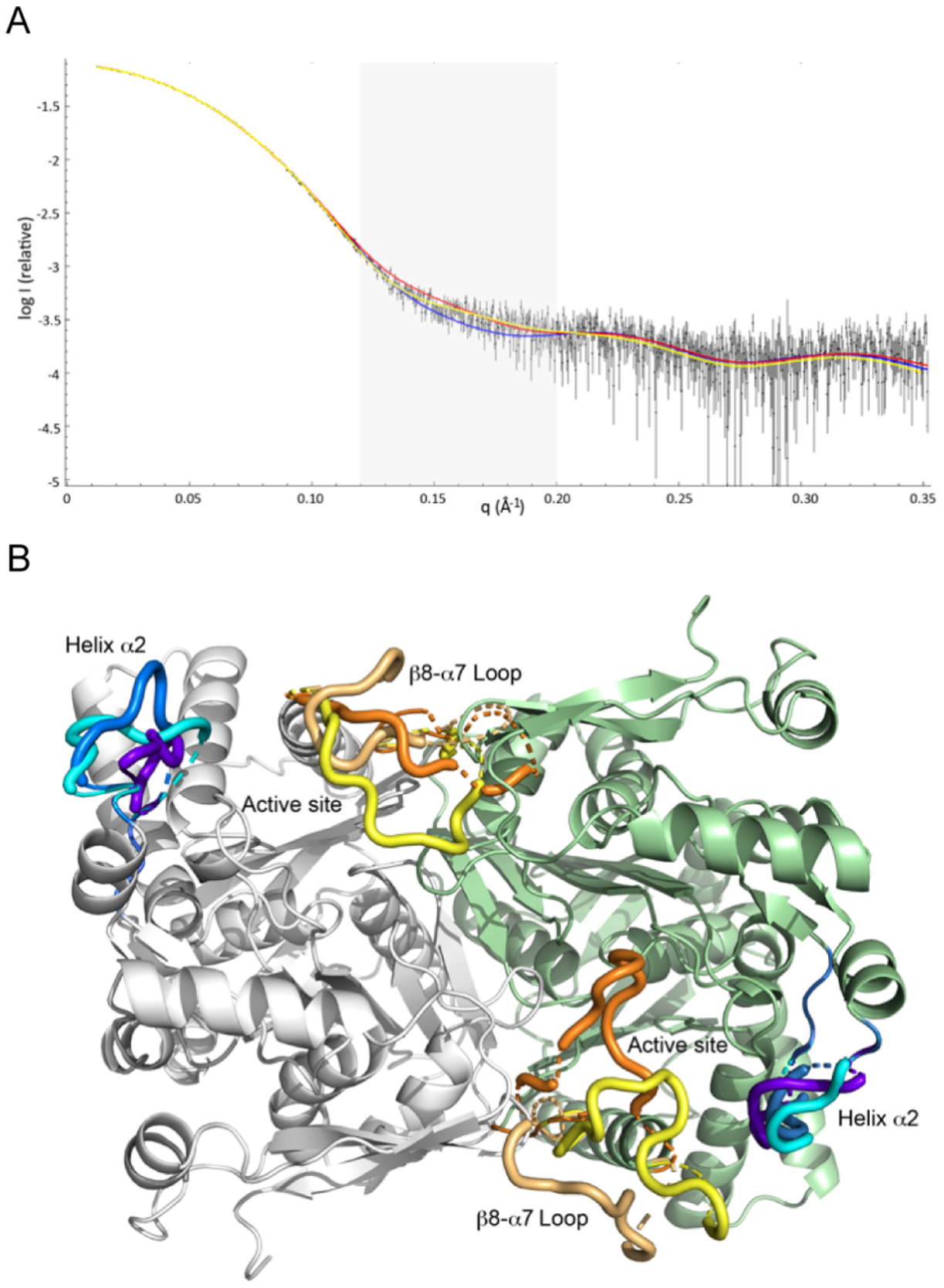


Figure 6. Reconstruction of the missing $\beta 8-\alpha 7$ loop using the Ensemble Optimisation Method (EOM 3.0),[31,32] and comparison with the VirC^{4A} crystal structure and AlphaFold2[17]-derived VirC^{4A} model. (A) Comparison of the theoretical scattering curve (in yellow) calculated from the EOM model that contains the $\beta 8-\alpha 7$ loop and helix $\alpha 2$ which are missing from the VirC^{4A} crystal structure, with those calculated from the VirC^{4A} crystal structure (in red) (χ^2 = 1.88) and the AlphaFold2 model of VirC^{4A} (in blue) (χ^2 = 1.42). The fit between the experimental and the reconstructed models using EOM was evidently improved in the range of 0.12−0.2 Å^{-1} (light gray shading), as was the overall fit (χ^2 = 1.11). (B) The three models of open forms of VirC^{4A} that contain the $\beta 8-\alpha 7$ loop which were calculated using EOM 3.0 are consistent with the SAXS data. All of the models reveal

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

Figure 6. continued

minor differences in the location of the helix α2 (shades of blue/violet). By contrast, the models exhibit major differences in terms of the conformation of the β8−α7 loop (shades of orange/yellow). The three EOM models further show that the β8−α7 loop does not cap the active site, but adopts multiple open conformations. The amplitude of the loop movement varies from one monomer to the other, and is also model-dependent. Our interpretation of these results is that the β8−α7 loop is likely to adopt a continuum of open conformations in solution.

with *holo*-ACP$_A$, consistent with the more extensive folding of VirC suggested by the SAXS analysis.[15]

Confirmation of VirC Flexibility Using SAXS

To further explore VirC^{4A} flexibility in solution, we analyzed the enzyme in solution by SAXS. This experiment yielded an R_g of 27.8 Å, and a molecular weight of 83.12 kDa calculated using Bayesian Interference in PRIMUS,[27] which is consistent with a dimeric state of the protein. The distance distribution function calculated with GNOM[28] yielded a D_{max} of 81.24 Å.

Comparison via CRYSOL[29] of the acquired SAXS data with that calculated from the VirC^{4A} crystal structure as well as from a VirC^{4A} model generated using AlphaFold2,[17,30] gave relatively good agreement[27] (χ^2 = 1.88 and 1.42, respectively) (Figure 5). Nonetheless, some divergence was noted in the range of 0.12−0.2 Å^{-1}, which we hypothesized originated in the disordered elements observed in the crystal structure—notably the absent helix α2 and the loop regions. (Note: consistent with the SRCD analysis, comparison of AlphaFold2 models of VirC and VirC^{4A} shows them to be structurally identical with an RMSD of 0.114 Å (for 716 Cα) including the β8−α7 loop, and thus conclusions based on the VirC^{4A} model are also valid for the VirC model).

To improve the fits, we next modeled the flexible β8−α7 loop using the Ensemble Optimization Method (EOM 3.0).[31,32] This approach yielded three distinct open-form models with an enhanced fit to the SAXS data (χ^2 = 1.11), and better agreement between the curves, particularly within the 0.12−0.2 Å^{-1} range (Figure 6A). In the three models, the conformations of the solvent-accessible β8−α7 loop notably differ in the amplitude of their movement, although none of the displacements results in closing of the active site (Figure 6B). The results are fully consistent with the inherent flexibility of the β8−α7 loop which may adopt a continuum of positions in solution, and support the idea that movement of the loop is implicated in substrate selection and accommodation within the active site. The three models also identify the α1−α2 loop and helix α2 as highly mobile elements, consistent with the absence of helix α2 from the electron density maps of the crystal structure.

We next extended this analysis with SREFLEX[33] to explore the contribution of movements beyond the β8−α7 loop to the conformations adopted by VirC^{4A} in solution. For this, five open-form models were computed using normal mode analysis (NMA) from the SAXS data, yielding further improved χ^2 ranging from 1.07 to 1.09 with RMSD of 2.13−3.94 Å compared with the AlphaFold2 model (312 Cα atoms) (Figure 7A). Relative to the scattering curves calculated from the crystal structure and the AlphaFold2 model, the fit between the experimental SAXS curve and the models derived from SREFLEX was further enhanced in the 0.12−0.2 Å^{-1} range. Comparison of these two sets of structures shows that the highest amplitude movements occur at the surface of the protein, while the core structure located at the dimer interface exhibits only minor displacements (Figure 7B).

This analysis further confirmed that the β8−α7 loop undergoes the largest displacement (by as much as 47 Å) between the most open and closed models. This substantial movement is allowed by the hinge region of the loop that encompasses residues Asp190 and Val203 that are located at 9 and 11 Å (Cα-to-Cα) respectively from the catalytic Cys114 (Ala in VirC^{4A}). Additional substantial motions were identified involving loops and secondary structures in and around the active site. For example, the β5−α6 loop bearing the catalytic Cys114 can travel by up to 6.7 Å, while the β-sheet formed by β6−β7−β1−β8 undergoes 4.7 Å of lateral displacement. Furthermore, the β6−β7 loop moves by a maximum of 9.3 Å (measured for His151) between the closed and open forms of the enzyme.

This motion is coupled to a 11.9 Å movement by helices α1 and α2 that are located at the entrance of the active site where they are positioned to interact with the ACPs. Interestingly, although the overall confidence scores (pLDDT) of the AlphaFold2[30] models for VirC and VirC^{4A} are superior to 90, they are substantially lower for precisely the identified most flexible regions (e.g., 90 > pLDDT > 70 for helix α2, and both the β6−β7 and β8−α7 loops). Overall, the SAXS data conclusively demonstrate the highly dynamic nature of unliganded VirC in solution, behavior which was suggested by the disordered regions observed in the crystal structure.

Implication of HMGS Flexibility for Substrate Binding

To take the analysis further, we wished to investigate substrate binding into the VirC active site. However, the crystal structure was unsuitable for this analysis as it lacks multiple elements (e.g., the β8−α7 loop and helix α2) necessary to determine both the full dimensions and chemical character of the active site necessary for docking studies. We therefore based our analysis on the VirC model generated with AlphaFold2,[30] which although inadequate to fully capture the flexibility of the enzyme in solution, was more representative of the crystal structure in the core region which includes the active site. Notably, the active site in the model is closed by the β8−α7 loop, and strongly resembles those of CurD[8] and its HMGCS homologues,[16,21−24] despite the modeling being carried out in the absence of a specified template.[17]

One important determinant of the active site volume in the model is the position of two conserved residues Arg193 and Glu199 located on the β8-α7 loop, whose side chains form a salt bridge and point toward the protein interior (Figure 8). Using CASTp 3.0,[34] we calculated the volume of the active site to be 427.8 Å^3, with a potential interaction surface of 580.2 Å^2. The total length of the cavity is 19 Å, extending from the protein surface under the β8−α7 loop to the catalytic Cys at the base of the active site—in other words, far short of the maximum 40 Å required to house the module 5 and 7 substrates (Figure 8A). Therefore, binding of the VirC substrates into the active site would minimally necessitate a structural rearrangement of Arg193 and Glu199.

As a rough test of this hypothesis, we computationally mutated both the Arg193 and Glu199 to Ala, and remodeled the enzyme using AlphaFold2.[17] The resulting cavity was 44 Å in length, with a volume of 666.2 Å^3 and surface interaction of 1071.9 Å^2, dimensions large enough to accommodate both Vir

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

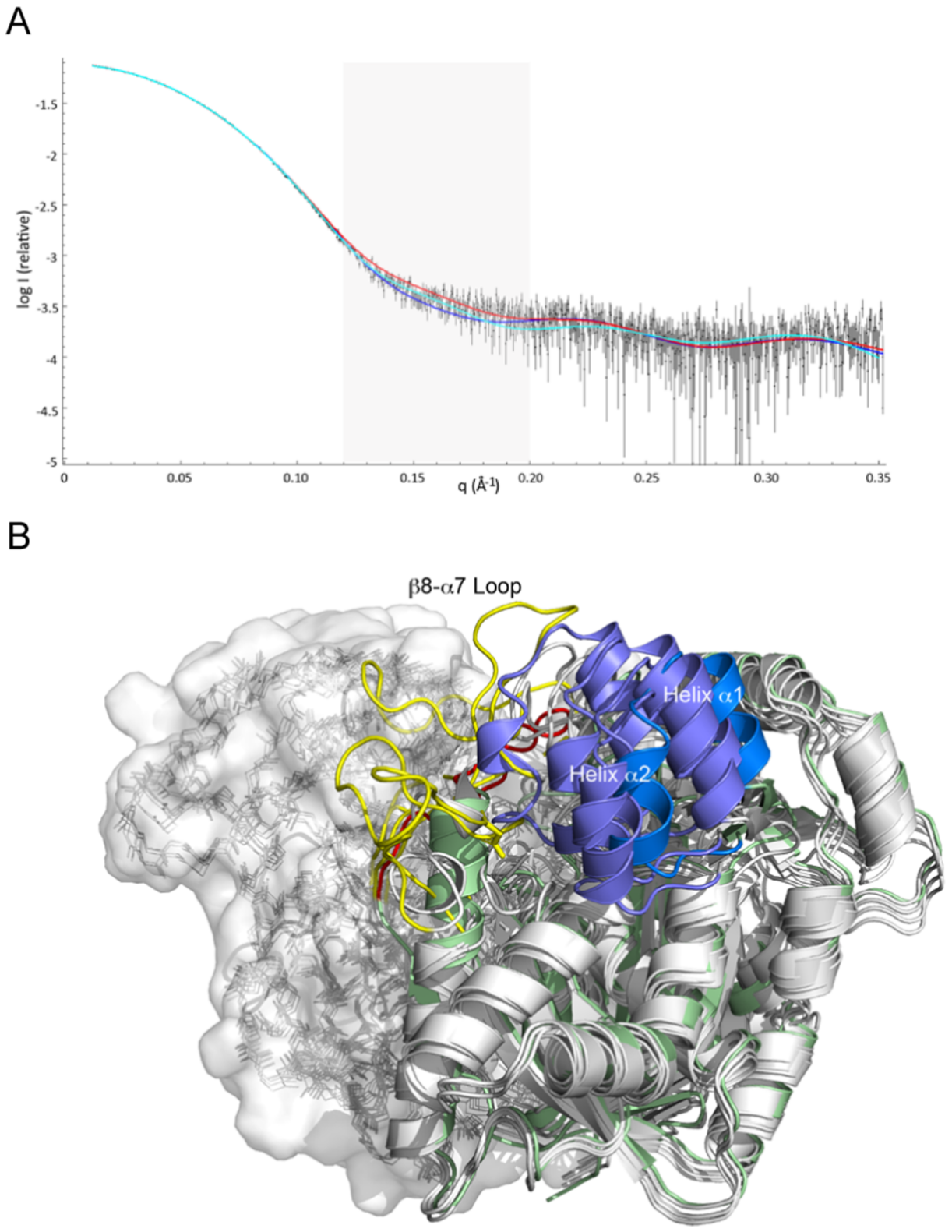


Figure 7. Overall flexibility of VirC^{4A} as judged using NMA[33] coupled with the SAXS data. (A) Comparison of the SAXS experimental data (in black) with the scattering curves calculated from the SREFLEX[33] open-form models leads to an improved fit (χ^2 = 1.07−1.09) in the range of 0.12−0.2 Å^{-1} (light gray shading), in comparison to that calculated from the crystal structure (in red) (χ^2 = 1.88) and the AlphaFold2[17] model (in blue) (χ^2 = 1.42). (B) Comparison of the closed state modeled by AlphaFold2 (in pale green) with the four open state models (in white) derived from SREFLEX. The closed conformation of the β8−α7 loop observed in the AlphaFold2 model is shown in red, while the open conformations found with SREFLEX are represented in yellow. Comparison of the models also reveals that the subdomain containing helices α1 and α2 (shown in marine blue for the AlphaFold2 model and in violet for the SREFLEX models) is mobile. Overall, NMA combined with SAXS data reveals multiple flexible regions in VirC^{4A}, implicating them in both accommodation of the substrate in the active site and binding of the ACP partner.

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

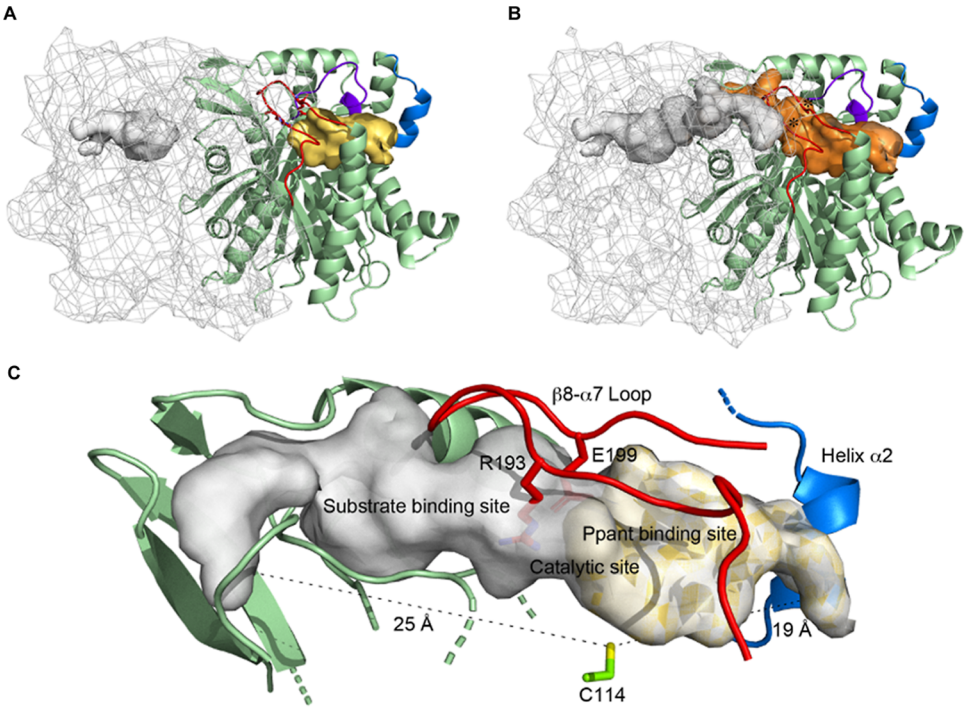


Figure 8. Overall shape and volume of the VirC active site. (A) The cavity in the closed AlphaFold2[17] model has been delineated using CASTp 3.0.[34] Monomer A is represented in mesh with the active site cavity shown in surface representation (in white), while monomer B is shown in cartoon representation (in pale green) along with its active site cavity (in gold). Helix α2, and the β6−β7 and β8−α7 loops are colored in marine blue, purple and red, respectively. The two residues Arg193 and Glu199 (shown as red sticks) form a lid over the active site in the closed conformation. The volume of the active site in this conformation is approximately 427.8 Å^3, with a potential interaction surface of 580.2 Å^2. The measured 19 Å depth of the cavity would not accommodate binding of either of the two identified acyl-Ppant substrates. (B) To mimic the minimal conformational motion necessary for substrate binding, residues Arg193 and Glu199 (indicated with *) have been mutated in silico to Ala. The resulting active site is relatively linear, with an increased volume of 666.2 Å^3 and approximately twice the surface area (1071.9 Å^2). Furthermore, the cavity depth increases to 44 Å, a distance compatible with substrate binding. (C) Detailed view of the enlarged active site. Secondary structure elements surrounding the active site, as well as the β8−α7 loop and helix α2, are colored in pale green, red and marine blue, respectively, and the catalytic Cys114 in green. Cavity measurements are shown using dashed lines.

intermediates (Figure 8B,C). Thus, movements of the β8−α7 loop directly influence the active site cavity volume as well as its surface features, and may additionally contribute to both substrate/ACP binding and catalysis. Contrary to the crystal structure of VirC, in the crystal structure of CurD (PDB ID: 5KP5[8]) and the three structures of PksG (PDB IDs: 4YXQ, 4YXT, 4YXV), the β8−α7 loops are ordered but the β6−β7 loops are partially disordered and incorporate 13−15 amino acid insertions. In these cases, we propose that both flexible β6−β7 and β8−α7 loops contribute to establishing an active site conformation appropriate for substrate binding.

Analysis of the HMGCS/Thiolase/DUF35 Complex Provides Additional Evidence for HMGCS Evolvability

The functional plasticity of the HMGCSs is emphasized by their ability to participate in distinct protein−protein interactions in the context of isoprenoid biosynthesis.[24] As demonstrated for archaea (PDB ID: 6ET9), two dimers of HMGCS form a macrocomplex with two dimers of the first enzyme in the isoprenoid pathway, acetoacetyl-CoA thiolase, an interaction which is mediated by four monomers of a scaffolding protein (DUF35).[24] Complex formation notably gives rise to a binding site for CoA shared by both the thiolase and HMGCS, which allows for coupling of the endergonic thiolase-catalyzed reaction to the exergonic chemistry of HMG-CoA formation. Such an architecture is also likely to be present in certain bacteria, including *Streptomyces albus* DSM 41398, where a DUF homologue is fused directly C-terminal to the HMGCS.[24] On the other hand, BlastP[35,36] analysis of the genome of the *S. virginiae* type strain (taxonomy ID: 1961) does not reveal a convincing DUF35 homologue despite the presence of a full isoprenoid cluster (GenBank: ADQ43372, ADQ43373, ADQ43374, ADQ43375, ADQ43376, ADQ43377, ADQ43378), and thus there is no indication for formation of an analogous complex in this species.

In the archaeal HMGCS/thiolase/DUF35 complex, the HMGCS secondary structure elements that interact with DUF35 include three α-helices (α3, α5 and α14), as well as loops β1−β2 and α12−α13 (VirC nomenclature). Relative to the archaeal HMGCS and to the HMGCS from the same *S. virginiae* strain, VirC incorporates a C-terminal extension (47 and 20 residues, respectively), which encompasses helix α15 and β-strands β14 and β15. Indeed, the archaeal HMGCS is truncated by some 50−100 residues compared to its eukaryotic and bacterial homologues.[24] Interestingly, in VirC, these supplemental structural elements form a mixed $\alpha\beta$ subdomain which occupies precisely the place of the DUF35 in the archaeal complex (Figure S4). Based on our SAXS analysis (Figure 6), this subdomain does not exhibit substantial flexibility, and thus is likely to stably occlude this potential interaction interface in solution. Furthermore, comparative sequence and structure-based analysis of this C-terminal subdomain shows it to be a unique feature of the PKS-integrated HMGS homologues VirC, PksG and CurD (Figures S2 and S4).

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

DISCUSSION

The functional divergence of enzymes over the course of evolution is argued to arise from their conformational heterogeneity.[2] More precisely, any given enzyme adopts multiple conformational substates in addition to its dominant conformation that may be specific for different substrates or catalyze alternative chemistry.[2,37] Thus, following duplication of the coding gene, such secondary characteristics can be honed through iterative rounds of mutation and selection that stabilize the promiscuous conformations and optimize the features of their active sites.[38]

This theory is attractive for explaining how the HMGCSs of isoprenoid biosynthesis were integrated into polyketide pathways, particularly as both types of system are present in certain bacteria (notably including *Streptomyces*) consistent with gene duplication events. This transition necessitated both adaptation of the specificity of HMGCS for the activating groups of the nucleophilic acetate and the acceptor electrophile (ACPs instead of coenzyme A), and the chemical nature of the acceptor substrate (chains of substantially increased length and structural diversity relative to acetoacetate). Our analysis of VirC, a model HMGS from the virginiamycin M *trans*-AT PKS in *S. virginiae*, shows it to be a highly dynamic enzyme. This conformational flexibility not only encompasses multiple loops ($\alpha1-\alpha2$, $\beta5-\alpha6$, $\beta6-\beta7$ and $\beta8-\alpha7$), but α-helices 1 and 2, and a β-sheet ($\beta6-\beta7-\beta1-\beta8$). Comparative analysis of the VirC crystal structure and the enzyme in solution supports the idea that the plasticity of these elements, and in particular helix $\alpha2$ and the $\beta6-\beta7$ loop, is critical for the ability of VirC to accommodate the structurally complex acceptor substrate. This feature is notably shared with a VirC homologue PksG (PDB ID: 4YXT, unpublished), which also acts on a long hybrid polyketide-peptide chain.

Concerning the $\beta8-\alpha7$ loop, in all HMGCS homologues characterized to date [including one solved in its *apo* form (PDB ID: 6ET9)], it caps the active site. Although the same loops lie across the active sites in the PksG and CurD structures, they show higher than average flexibility relative to the rest of the proteins, while that in VirC extends into solution where it likely adopts a continuum of open conformations. The transition from the catalytically inactive open conformation to a catalytically closed conformation likely takes place upon substrate binding (i.e., ligand-gated loop motion[39]). In any event, taken together, these data are consistent with a critical role for this flexible lid loop in enlarging and reshaping the ancestral HMGCS active site to accommodate polyketide substrates.[39]

Mobile elements are similarly involved in allowing the HMGSs to interact effectively with ACP domains instead of CoA, as helix $\alpha2$ was previously implicated in forming the interface between CurD and ACP_D.[8] Such regions may also contribute to catalysis, as the catalytic Cys114 is located on the flexible $\beta5-\alpha6$ loop. Taken together with data obtained previously on VirD and VirE,[15] these results demonstrate that the structures of three enzymes of the β-methylation cassette adapt to their substrates, and in this way act as specificity gatekeepers that regulate the reaction series. Indeed, we provide direct evidence using SRCD that the VirC structure is substantially ordered by specific interaction with its partner ACP_A.

Our data also suggest a mechanism by which HMGSs avoid formation of higher-order oligomers that would presumably interfere with their ability to act in trans on substrates tethered to PKS assembly lines. Specifically, VirC possesses a C-terminal extension relative to its HMGCS homologues. This structural element sterically occludes a region in the enzyme that in the HMGCSs serves as a platform for binding DUF35 scaffolding protein.[24] In this way, the HMGSs maintain the homodimeric state which is evidently crucial for their interactions with their two acyl-ACP substrates. Thus, in addition to exploiting the intrinsic conformational plasticity of the HMGCSs to expand their substrate specificity, Nature has modified the C-terminal region both by truncation and extension to regulate interactions with potential partners.

While the HMGSs are highly plastic enzymes, this feature was not previously noted for the putative ancestral HMGCSs, but could explain the relative paucity of such structures in the PDB. This behavior may have been suppressed in the solved HMGCS crystal structures for several reasons, including the almost uniform presence of substrates/substrate analogs (CoA or HMG-CoA) in the active sites, as well as the crystal packing, both of which would have had stabilizing effects on the enzymes. Indeed, the strongly dynamic character of unliganded VirC was also not entirely apparent from the crystal structure, but was revealed by analysis of its solution behavior by SAXS coupled with modeling of flexible elements using the ensemble optimization method[31,32] and SREFLEX.[33] We further show that the AlphaFold2[17,30] model of substrate-free VirC fails to fully represent both the crystal structure of the enzyme and its solution state. In particular, the model provides erroneous information concerning multiple key structural elements of the protein, including the flexible loops which define the active site. This current limitation emphasizes the need to characterize enzymes directly using an integrative structural biology approach.

Overall, the insights obtained here increase our understanding of functional adaptation within the thiolase superfamily, and can inform strategies by which enzyme activity is engineered productively in the laboratory.

METHODS

Bioinformatics Analysis

HMG(C)S sequences in the Protein Data Base (http://www.ncbi.nlm.nih.gov/protein) were identified using Blast.[35,36] Multiple sequence alignment was carried out using ClustalW[40] and the associated figure (Figure S2) created using ESPript.[41] We also identified the closest structural counterparts using the DALI[20] server. The CASTp 3.0[34] server was used to determine the size of the internal cavity of the AlphaFold2[17] models of VirC wild type and the VirC Arg193 and Glu199 in silico mutant.

Materials and DNA Manipulation

Biochemicals and media were purchased from VWR (glycerol, NaPi, NaCl, $MgSO_4$), BD (tryptone, yeast extract), Thermo Fischer Scientific (Tris), Euromedex (isopropyl β-D-1-thiogalactopyranoside, IPTG), and Sigma-Aldrich [imidazole, tris(2-carboxyethyl)phosphine hydrochloride (TCEP), CoASH]. Isolation of DNA fragments from agarose gel, purification of PCR products and extraction of plasmids were carried out using the NucleoSpin Gel and PCR Clean-up or NucleoSpin Plasmid DNA kits (Macherey Nagel). Standard PCR reactions were performed with Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific), and reactions were carried out on a Mastercycler Pro (Eppendorf). DNA sequencing was carried out by Eurofins Genomics (Köln, DE).

Gene Cloning and Site-Directed Mutagenesis

Wild type VirC, ACP_{5a} and ACP_A (ACP_{5b}[15]) were amplified directly from *S. virginiae* genomic DNA using forward and reverse primers incorporating *Bam*HI and *Hind*III restriction sites, respectively, and

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

were ligated into the corresponding sites of vector pBG-102 for ACP_A and ACP_{5a}, and pLM-302 for VirC. Vector pBG-102 codes for a His_6-SUMO tag and pLM-302 codes for a His_6-maltose binding protein (MBP) tag (Centre for Structural Biology, Vanderbilt University). Following cleavage of the tags, the proteins incorporated a non-native N-terminal Gly-Pro-Gly-Ser sequence. Mutations were introduced via site-directed mutagenesis using pLM-302_$VirC^{wt}$ as a template and the Phusion high-fidelity polymerase from Thermo Fisher Scientific, followed by digestion of the parental DNA by 1 μL of FastDigest *Dpn*I (Thermo Fischer Scientific). All primers used in this work are listed in Table S3. *E. coli* strain DH5α (Novagen) was used for cloning. Mutations were confirmed by DNA sequencing prior to protein expression.

Expression and Purification of Recombinant ACPs

The plasmids encoding ACP_A and ACP_{5a} were transformed into *E. coli* BL21(DE3) (Novagen) cells and grown at 37 °C in LB medium supplemented with 50 μg mL^{-1} kanamycin to an A_{600} of 0.8, and then IPTG added to a final concentration of 0.5 mM. Following incubation at 20 °C for 18 h, the cells were harvested by centrifugation at 3000*g* for 30 min at 4 °C, and cell pellets stored immediately at −80 °C.

Expression of Labeled Protein for X-ray Crystallography

Vir^{WT} and $VirC^{4A}$ were transformed into *E. coli* BL21(DE3) (Novagen) for producing seleniated proteins. After an overnight preculture at 37 °C, the cultures were grown in LB medium (yeast extract 10 g L^{-1}, tryptone 5 g L^{-1}, NaCl 10 g L^{-1}, adjusted to pH 7.0 with NaOH) supplemented with 50 μg mL^{-1} kanamycin. Seleniated proteins were produced in M9 minimal medium (50 mM Na_2HPO_4, 22 mM KH_2PO_4, 10 mM NaCl, 20 mM NH_4Cl, adjusted to pH 7.2 with NaOH) supplemented with 50 μg mL^{-1} kanamycin to an OD_{600} of 0.8, and then IPTG added to a final concentration of 0.5 mM. Autoclaved M9 medium was supplemented with 50 mg L^{-1} of thiamine and riboflavin, 4 g L^{-1} glucose, 100 mM $CaCl_2$, 2 mM $MgSO_4$, 40 mg L^{-1} selenomethionine, and 40 mg L^{-1} of the 19 amino acids, based on the methionine biosynthesis inhibition method.[46] Following incubation at 20 °C for 18 h, the cells were harvested by centrifugation at 3000*g* for 30 min at 4 °C, and cell pellets stored immediately at −80 °C.

Purification of Recombinant ACPs, VirC and $VirC^{4A}$

The cell pellets were resuspended in His-buffer (50 mM NaPi pH 7.5, 250 mM NaCl, 10% glycerol in the case of VirC and $VirC^{4A}$, and 30 mM Tris−HCl pH 8.5, 250 mM NaCl for the ACPs) containing 8 U mL^{-1} of benzonase (Merck) and 5 mM $MgSO_4$. The cells were lysed by sonication and clarified by centrifugation (35,000*g* for 40 min). Cell extracts were loaded onto a 5 mL HisTrap column (Cytiva) and washed with resuspension buffer supplemented with 20 mM imidazole. The supernatant was loaded onto a HisTrap 5 mL column equilibrated with His-buffer using an Akta Pure system (Cytiva). The proteins were eluted using a linear gradient of 0−50% His-elution buffer (50 mM NaPi pH 7.5, 250 mM NaCl, 300 mM imidazole for VirC and $VirC^{4A}$ or 20 mM Tris−HCl pH 8.5, 300 mM NaCl for the ACPs) over ten column volumes. The constructs were then incubated with His-tagged human rhinovirus 3C protease (1 mM) for 12−16 h at 4 °C to cleave off the MBP or SUMO tags. Proteins were then separated from the remaining His-tagged proteins by loading onto a HisTrap 5 mL column, followed by elution in resuspension buffer containing 20 mM imidazole. Eluted fractions found to contain protein of the correct molecular weight as judged by SDS-PAGE analysis were pooled, concentrated using an Amicon Ultracel-10 (Merck Millipore) by centrifugation at 4000*g*, and loaded onto a Superdex 75 16/60 column (Cytiva) equilibrated with 20 mM Tris−HCl pH 8.5, 300 mM NaCl, 5% glycerol. Following a concentration step, the purity of the labeled proteins was determined by SDS−PAGE and their concentrations were determined by NanoDrop (Thermo Scientific), with extinction coefficients calculated using the ExPASy[42] ProtParam tool.

Svp-Catalyzed Modification of the ACPs

apo-ACPs (1 mM) were incubated in buffer (20 mM Tris−HCl pH 8.5) with 5 mM CoASH, 40 μM PPTase Svp,[43] 10 mM $MgCl_2$ and 50 mM TCEP for 30 h at 20 °C. The ACPs were purified using a Superdex 75 16/60 column (Cytiva) equilibrated in 20 mM Tris−HCl pH 8.5, 300 mM NaCl, 50 mM TCEP. Quantitative modification was verified by HPLC-MS (Thermo Scientific).[15]

Crystallization, X-ray Data Collection and Structure Refinement

Our efforts to crystallize native VirC were unsuccessful. We therefore generated a quadruple alanine variant ($VirC^{4A}$: Cys114Ala/Gln334Ala/Arg335Ala/Arg338Ala) by site-directed mutagenesis to make it more amenable to crystallization, as previously reported.[8] Se-$VirC^{4A}$ was purified and stored in buffer (20 mM Tris−HCl pH 8.5, 300 mM NaCl, 5% glycerol) at a final concentration of 13 mg mL^{-1}. Prior to crystallization trials, sample homogeneity was checked by dynamic light scattering (DLS) using a Zetasizer NanoS (Malverne). Initial crystallization hits were obtained using the Morpheus screen (Molecular Dimensions).

A few crystals grew via the vapor diffusion method using a 1:1 ratio, with the well solution containing 11% PEG 3350, 12.5% 2-methyl-2,4-pentanediol (MPD), 100 mM amino acids (0.02 M DL-glutamic acid monohydrate, 0.02 M glycine, 0.02 M DL-serine, 0.02 M DL-alanine, 0.02 M DL-lysine monohydrochloride), 100 mM MES-imidazole buffer, pH 6.5. Crystals were then soaked in crystallization buffer containing 20% MPD prior to freezing by a stream of gaseous nitrogen. X-ray diffraction data were collected on the beamline Proxima-2A at the SOLEIL synchrotron (Saint-Aubin, France). The data set was indexed and integrated with XDS and scaled using Pointless and Aimless (CCP4 package).[19] The crystals belong to the space-group H32. The structure was solved by molecular replacement using MOLREP[18] with a computed AlphaFold2[23] structure of VirC as a search model. The initial backbone was then built using Buccaneer.[44] The final model was manually built and refined to 1.99 Å through iterative processing using Coot[45] and REFMAC5.[46] Structure geometry was validated using the program MolProbity.[47] The structure contains 97.57% of residues in the allowed regions of the Ramachandran plot and 0.26% outliers (two residues). Statistics are reported in Table S1. Figures were prepared using PyMOL.[48] The crystal structure of $VirC^{4A}$ has been deposited in the Protein Data Bank under the accession code PDB ID: 8S81.

Synchrotron Radiation Circular Dichroism

SRCD data were recorded on the DISCO beamline at the SOLEIL synchrotron. Prior to data collection, the protein samples were exchange into identical buffer (20 mM Tris−HCl pH 8.5, 300 mM NaCl, 50 mM TCEP) using a desalting column (Thermo Fisher Scientific). VirC wild type and $VirC^{4A}$ were analyzed at 370 μM (17 mg mL^{-1}), *holo*-ACP_{5a} at 1.1 mM (12 mg mL^{-1}) and *holo*-ACP_A at 900 μM (8.5 mg mL^{-1}), and the protein complexes were formed by mixing 2 μL of $VirC^{4A}$ at 740 μM (34 mg mL^{-1}) with 2 μL of *holo*-ACP_A at 1.8 mM (17 mg mL^{-1}) or 2 μL of *holo*-ACP_{5a} at 2.2 mM (24 mg mL^{-1}) respectively (under the conditions that allowed previously for obtaining the protein complex by SAXS[15]). $VirC^{4A}$ was analyzed at 17 mg mL^{-1} and *holo*-ACP_A at 8.5 mg mL^{-1}, and the protein complex was formed by mixing 2 μL of $VirC^{4A}$ at 34 mg mL^{-1} with 2 μL of *holo*-ACP_A at 17 mg mL^{-1} respectively. Protein samples were deposited between two CaF_2 coverslips with a guaranteed path length of 2 μm.[49] Use of a beam size of 4 mm × 4 mm and the photon-flux per nanometer step of 2×10^{10} photons s^{-1} in the spectral band from 270 to 170 nm, prevented radiation induced damage.[50] Spectra were collected consecutively over time and represent the mean of 3 independent acquisitions. Sample spectra were corrected for buffer background by subtracting the average spectrum of buffer alone before taking into account the protein concentration and the number of residues. Data were processed using CDToolX.[51]

SAXS Data Collection

SAXS data were acquired on the SWING beamline at the SOLEIL synchrotron. The frames were recorded using an Eiger 4 M detector at an energy of 12 keV. The distance between the sample and the detector was set to 2000 mm leading to scattering vectors q ranging from 0.0005 to 0.5 $Å^{-1}$. The protein samples were injected using the online automatic sample changer[49] into a preequilibrated HPLC-coupled size-exclusion chromatography column (Bio-SEC 100 Å, Agilent), at a

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

temperature of 15 °C. After equilibrating the column in the protein buffer (20 mM Tris−HCl, pH 8.5, 300 mM NaCl, 5% glycerol), 50 mL at 15 mg mL^{-1} of the protein sample was then injected. The buffer background and the protein elution peak were recorded with 600 successive frames. The protein concentration downstream of the elution column was followed via the absorbance at 280 nm using an in situ spectrophotometer.

The dedicated in-house application FOXTROT[52] was then used to perform data reduction to absolute units, frame averaging, and solvent subtraction. Each acquisition frame of the experiment yielded a scattering spectrum, which was then analyzed by FOXTROT to determine an R_g (radius of gyration), as well as an $I(0)$ value [the $I(0)$ depends on the protein concentration at that position in the elution peak, as described by the Guinier law (approximation $I(q) = I(0)\exp(-q^2R_g^2/3)$ for $qR_g < 1.3$)]. Notably, observing a constant R_g for a significant proportion of the concentrations present in the gel filtration peaks showed that the measurements were concentration-independent, and thus that they were effectively carried out under conditions of infinite dilution. All the frames exhibiting identical R_g as a function of $I(0)$ were averaged. Finally, the distance distribution function $P(r)$ and the maximum particle diameter D_{max} were calculated by Fourier inversion of the scattering intensity $I(q)$ using GNOM.[50] The molecular weight of $VirC^{4A}$ was determined from the acquired SAXS data using Bayesian Interference in PRIMUS.[51] The SAXS data are presented in Table S2.

Modeling of Flexible Regions in VirC

For data analysis, the crystal structure of a $VirC^{4A}$ homodimer was used to model flexible regions using EOM 3.0.[31,32] EOM 3.0 is a suite of programs used to fit experimental SAXS data to an averaged theoretical scattering intensity calculated from an ensemble of conformations. For this, a pool of n independent models based upon sequence and structural information is generated using the program RANCH,[32] and the theoretical scattering intensities of the models in the pool are calculated using FFMAKER.[53] In addition, to explore the contribution of movements beyond the $\beta8-\alpha7$ loop to the conformations adopted by $VirC^{4A}$ in solution, the overall flexibility of VirC was evaluated with SREFLEX[33] using the model of a VirC homodimer generated by AlphaFold2.[23] The quality of the models was determined using CRYSOL,[21] which allows for comparing the fit between the theoretical scattering curves calculated from atomic coordinates with experimental scattering curves, and the fits judged using the discrepancy χ^2, defined according to Konarev and colleagues.[52]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.4c00477.

Four Supporting Information Figures, 3 Supporting Information Tables, Supporting Information references (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Sabrina Collin − *Université de Lorraine, CNRS, IMoPA, Nancy F-54000, France*; Email: sabrina.collin@univ-lorraine.fr

Kira J. Weissman − *Université de Lorraine, CNRS, IMoPA, Nancy F-54000, France*; orcid.org/0000-0002-3012-2960; Email: kira.weissman@univ-lorraine.fr

Arnaud Gruez − *Université de Lorraine, CNRS, IMoPA, Nancy F-54000, France*; Email: arnaud.gruez@univ-lorraine.fr

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.4c00477

Author Contributions

Sabrina Collin: Conceptualization, investigation, formal analysis, project administration, validation, visualization, writing—original draft, review, and editing; **Kira J. Weissman**: Funding acquisition, project administration, writing—original draft, review, and editing; **Arnaud Gruez**: Conceptualization, investigation, formal analysis, project administration, validation, visualization, writing—original draft, review, and editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful for funding of this work by the Agence Nationale de la Recherche (ANR-11-JSV8-003-01, PKS−PPIs; ANR-16-CE92-0006-01, PKS STRUCTURE; ANR-20-CE93-0002-01 PKSOx to K.J.W.), the Université de Lorraine, and the Centre National de la Recherche Scientifique (CNRS). We acknowledge W. Shepard and M. Savko (Soleil Synchrotron, PROXIMA-2), J. Perez and A. Thureau (Soleil Synchrotron, SWING), and F. Wien (Soleil Synchrotron, DISCO) for help with data acquisition. Crystal screening for diffraction quality was carried out on the in-house Bruker X8-Proteum diffractometer of the Service Mutualisé de Plateformes (SMP, UL).

■ ABBREVIATIONS

HMGCS 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase; HMGS 3-hydroxy-3-methylglutaryl synthase; ACP acyl carrier protein; SAXS small-angle X-ray scattering; HMG 3-hydroxy-3-methylglutaryl; CoA coenzyme A; MVA mevalonate-dependent pathway of isoprenoid assembly; AT acyl transferase; PKS polyketide synthase; Ppant phosphopantetheine; ACP_D ACP donor to which the nucleophile is tethered during β-methylation; ACP_A ACP acceptor which bears the polyketide intermediate targeted by the β-methylation cassette; Cur curacin; AACT acetoacetyl-CoA thiolase; HMGR HMG-CoA reductase; MK mevalonate-5-kinase; PMK phosphomevalonate kinase; MDD mevalonate diphosphate decarboxylase; IPI isopentyl pyrophosphate isomerase (IPI); KS^0 ketosynthase domain which is inactive for chain extension but retains decarboxylation activity; Vir virginiamycin M; RMSD root-mean-square deviation; NRPS nonribosomal peptide synthetase; EOM ensemble optimization method; NMA normal-mode analysis; DUF domain of unknown function; MBP maltose binding protein; DLS dynamic light scattering; MPD 2-methyl-2,4-pentanediol; Rg radius of gyration; D_{max} maximum particle dimension

■ REFERENCES

(1) Yang, G.; Miton, C. M.; Tokuriki, N. A Mechanistic View of Enzyme Evolution. *Protein Sci.* **2020**, *29* (8), 1724−1747.

(2) James, L. C.; Tawfik, D. S. Conformational Diversity and Protein Evolution - a 60-Year-Old Hypothesis Revisited. *Trends Biochem. Sci.* **2003**, *28* (7), 361−368.

(3) Fernández-Martínez, L. T.; Hoskisson, P. A. Expanding, Integrating, Sensing and Responding: The Role of Primary Metabolism in Specialised Metabolite Production. *Curr. Opin. Microbiol.* **2019**, *51*, 16−21.

(4) Calderone, C. T. Isoprenoid-like Alkylations in Polyketide Biosynthesis. *Nat. Prod. Rep.* **2008**, *25* (5), 845−853.

(5) Walker, P. D.; Weir, A. N. M.; Willis, C. L.; Crump, M. P. Polyketide β-branching: diversity, mechanism and selectivity. *Nat. Prod. Rep.* **2021**, *38* (4), 723−756.

(6) Hoshino, Y.; Gaucher, E. A. On the Origin of Isoprenoid Biosynthesis. *Mol. Biol. Evol.* **2018**, *35* (9), 2185−2197.
(7) Daum, M.; Herrmann, S.; Wilkinson, B.; Bechthold, A. Genes and Enzymes Involved in Bacterial Isoprenoid Biosynthesis. *Curr. Opin. Chem. Biol.* **2009**, *13* (2), 180−188.
(8) Maloney, F. P.; Gerwick, L.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Anatomy of the β-branching enzyme of polyketide biosynthesis and its interaction with an acyl-ACP substrate. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113* (37), 10316−10321.
(9) Staunton, J.; Weissman, K. J. Polyketide Biosynthesis: A Millennium Review. *Nat. Prod. Rep.* **2001**, *18* (4), 380−416.
(10) Helfrich, E. J. N.; Piel, J. Biosynthesis of Polyketides by Trans-AT Polyketide Synthases. *Nat. Prod. Rep.* **2016**, *33* (2), 231−316.
(11) Gu, L.; Wang, B.; Kulkarni, A.; Geders, T. W.; Grindberg, R. V.; Gerwick, L.; Håkansson, K.; Wipf, P.; Smith, J. L.; Gerwick, W. H.; Sherman, D. H. Metamorphic Enzyme Assembly in Polyketide Diversification. *Nature* **2009**, *459* (7247), 731−735.
(12) Sato, K.; Katsuyama, Y.; Yokota, K.; Awakawa, T.; Tezuka, T.; Ohnishi, Y. Involvement of β-Alkylation Machinery and Two Sets of Ketosynthase-Chain-Length Factors in the Biosynthesis of Fogacin Polyketides in Actinoplanes missouriensis. *Chembiochem* **2019**, *20* (8), 1039−1050.
(13) Calderone, C. T.; Iwig, D. F.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. Incorporation of Nonmethyl Branches by Isoprenoid-like Logic: Multiple β-Alkylation Events in the Biosynthesis of Myxovirescin A1. *Chem. Biol.* **2007**, *14* (7), 835−846.
(14) Calderone, C. T.; Kowtoniuk, W. E.; Kelleher, N. L.; Walsh, C. T.; Dorrestein, P. C. Convergence of Isoprene and Polyketide Biosynthetic Machinery: Isoprenyl-S-Carrier Proteins in the pksX Pathway of Bacillus subtilis. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (24), 8977−8982.
(15) Collin, S.; Cox, R. J.; Paris, C.; Jacob, C.; Chagot, B.; Weissman, K. J.; Gruez, A. Decrypting the programming of β-methylation in virginiamycin M biosynthesis. *Nat. Commun.* **2023**, *14* (1), 1327.
(16) Campobasso, N.; Patel, M.; Wilding, I. E.; Kallender, H.; Rosenberg, M.; Gwynn, M. N. Staphylococcus aureus 3-Hydroxy-3-Methylglutaryl-CoA Synthase: Crystal Structure and Mechanism. *J. Biol. Chem.* **2004**, *279* (43), 44883−44888.
(17) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.
(18) Vagin, A.; Teplyakov, A. MOLREP: An Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997**, *30* (6), 1022−1025.
(19) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (4), 235−242.
(20) Holm, L.; Rosenström, P. Dali Server: Conservation Mapping in 3D. *Nucleic Acids Res.* **2010**, *38*, W545−W549.
(21) Theisen, M. J.; Misra, I.; Saadat, D.; Campobasso, N.; Miziorko, H. M.; Harrison, D. H. T. 3-Hydroxy-3-Methylglutaryl-CoA Synthase Intermediate Complex Observed in "Real-Time. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (47), 16442−16447.
(22) Pojer, F.; Ferrer, J.-L.; Richard, S. B.; Nagegowda, D. A.; Chye, M.-L.; Bach, T. J.; Noel, J. P. Structural Basis for the Design of Potent and Species-Specific Inhibitors of 3-Hydroxy-3-Methylglutaryl CoA Synthases. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (31), 11491−11496.
(23) Shafqat, N.; Turnbull, A.; Zschocke, J.; Oppermann, U.; Yue, W. W. Crystal Structures of Human HMG-CoA Synthase Isoforms Provide Insights into Inherited Ketogenesis Disorders and Inhibitor Design. *J. Mol. Biol.* **2010**, *398* (4), 497−506.
(24) Vögeli, B.; Engilberge, S.; Girard, E.; Riobé, F.; Maury, O.; Erb, T. J.; Shima, S.; Wagner, T. Archaeal Acetoacetyl-CoA Thiolase/HMG-CoA Synthase Complex Channels the Intermediate via a Fused CoA-Binding Site. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (13), 3380−3385.
(25) Butcher, R. A.; Schroeder, F. C.; Fischbach, M. A.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Clardy, J. The Identification of Bacillaene, the Product of the PksX Megacomplex in Bacillus subtilis. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (5), 1506−1509.
(26) Johansson, K.; Bourhis, J.-M.; Campanacci, V.; Cambillau, C.; Canard, B.; Longhi, S. Crystal Structure of the Measles Virus Phosphoprotein Domain Responsible for the Induced Folding of the C-Terminal Domain of the Nucleoprotein. *J. Biol. Chem.* **2003**, *278* (45), 44567−44573.
(27) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. PRIMUS: A Windows PC-Based System for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2003**, *36* (5), 1277−1282.
(28) Svergun, D. I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. Appl. Crystallogr.* **1992**, *25* (4), 495−503.
(29) Svergun, D.; Barberato, C.; Koch, M. H. J. CRYSOL—a Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **1995**, *28* (6), 768−773.
(30) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19* (6), 679−682.
(31) Tria, G.; Mertens, H. D. T.; Kachala, M.; Svergun, D. I. Advanced Ensemble Modelling of Flexible Macromolecules Using X-Ray Solution Scattering. *IUCrJ* **2015**, *2* (2), 207−217.
(32) Bernadó, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. Structural Characterization of Flexible Proteins Using Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2007**, *129* (17), 5656−5664.
(33) Panjkovich, A.; Svergun, D. I. Deciphering Conformational Transitions of Proteins by Small Angle X-Ray Scattering and Normal Mode Analysis. *Phys. Chem. Chem. Phys.* **2016**, *18* (8), 5707−5719.
(34) Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res.* **2018**, *46* (W1), W363−W367.
(35) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403−410.
(36) Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T. L. NCBI BLAST: A Better Web Interface. *Nucleic Acids Res.* **2008**, *36* (Web Server), W5−W9.
(37) Damry, A. M.; Jackson, C. J. The Evolution and Engineering of Enzyme Activity through Tuning Conformational Landscapes. *Protein Eng., Des. Sel.* **2021**, *34*, gzab009.
(38) Tokuriki, N.; Tawfik, D. S. Protein Dynamism and Evolvability. *Science* **2009**, *324* (5924), 203−207.
(39) Corbella, M.; Pinto, G. P.; Kamerlin, S. C. L. Loop Dynamics and the Evolution of Enzyme Activity. *Nat. Rev. Chem* **2023**, *7* (8), 536−547.
(40) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **1994**, *22* (22), 4673−4680.
(41) Gouet, P.; Robert, X.; Courcelle, E. ESPript/ENDscript: Extracting and Rendering Sequence and 3D Information from Atomic Structures of Proteins. *Nucleic Acids Res.* **2003**, *31* (13), 3320−3323.
(42) Protein Identification and Analysis Tools in the ExPASy Server. https://web.expasy.org/compute_pi/pi_tool-doc.html.
(43) Sánchez, C.; Du, L.; Edwards, D. J.; Toney, M. D.; Shen, B. Cloning and Characterization of a Phosphopantetheinyl Transferase from Streptomyces verticillus ATCC15003, the Producer of the Hybrid Peptide-Polyketide Antitumor Drug Bleomycin. *Chem. Biol.* **2001**, *8* (7), 725−738.
(44) Cowtan, K. The Buccaneer Software for Automated Model Building. 1. Tracing Protein Chains. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62* (9), 1002−1011.

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847

(45) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (12), 2126−2132.
(46) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (4), 355−367.
(47) MolProbity. http://molprobity.biochem.duke.edu/.
(48) PyMOL: The PyMOL Molecular Graphics System Version 2.0. Schrödinger, LLC: New York, NY; 2015.
(49) Réfrégiers, M.; Wien, F.; Ta, H. P.; Premvardhan, L.; Bac, S.; Jamme, F.; Rouam, V.; Lagarde, B.; Polack, F.; Giorgetta, J. L.; Ricaud, J. P.; Bordessoule, M.; Giuliani, A. DISCO Synchrotron-Radiation Circular-Dichroism Endstation at SOLEIL. *J. Synchrotron Radiat.* **2012**, *19* (5), 831−835.
(50) Miles, A. J.; Janes, R. W.; Brown, A.; Clarke, D. T.; Sutherland, J. C.; Tao, Y.; Wallace, B. A.; Hoffmann, S. V. Light Flux Density Threshold at Which Protein Denaturation Is Induced by Synchrotron Radiation Circular Dichroism Beamlines. *J. Synchrotron Radiat.* **2008**, *15* (4), 420−422.
(51) Miles, A. J.; Wallace, B. A. CDtoolX, a Downloadable Software Package for Processing and Analyses of Circular Dichroism Spectroscopic Data. *Protein Sci.* **2018**, *27* (9), 1717−1722.
(52) Thureau, A.; Roblin, P.; Pérez, J. BioSAXS on the SWING Beamline at Synchrotron SOLEIL. *J. Appl. Crystallogr.* **2021**, *54* (6), 1698−1710.
(53) Franke, D.; Petoukhov, M. V.; Konarev, P. V.; Panjkovich, A.; Tuukkanen, A.; Mertens, H. D. T.; Kikhney, A. G.; Hajizadeh, N. R.; Franklin, J. M.; Jeffries, C. M.; Svergun, D. I. ATSAS 2.8: A Comprehensive Data Analysis Suite for Small-Angle Scattering from Macromolecular Solutions. *J. Appl. Crystallogr.* **2017**, *50* (4), 1212−1225.

https://doi.org/10.1021/jacsau.4c00477
JACS Au 2024, 4, 3833−3847