Supporting Information

Solution Structure and Conformational Flexibility of a Polyketide Synthase Module

Maja Klaus[a+] and Emanuele Rossini[b+], Andreas Linden[c,d], Karthik S. Paithankar[a], Matthias Zeug[a,] Zoya Ignatova[e], Henning Urlaub[c,d], Chaitan Khosla[f], Jürgen Köfinger[b], Gerhard Hummer[b,g]*, and Martin Grininger[a]**

[a] *Institute of Organic Chemistry and Chemical Biology, Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Max-von-Laue Str, 15, 60438 Frankfurt am Main, Germany*

[b] *Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von-Laue Str, 3, 60438 Frankfurt am Main, Germany,*

[c] *Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077, Goettingen, Germany*

[d] *Institute for Clinical Chemistry, University Medical Center Göttingen, Robert Koch Strasse 40, 37075, Goettingen, Germany*

[e] *Institute for Biochemistry and Molecular Biology, University of Hamburg, Notkestr, 85, 22607 Hamburg, Germany*

[f] *Department of Chemistry, Stanford ChEM-H, Department of Chemical Engineering Stanford University, Stanford, California 94305, United States*

[g] *Institute of Biophysics, Goethe University Frankfurt, Max-von-Laue Str, 1, 60438, Frankfurt am Main, Germany*

Materials and Methods

Additional Notes
Note 1 – SAXS analysis of construct *A* and *B*.
Note 2 – Correction for protein aggregation in solution and radius of gyration of A_{ext} and A_{arch}.
Note 3 – Simulations of the arch-shaped architecture at different temperatures.
Note 4 – Flexibility of KR in A_{ext} and A_{arch}.
Note 5 – Cross-link validation.
Note 6 – General cross-link statistics of A_{ext} and A_{arch}.
Note 7 – The BioEn intensities of B_{ext} match the experiment.
Note 8 – Flexibility of KR1, ACP1 and KR in B_{ext}.
Note 9 – Differences in cross-links between holo- and apo-constructs .
Note 10 – B_{ext} agrees with measured XL-MS data
Note 11 – XL-MS data can be transferred across constructs A and B.
Note 12 – Turnover analysis of phage display enriched mutants.
Note 13 – Analysis of Second Generation ACP1 Libraries.

20 Supplementary Figures

10 Supplementary Tables

1 Supplementary Movie

Materials and Methods:

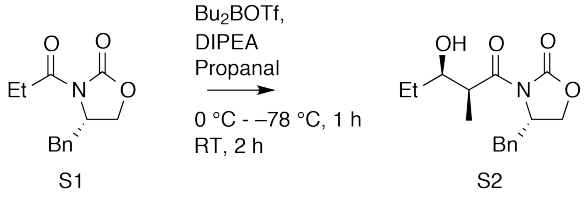
Reagents. For cloning: CloneAmp HiFi PCR Premix (Takara); primers were synthesized by Sigma Aldrich; the GeneJET Plasmid Miniprep Kit and the GeneJET Gel Extraction Kit from Thermo Scientific were used for DNA purification; Stellar Competent Cells were from Takara and One Shot® BL21 (DE3) Cells were from Invitrogen. Protein purification: All chemicals for buffer preparations were from SigmaAldrich; Isopropyl-β-D-1-thiogalactopyranoside (IPTG), carbenicillin and spectinomycin were from Sigma Aldrich/Carl Roth; LB broth (Lennox) and 2xYT media for cell cultures were from Carl Roth; Ni-NTA affinity resin (Clontech); the HiTrapQ column for anion exchange chromatography was from GE Healthcare; Amicon Ultra centrifugal filters (Millipore). PKS enzymatic assay: Coenzyme A (CoA), reduced β-nicotinamide adenine dinucleotide 2'-phosphate (NADPH), methylmalonic acid, malonic acid, and magenesium chloride hexahydrate were from Carl Roth. Adenosine-5'-triphosphate (ATP) was from SigmaAldrich. Reducing agent tris(2-carboxyethyl)-phosphine (TCEP) was from Thermo Fisher Scientific. UV-Star half area microtiter plates (Greiner). Phage Display reagents: For phage display protocols [1], T4 polynucleotide kinase, T7 DNA polymerase, T4 DNA polymerase, 40 mM dNTP mix, and M13K07 helper phage were from NEB. The QIAprep Spin M13 Kit and the QIAquick Gel Extraction Kit were from Qiagen. Mouse monoclonal anti-M13-HRP antibody (horseradish peroxidase conjugated) was from GE Healthcare 3,3',5,5'-Tetramethylbenzidine (TMB) Liquid Substrate System was from Sigma Aldrich. Competent SS230 and CJ236 cells were purchased from Lucigen and XL1 Blue cells from Agilent Technologies.

Plasmids. Plasmids harboring genes encoding individual PKS modules were generated in this study via In-Fusion Cloning (Takara) and restriction and ligation-based techniques. Tables S10 specifies the primer sequences and the resulting plasmids. Plasmids pMK53 and pADD01 were used for expression of construct *A* resp. construct *B*. Plasmids sequences were verified by DNA sequencing (SeqLab). Gene sequences were ordered from Thermo Fisher Scientific.

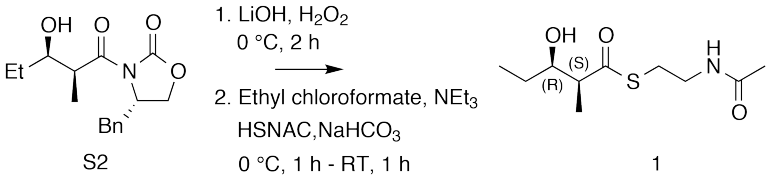
Bacterial Cell Culture and Protein Purification. All PKS proteins were expressed and purified using similar protocols. For *holo*-proteins (where the ACP domain is post-translationally modified with a phosphopantetheine arm) *E. coli* BL21 cells were co-transformed with a plasmid encoding for the phosphopantetheine transferase Sfp from *B. subtilis* (pAR357 [2]). Proteins either contained a C-terminal His_6-tag or an N-terminal MBP-tag for purification. Cultures were grown on a 1-2 L scale in 2xYT media at 37 °C to an OD_{600} of 0.3, whereupon the temperature was adjusted to 18 °C. At OD_{600} of 0.6, protein production was induced with 0.1 mM IPTG, and the cells were grown for another 18 h. Cells were harvested by centrifugation at 5000 g for 15 min and lysed by French Press (lysis buffer: 50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 10% glycerol, pH 7.6 (without imidazole for MBP-tagged proteins). The cell debris was removed by centrifugation at 50,000 g for 50 min. The supernatant was further purified using affinity chromatography. For Ni-NTA chromatography (column volume 5 mL) the supernatant was applied to the column, a first wash step was performed with the above lysis buffer (10 column volumes), followed by a second wash step using 10 column volumes of wash buffer (50 mM phosphate, 25 mM imidazole, 300 mM NaCl, 10 % glycerol, pH 7.6). Proteins were eluted with 6 column volumes elution buffer (50 mM phosphate, 500 mM imidazol, 10% glycerol, pH 7.6). MBP-tagged proteins were purified using 5 mL amylose resin per 2 L of culture. After applying the lysate, the column was washed with 12 column volumes of lysis buffer prior to elution with 6 column volumes of lysis buffer containing 10 mM maltose. Twin-Strep-tagged proteins (KS3-AT3 for phage display) were purified using 5 mL StrepTactin column. The column was washed with 5 CV lysis buffer to elution with 3 column volumes lysis buffer containing 2.5 mM destiobiotin.
Eluates from His_6-tag, Strep-tag and MBP-tag purifications were further purified by anion exchange chromatography using a HitrapQ column on an ÄKTA FPLC system. Buffer A consisted of 50 mM

phosphate, 10% glycerol, pH 7.6, whereas buffer B contained 50 mM phosphate, 500 mM NaCl, 10% glycerol, pH 7.6. The proteins were further purified using size exclusion chromatography (SEC). For SEC, an ÄKTA FPLC system using a Superose 6 Increase 10/300 GL column in buffer B was used. Enzymes *S. coelicolor* malonyl-CoA synthetase (MatB) and *S. coelicolor* methylmalonyl-CoA epimerase (SCME) were purified as decribed [3]. Protein concentrations were determined with the BCA Protein Assay Kit (Thermo Scientific). Samples were stored as aliquots at −80 °C until further use.

Synthesis of (2S, 3R) 3-hydroxy-2-methylpentanoyl-*S*-*N*-acetylcysteamine thioester (1) (provided by Mirko Joppe)
Diketide SNAC **1** was synthesized via compound **(4S, 2'S, 3'R)-3-(2'-Methyl-3'-hydroxypentanoyl)-4-benzyl-2-oxazolidinone (S2)** following protocols of Sharma et al. [4] and Peter et al. [5].



S2: Yield: 69 % (4.3 g), white crystals.
^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.37-7.21 (m, 5H), 4.75-4.69 (m, 1H), 4.27-4.19 (m, 2H), 3.90-3.86 (m, 1H), 3.81 (dq, J = 2.7, 7.0 Hz, 1H), 3.27 (dd, J = 3.4, 13.5 Hz, 1H), 2.81 (dd, J = 9.4, 13.4 Hz, 1H), 1.65-1.43 (m, 2H), 1.27 (d, J = 7.0 Hz, 3H), 0.99 (t, J = 7.5 Hz, 3H) ppm.



1: Yield: 33 % (128 mg), clear liquid.
^{1}H NMR (400 MHz, $CDCl_3$) δ = 5.76 (s, br, 1H), 3.88-3.83 (m, 1H), 3.53-3.41 (m, 2H), 3.10-2.99 (m, 2H), 2.78-2.72 (m, 1H), 1.98 (s, 3H), 1.57-1.43 (m, 2H), 1.23 (d, *J* = 7.0 Hz, 3H) 0.99 (t, *J* = 7.4 Hz, 3H) ppm. MS (ESI+) found 234.12, for $[M+H]^+$ calculated 234.12; found 256.24, calculated for $[M+Na]^+$ 256.10.

Enzymatic Assays. PKS enzymatic assay of KR1-ACP1-M2-TE was performed according to published procedures [6]. KR1-ACP1-M2-TE was used at 4 µM final concentration and enzymes MatB and SCME for *in situ* substrate generation were used at final concentrations of 4 µM and 8 µM, respectively. (2*S*,3*R*)-2-methyl-3-hydroxy-diketide-SNAC (5 mM) served as a starting substrate to prime the reaction [4,7] After certain time points, reactions were quenched with ethylacetate and the resulting product extracted twice with 450 µL dried *in vacuo* and submitted for LC-MS analysis.

Liquid Chromatography-Mass Spectrometry Analysis of Triketide Products. Dried samples were reconstituted in 100 µL methanol, separated on a Acquity UPLC BEH C18 1.7 µm RP 2.1 x 50 mm column (Waters) with an Acquity UPLC BEH C18 1.7 µm RP 2.1 x 5 mm pre-column (Waters), connected to an Ultimate 3000 LC (Dionex) HPLC over a 16 min linear gradient of acetonitrile from 5% to 95% in water. Separated samples were injected into an AmaZonX (Bruker) equipped with an

ESI Source set to positive ionization mode. The reduced triketide product was located by searching for the theoretical *m/z* for the $[M+Na]^+$- and $[M+H]^+$-ion. $[M+Na]^+$ = 195.100 and $[M+H]^+$ =173.118.

Tandem Size Exclusion Chromatography and Small Angle X–Ray Scattering. SEC-SAXS analysis was performed on the Bio-SAXS beamline BM29 at the European Synchrotron Radiation Facility (ESRF) [8]. All measurements were performed in a buffer consisting of 50 mM sodium phosphate, 500 mM NaCl, 5% glycerol, pH 7.55 using a Superose 6 Increase 10/300 GL column. Details of SAXS data collection and analysis are listed in Table S3.

Cross-link Mass-Spectrometric Analysis. We tested the homobifunctional crosslinkers bis(sulfosuccinimidyl)suberate (BS3) and disuccinimidyl suberate (DSS), as well as the heterobifunctional reagent NHS–Diazirine/succinimidyl 4,4–azipentanoate (SDA) on MBP–M2(2) for their ability to capture crosslinks, and initially judged crosslinking efficiency by SDS-PAGE (data not shown). BS3 or DSS showed few crosslinks, which we attributed to the low number lysine residues present in the DEBS core of the protein. The photoactive reagent SDA performed better in this regards owing to its reaction with primary amino groups of e.g. lysines and insert under radical formation into C–C and C–heteroatom bonds. By taking into consideration the lysine chain length, the 3.9 Å spacer, the chain length of long amino acids (e.g. lysine and arginine), and conformational flexibility, a maximum Cα–Cα distance of 20 Å is achieved.

SDA cross-linking: 150 to 175 µg of construct *A* and *B* at a final protein concentration of 1 µg/µL or less were incubated with 2 mM SDA (100 mM stock in DMSO) for 30 min at room temperature, in a buffer consisting of 50 mM sodium phosphate, 500 mM NaCl, 10% glycerol, pH 7.55 (construct *A*) or 200 mM sodium phosphate, 200 mM NaCl, 10% glycerol, pH 7.55 (construct *B*, second experiment). The cross-linking reactions were quenched with 50 mM Tris-HCl. The samples were dialyzed against reconstitution buffer via a membrane filter (MF Membrane Filters, 0.025 µm VSWP, Merck). Afterwards, samples were irradiated with UV light (365 nm) for 5 min at 4 °C.

Protein Digestion and Enrichment of Cross-linked Peptides: Proteins of cross-linked samples were reduced and alkylated with 10 mM DTT and 40 mM iodoacetamide, respectively. Proteins were digested by trypsin in an enzyme-to-protein ratio of 1:50 at 37 °C overnight at a final concentration of 1 M urea. Peptides were acidified with trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v), desalted on MicroSpin Colums (Harvard Apparatus) following manufacturer's instructions and vacuum dried. Peptides were resuspended in 50 µL 30% acetonitrile (v/v)/0.1% TFA (v/v) to enrich cross-linked peptide species by peptide size exclusion chromatography (SuperdexPeptide 3.2/300 column, GE Healthcare). Fractions of 50 µL were collected at a flow rate of 50 µL/min and those that eluted first and contained the cross-linked peptide pairs were subjected to LC-MS/MS analysis.

LC-MS/MS Analysis: Cross-linked peptides were measured in technical duplicates on an Orbitrap Fusion Tribrid Mass Spectrometer or on a Q Exactive HF-X coupled to a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific) equipped with an in house-packed C_{18} column (ReproSil-Pur 120 C_{18}-AQ, 1.9 µm pore size, 75 µm inner diameter, 30 cm length, Dr. Maisch GmbH). Samples were separated applying the following gradient: mobile phase A consisted of 0.1% formic acid (FA, v/v), mobile phase B of 80% ACN/0.08% FA (v/v). The gradient started at 5% B, increasing to 12, 15 or 20% B within 3 min (according to fraction), followed by a continuous increase to 46% B within 45 min, then keeping B constant at 90% for 8 min. After each gradient the column was again equilibrated to 5% B for 2 min. The flow rate was set to 300 nL/min. MS1 survey scans were acquired in the orbitrap (OT) with a resolution of 120,000, an injection time (IT) of 60 ms (50 ms on HF-X) and an automatic gain control (AGC) target of 5×10^5. Dynamic exclusion was set to 10 s

(30 s on HF-X) and only charge states between +3 and +8 were considered for fragmentation. MS2 spectra were acquired in the OT of the 20 (25 on HF-X) most abundant precursor ions, resolution 30,000, IT 128 ms and AGC target 5×10^4. Fragmentation was enforced by higher-energy collisional dissociation (HCD) at 30%.

Data Analysis: ProteomeDiscoverer 1.4 (Thermo Fisher Scientific) was used for converting raw files into .mgf format (signal-to-noise ratio 1.5, 1000–10000 Da precursor mass). The generated .mgf files were subjected to pLink v. 1.23 (pFind group)[9] to identify cross-linked peptides. Here, default settings were applied with carbamidomethylation of cysteines as fixed and oxidation of methionines as variable modification. FDR was set to 0.01. SDA was selected as cross-linker. All spectra were evaluated manually.

Protein Homology Modeling. We constructed homology models of the different protein domains of DEBS using MODELLER version 9.16 [10]. The X-ray crystallographic structures of *S. erythraea* KS-AT_2 (PDB ID: 2hg4, 2qo3, 6c9u) [6, 11-12], KR2 (2fr0, 2fr1) [13], ACP2 (2ju2) [14], and DD (1pzq)[15] were used as templates. The C-terminus 1883:1914 region (DD) was not modeled due to alignment mismatch. A three-dimensional structure was already available for the domains MBP (1anf) [16], KR1 (2fr0) [13], ACP2 (2ju2) [14], and TE (1mo2) [17]. For each protein domain we constructed a set of $2x10^2$ models. The homology models were then selected based on their discrete optimized protein energy (DOPE) score. Pymol software was used for visual inspection of the selected structures [18].

Construction of the *in silico* Starting Models (Fig. S5). *Model ext:* We built initial structural arrangements for A_{ext} and B_{ext} as informed by previous SAXS investigations of DEBS [19]. Both models carry an extended conformation of KS-AT_2 and their post-AT linkers are fully structured as produced by the homology modeling. The non-native MBP domains were located N-terminally of the KS. The length of the modeled domains is highlighted in Table S9. The models were constructed using the software Chimera [20].

Model arch: The initial structural arrangement of A_{arch} was constructed by docking the modeled protein domains within the cryoEM density of the related PikAIII complex (EMD-5664) [21-22]. The rigid docking was performed using the software Chimera [20]. To obtain an arched conformation, we separately docked a KS dimer and two copies of LD-AT generated from the extended KS-AT_2 homology model as informed from PikAIII complex [22]. Next, we merged the structures of the split domains to reconstruct the full KS-AT_2 domain. The post-AT linkers were not resolved within the cryoEM density and excluded from the docking [22]. We then docked the remaining KR, ACP, and DD domains within the density. The initial conformation of the native A_{arch} displays a cross-correlation to the density equal to 0.71 (Fig. S6). The non-native MBPs were located at the N-terminal regions of KS-AT_2 using the software Chimera [20]. The length of the modeled domains is highlighted in Table S9.

Coarse-Grained Simulations. We generated a canonical ensemble of models A_{ext}, A_{arch}, and B_{ext} using Monte Carlo (MC) simulations based on coarse-grained (CG) potential energies [23]. Folded protein domains were treated as rigid bodies described by one interaction site per residue located at their Cα atoms. The protein domains were linked by flexible linkers represented as Gaussian chain polymers and placed in a cubic box of 50 nm size. The solvent dielectric constant was set to $\varepsilon=80$. The post-AT regions of A_{arch} were modeled as flexible linkers (Table S9). We kept KS-AT_2 frozen and used step sizes of 3 Å and 3.14 rad for the MC translational and rotational movements of all other domains. For all models, conformations were sampled in a temperature replica-exchange simulation scheme covering the range of 298 to 598 K in steps of 20 K. For A_{arch}, we also performed replica-exchange simulations collecting replicas between 298 and 398 K every 20 K to preserve the

arch-shaped conformation otherwise disrupted at higher temperatures (Note 2). For each model, after an initial equilibration, a total of 9.0×10^5 MC conformations were produced to exhaustively sample relevant conformations. To augment model flexibility and the sampled conformational space of the models, only the conformations produced at highest temperature were considered for further analysis. A_{arch} conformations sampled at 598 K are denoted as A_{arch}^{598}.

Computation of SAXS Intensities. We computed the SAXS intensities for each of the individual CG conformations using the package FoXS [24-25].

Ensemble Refinement of SAXS Intensities. We used the Bayesian Inference of Ensembles (BioEn) method to determine the statistical weights of A_{ext}, A_{arch}, and B_{ext} ensembles of conformations underlining the measured SAXS intensities [26-28]. BioEn refinement optimizes the weights of the structures in the ensemble to improve the agreement with experimental intensities while keeping the distribution close to the reference. We performed the BioEn computations considering the SAXS intensities in the range of $0.02 < q < 0.3$ Å^{-1}. We used L-curve analysis to identify the confidence parameter θ for an optimal tradeoff between the consistency of simulated and experimental data (chi-squared metric χ^2) and changes in weights based on relative entropy S_{KL} (Fig. S7). From the scattering curves, the radius of gyration R_g was calculated using the Guinier approximation [29]. We reduced the sampled ensembles to the conformations that together fulfill more than 99% of the statistical cumulative weight for evaluation of the measured cross-links (Fig. S5).

Correction for Protein Aggregation in Solution. Systematic deviations between the measured and simulated scattering curves at low q range suggested moderate aggregation of the systems in solution (Fig. S2 and S8). To quantify and account for the fraction of aggregation, we combined each simulated curve with the intensities of a globular model at a statistical weight that matched the experiment. We approximated the globular aggregate in solution as a sphere and considered the BioEn intensities at low q range between 0 and 0.02 Å^{-1}. We optimized the radius of gyration of the sphere, R_g^{sph}, and its statistical weight, w_{sph}, to match the experimental data as:

$$I_{sph}(q) = \frac{9\left(sin\left(qR_g^{sph}\right) - qR_g^{sph}cos\left(qR_g^{sph}\right)\right)^2}{\left(qR_g^{sph}\right)^6}$$

$$I_{agg}(q) = \left(1 - w_{sph}\right)I_{BioEN}(q) + w_{sph}\frac{I_{sph}(q)}{I_{sph}(0)}\bar{I}(0)$$

with q the scattering angle and $I_{agg}(q)$ the intensity of the complex in solution, $I_{BioEN}(q)$ the BioEn intensities, and $I_{sph}(q)$ the intensities of the sphere, $I_{sph}(0)$ the sphere intensity at zero q, and $I(0)$ the intensity prefactor at zero q (see also SI Note 1).

Quantification of Module Flexibility. We evaluated the flexibility across the conformations that explain 90% of the BioEn scattering curves (Fig. 2). For each ensemble, we monitored the distance of the center of mass of the KR domains between the conformations along the ranking and the one with highest statistical weight. Similarly, we evaluated the flexibility of the MBP of A_{ext}, and the KR1 and ACP1 domains of B_{ext}.

Structural Mapping of the XL-MS data. We mapped the measured cross-links of construct *A* over the 1000 A_{ext} and A_{arch} conformations ranked by weight that fulfill 99.9% of the BioEn scattering curves. For all selected conformations, we calculated the Cα pair-distances across the reported cross-linked residues. Additionally, we quantified the frequencies of interaction across all exposed

residues of KS-AT_2 and MBP in A_{ext}. To account for the CG level of theory of the simulations, the residues were considered interacting if their Ca atoms were located within a distance of 21 Å. The same analysis was performed to map the measured XL-MS data of construct *B* over the 5000 conformations of model B_{ext} that fulfill 99% of the BioEn scattering curve. To evaluate transferability of the XL-MS data across the different constructs, we additionally mapped the 1000 top-ranked B_{ext} conformations (92% in cumulative weight) on the XL-MS data measured for construct *A*. We then quantified the frequencies of interaction across the rigid domains shared across the two constructs.

Explorative Analysis. We used the Python libraries Numpy [30], Scipy [31], Pandas [32], and MDAnalysis[33-34] to manage cross-link data, derive the bead coordinates along the CG trajectories, and perform distance computations. We used the Python library Matplotlib for the visualization of our results [35]. We used VMD and Pymol to visualize the simulated frames [18, 36].

Phage Display Methodology. Phage display experiments were carried out according to the protocol described by Tonikian *et al*. [1]. Specific experimental setup is described below.

Library generation: A library of ACP1(2) mutants fused to the N-terminus of the minor coat protein P3 of the M13 bacteriophage was generated as described before [1]. A primer for randomization of five amino acids in the chain translocation epitope of ACP1 was designed using NNK at the target positions (P-MK162: 5' CTG GCG TCG CTG CCC GCG NNK GAG CGC NNK NNK GCG CTG TTC NNK CTC GTG CGC NNK CAC GCG GCC GCC GTC CTC 3' (N: A/C/G/T K: G/T).

Biopanning: For all panning steps a buffer containing 50 mM sodium phosphate, 150 mM NaCl, pH 7.55 was used. All incubations were done at room temperature while shaking at 200 rpm. Panning was performed in a StrepTactinXT coated microtiter plate. As a first step, four wells were coated with (3)KS3-AT3--Strep (100 µL at 5 µg/mL) for 2 h. Afterwards, the plate was blocked using 200 µL of blocking buffer (50 mM sodium phosphate, 150 mM NaCl, 0.5% BSA, 0.05% Tween20, pH 7.55). At the same time the purified phage library was also blocked in the same buffer (procedure as described as before [1]). The blocked phage library (100 µL per well) was incubated for 1 h on the coated and blocked wells, prior to washing with 200 µL of wash buffer (50 mM sodium phosphate, 150 mM NaCl, 0.05% Tween20, pH 7.55). Wash steps were increased each round over the course of the panning protocol (10x, 12x, 14x, 16x, 18x). Elution of bound phage was done with 100 µL of 100 mM hydrochloric acid for 20 min. The eluted phage were collected and the solution neutralized by adding 1/10 volume of 1 M tris base pH 8.0 and 1/10 volume of blocking buffer. The elution was propagated and the titer was determined.

For propagation 2 mL of log phase *E. coli* SS320 (OD_{600} 0.8) were added to 200 µL of eluted phage and infection was allowed for 30 min and 37 °C, 200 rpm whereupon 20 µL M13K07 helper phage ($1.0x10^{12}$ PFU/mL) were added. Following incubation for 1h at 37 °C, 200 rpm 45 mL of 2xYT/carb/kan were added to the culture and cells continued to grow over night. The next day, phages were harvested by centrifugation at 5000 x g for 10 min to remove the cells. Phage particles remained in the supernatant and were precipitated with a solution containing 20% PEG8000 (w/v) and 2.5 M NaCl for 15 min at room temperature. Subsequently, phages were pellet at 13000 x g for 10 min. The pellet containing the phage was resuspended in 1 mL of blocking buffer. Another spin for 5 min at 15000 x g removed the remaining cells. Phages were either stored at 4 °C, or used directly.

ELISA: To test the eluted phage for increased and specific binding, a specificity ELISA was performed. Colonies from the target elution rounds were grown in a minitube rack containing 400 µL 2xYT/carb/K07 medium (supplemented with 10^{10} PFU/mL M13K07).

A MaxiSorp plate was coated for 2 h at room temperature with 100 µL of 2 µg/mL target protein (3)KS3-AT3 and as a negative control with BSA. Afterwards the plate was blocked for 1 h with 200 µL of blocking buffer (50 mM sodium phosphate, 150 mM NaCl, 0.5% BSA, 0.05% Tween20, pH 7.55), followed by incubation of 100 µl of phage containing supernatant from the minitube rack for 1 h. The plate was washed twice with 200 µL of wash buffer (50 mM sodium phosphate, 150 mM NaCl, 0.05 % Tween20, pH 7.55), before incubating 100 uL of a 1:5000 anti-M13-HRP antibody solution in blocking buffer for 1 h. Following antibody incubation the plate was washed three times before development with 100 uL TMB substrate and quenching upon color change with 1 M phosphoric acid. Read out was carried out at 450 nm on a Tecan Synergy HT plate reader.

Titration ELISA: To further confirm the relative binding intensity of the newly enriched mutants, a titration ELISA using purified phage was used. Phages were purified out of a 45 mL culture using described protocols [1]. The OD_{268} was used as a measure to determine phage amounts (OD_{268}= $5x10^{12}$ PFU/mL). The phage stock was serial diluted in ELISA blocking buffer and applied to the ELISA plate using the same protocol as described above. The advantage of this assay is that defined phage amounts can be used and the signal can be normalized across different phage preparations.

Additional SI Notes:

Note 1 – SAXS analysis of constructs *A* and *B*. Both constructs *A* and *B* elute as a single peaks, but display a fronting shoulder (Fig S2A, B). Construct *A* elutes slightly later than construct *B* (26.77 min / #1627 vs. 26.75 min / #1623), consistent with its lower M_W. First, a basic SEC-SAXS analysis was conducted to characterize the sample. Data was collected up to 5 $Å^{-1}$, but truncated at 3 $Å^{-1}$ due to the low signal to noise ratio at high *q*. A suitable frame range displaying a stable R_g at the intensity maximum of the SEC-SAXS peak was selected (Fig S2B). The sample average was generated using frames 1618-1641 for construct *A* and frames 1619-1646 for construct *B*. For buffer correction, two regions to the left and right of the peak were averaged to obtain the background intensity and subtracted from the sample average (frames 215-486 and 2087-2469 for construct *A*; frames 500-750 and 2500-2750 for construct *B*, Fig S2C). Care was taken to not include any contaminations by inspecting the UV trace (Fig S2D).

The Kratky-plots for constructs *A* and *B* display a second peak at ~ 7 R_g, indicating a multidomain protein, which matches the expectations for a PKS system. The curves converge to 0 at high *q*, showing that the bulk of the particles is properly folded, and that they are not overly flexible (Fig. S2E).

The Guinier-region was determined as 0.5 qR_g – 1.3 qR_g, truncating the plot at the canonical 1.3 Rg and omitting the first data points. The downward slope of the scattering profile introduced by these first data points does not resemble inter-particle repulsion but is an artifact arising from beamstop/masking issues. Aggregation and repulsive interactions can be distinguished by the shape of the residuals of the Guinier fit. The upward facing parabola shape of the *q*-region extended up to $q*R_g$ of 3, is indicative of aggregation [37]. Repulsive interactions would show as an downward facing parabola instead.

Both regressions display R-values > 0.999 and a flat distributions of residuals (Fig. S3A). The Guinier-analysis yielded similar results for both constructs. For construct *A*, I_0 was determined as 157 ± 1 and R_g as 77.2 ± 0.2 Å. For construct *B*, I_0 and R_g derived as 139 ± 1 and 76.6 ± 0.2 Å, respectively (Fig. S3B).

Molecular weight estimates are in agreement with previous reports (Fig. S3B) [19, 38]. Here, the estimates based on the volume of correlation V_c are in excellent agreement with the M_W derived from the construct's atomic composition[39]. The integrated intensities converge at high *q*, indicating accurate estimates. The M_W calculations based on the Porod-volume V_p are larger than the expected values. V_p is known to be affected by partially extended/unfolded or non-globular conformations [40]. Computing *P(r)* functions with BIFT[41] and GNOM[29] (BioXtas RAW suite [42]) revealed similar pair distance distributions, with good agreement between the experimental and back calculated scattering profiles for construct *B* and the BIFT solution of construct *A* ($1.057 < \chi^2 > 1.087$). The GNOM solution of construct *A* displays a slightly higher χ^2 value of 1.21 and displays slight periodicity in the high *r* range. Larger D_{max} values for construct *B* than for construct *A* were received, consistent with the increased size/domain number of construct *B*. The sharp descent of the curves for the automated BIFT solutions indicate slight overfitting (Fig 3C). When the D_{max} value is adjusted manually in GNOM, the *P(r)* functions only slowly converge to 0, indicating either the presence of aggregation in the sample or partial unfolding of the protein. The I_0/R_g values from the *P(r)* function match the Guinier derived values.

The basic sample characterization revealed that the sample is both folded and displays the expected dimeric composition. The constructs behave as expected for modPKS proteins. However, the SEC

procedure could not completely remove minor aggregates from the main fraction. As already the presence of small amounts of aggregation can impair SAXS based calculations, we continued to deconvolute the scattering profiles *in silico*. By this, a larger frame range spanning the complete peak could be considered for modeling (1536-1717 for construct *A*, and 1539-1704 for construct *B*), thus increasing the signal to noise ratio.

Note 2 – Correction for protein aggregation in solution and radius of gyration of A_{ext} and A_{arch}. For ensemble refinement via BioEN, we used SEC-SAXS frames over the entire dimer elution peak. Systematic deviations between the measured and simulated scattering curves at low q range suggested moderate aggregation of the systems in solution, as expected (Fig. S2A and S9; see also SI Note 1). To quantify and account for the fraction of aggregation, we combined each simulated curve with the intensities of a globular model at a statistical weight that matched the experiment. We approximated the globular aggregate in solution as a sphere and considered the BioEn intensities at low q range between 0 and 0.02 Å^{-1}. We optimized the radius of gyration of the sphere, R_g^{sph}, and its statistical weight, w_{sph}, to match the experimental data as:

$$I_{sph}(q) = \frac{9\left(sin\left(qR_g^{sph}\right) - qR_g^{sph}cos\left(qR_g^{sph}\right)\right)^2}{\left(qR_g^{sph}\right)^6}$$

$$I_{agg}(q) = \left(1 - w_{sph}\right)I_{BioEN}(q) + w_{sph}\frac{I_{sph}(q)}{I_{sph}(0)}\bar{I}(0)$$

with q the scattering angle and $I_{agg}(q)$ the intensity of the complex in solution, $I_{BioEN}(q)$ the BioEn intensities, and $I_{sph}(q)$ the intensities of the sphere, $I_{sph}(0)$ the sphere intensity at zero q, and $\bar{I}(0)$ the intensity prefactor at zero q (see also SI Note 1).

The R_g-value calculated from SEC-SAXS frames over the entire dimeric elution peak is 79.8 Å for construct *A*. Computed R_g values from the BioEn scattering curves, based on the Guinier approximation, for A_{ext} and A_{arch} were equal to 67.6 Å and 64.0 Å, respectively. When accounting for the fraction of protein aggregation and correcting the BioEn scattering curves, the R_g-value computed for A_{ext} in presence of aggregation was equal to 79.6 Å. The value perfectly agrees with the experimental data. The computation predicts an aggregate present in solution with R_g^{sph} of 152.7 Å at a fraction of 6.9%. When the fraction of aggregation predicted for A_{ext} was applied to A_{arch}, it resulted in $R_g \approx 76.7$ Å, thus matching the experimental data less well (Fig. 2). For A_{arch}, a spherical model with R_g^{sph} of 144.0 Å and a fraction of aggregates in solution equal to 10.1% are required to obtain a R_g closer to experiment and equal to 81.0 Å (Fig. S9 & S11). The scattering curve of A_{arch} obtained at 598 K (A_{arch}^{598}, disrupting the arch-shaped fold) match the experimental Guinier region with a R_g of 79.6 at 7.1% aggregation and R_g^{sph} of 153.7 Å (Fig. S8).

In all cases, a simple model for an aggregate is sufficient to gently refine our physically meaningful simulation ensembles of conformers to fit the data. Moreover, the inferred aggregate sizes are much larger than the conformers in the structural ensembles and the statistical weights of the aggregates are small in all cases. These results provide strong evidence that our simple model of the aggregate is adequate and that the scattering intensities at larger q-values are not affected by aggregation.

Note 3 – Simulations of the arch-shaped architecture at different temperatures. We simulated A_{arch} in two alternative rigid-body temperature Replica-Exchange simulations ranging from 298 K to 398 K and from 298 K to 598 K, respectively. For both simulations, we analyzed the

conformations produced at the highest temperature to augment the sampled conformational space. The conformations produced at 398 K retain the arch-shaped architecture, although the flexible post-AT linkers allow occasional formation of globular models with KR approaching KS more closely (Fig. 2). The conformations relaxed at 598 K, here defined A_{arch}^{598}, disrupt the structural organization of the cryoEM density (Fig. S7). The A_{arch}^{598} scattering curves match the experimental Guinier region when accounting for protein aggregation in solution of 7.1 % and R_g^{sph} of 153.7 Å (SI Note 2). Analogously to A_{arch}, discrepancies from experimental data are highlighted at q values ranging from 0.05 and 0.1 $Å^{-1}$ (Fig. 2. S9 , S10 & S11). After refinement, 5 conformations together explain 90% of the fitting and the A_{arch}^{598} conformation ranked at highest statistical weight is rather globular (Fig. S11). Compared to the other models, A_{arch}^{598} displays the highest KR flexibility with an average value of 36.0 Å (Fig. S11). Such flexibility results in irrelevant architectures where KR deeply leaks towards KS-AT_2 and excides the N-terminal region of the module (Fig. S11). Having the general topology of module 2 disrupted, A_{arch}^{598} was not considered for further structural analysis.

Note 4 – Flexibility of KR in A_{ext} and A_{arch}. We monitored the flexibility of the KR domains across the conformations that jointly explain 90% of the BioEn fittings (Fig. S12). A_{ext} displays high KR mobility with an average distance from the reference of 16.1 Å. With its large mobility, the model provides a convenient platform to investigate the conformational space sampled through CG simulation. The KR average distance of A_{arch} at 398 K is equal to 11.2 Å and two clusters are displayed (Fig. S12). A total of 8 over 56 A_{arch} conformations ranked at high weight disrupt the cryoEM conformation, contributing together to 19% of the fitting. Alternatively, 48 conformations preserve the fully arch-shaped conformation. These conformations display an average distance of 3.6 Å and contribute together to explain 71% of the fitting. The arch-shaped organization emerges as a stable architecture that, however, does not fully match the experimental SAXS data.

Note 5 – Cross-link validation. The structures used to build the models match 83% of the cross-links measured within single protein domains. This confirms the high specificity of the measured data (Table S5). Based on these results, we used the cross-links measured across different protein domains to locate and describe candidate regions of DDI.

Note 6 – General cross-link statistics of A_{ext} and A_{arch}. A total of 62 cross-links were measured across the different protein domains of construct *A*. We mapped the measured data over 1000 A_{ext} and A_{arch} conformations that contributed to 99.9% of the BioEn fittings. A_{ext} originally satisfied 46 cross-links (Table S6). MBP displays high flexibility and so we excluded the cross-links based on its lysine residues. XL7, 19, 48, 49, and 50 were not captured by the analyzed conformations. Interestingly, XL19, 48, 49, and 50 were measured at low frequency and were described by the simulation when considering all 45000 relaxed conformations of A_{ext} (Table S6). This highlights the low influence that ACP dynamics plays on the scattering curves. The cumulative weights of A_{ext} conformations matching the measured cross-links range between a minimum close to 0 (MBP and ACP based cross-links) and a maximum of 0.65 (XL39 across KS-AT_2 and KR). In contrast, A_{arch} satisfies 26 XLs involving mostly residues from MBP and ACP (Table S6). These XLs are of minor structural relevance due to the large conformational space sampled by the non-native MBP domains and the scarce role played by ACP on the BioEn results. In conclusion, model A_{ext} agrees with the experimental data overall better than A_{arch}.

Note 7 – The BioEn intensities of B_{ext} match the experiment. We computed R_g values from the BioEn scattering curve. Based on the Guinier approximation, the R_g of B_{ext} was equal to 72.6 Å and poorly agreed with the measured R_g of 76.6 Å (Table S3). The discrepancy between measured and simulated data is compatible with a fraction of aggregation in solution (Fig. S16). Analogously to construct *A*, an aggregate present in solution with R_g^{sph} of 150.8 Å at a fraction of 2.4% shift the R_g of B_{ext} to 78.1 Å. The *in silico* SAXS intensities of the extended model agree perfectly with the experiment (Fig. S16). These results reinforce the conclusion that M2 presents an extended conformation of the KS-AT dimer in solution.

Note 8 – Flexibility of KR1, ACP1 and KR in B_{ext}. KR1 presents an average distance from the reference of 30.7 Å (Fig. S17) that is analogous to the one observed for MBP (39.0 Å). Both domains are placed terminally in the biological constructs and their high degree of flexibility depends on the absence of a further N-terminal domain. Interestingly, ACP1 displays a lower average distance of 11.7 Å and contacts the KS-AT_2 domain in defined regions located in proximity of the DDI across both KS and AT (Fig. S17). The model is characterized by a lower flexibility compared to A_{ext}, with an average KR distance of 10.7 Å (Fig. S16). The C-terminal TE of B_{ext} is larger compared to the DD of A_{ext} and act as a structural constraint that limits KR mobility (Fig. S18).

Note 9 – Differences in cross-links between holo- and apo-constructs. XL-MS experiments were preformed with construct *A* in apo- and holo-form, as well as with construct *B* in holo-form (Table S6 & 7). For construct *A*, we see different cross-linked peptide spectral matches (CSMs) between the apo and holo-protein, mainly comprising ACP in its docking to other domains. In holo-construct *A*, just 6 of overall 15 high frequency CSMs were found preserved (XL14-18 and XL21) and all medium frequency CSMs omitted. 4 new CSMs between MBP and KS were found in holo-construct *A*. The reproducibility of CSMs between the three experiments with apo-protein was good, so that the difference in cross-link patterns between apo- and holo-construct *A* seem significant. However, XL-MS has just been performed once for holo-construct *A*. For holo-construct *B*, a similar small number of CSMs was observed as for holo-construct *A*, however CSMs were often unique. Just 4 of overall 12 crosslinks, B-XL16, B-XL17, B-XL18, and B-XL27, were found in both constructs and matched with high frequency and high statistical weight by B_{ext} (Table S7). Just a few CSMs involving ACP were observed, similar as found for holo-construct *A*. XL-MS with holo-construct *B* features an interaction between AT and KR that has not been observed in constructs *A* (apo and holo), which may reflect a differently constrained KR, by either DD2 (construct *A*) or TE (construct *B*) (see Table S6). Overall, our data on construct *B*, i.e. the fewer CSMs with ACP, are supportive of phosphopantetheine constraining ACP docking. Further studies with phosphopantetheinylated vs. non-phosphopantetheinylated samples will be needed to understand the impact of phosphopantetheinylation on ACP docking.

Note 10 – B_{ext} agrees with measured XL-MS data. A total of 30 cross-links were measured across the different protein domains of construct *B*. We mapped the measured CSMs over the 5000 conformations that contributed to 99% of the BioEn cumulative weights. Compared to construct *A*, the number of CSMs measured for construct *B* exhibits lower reproducibility. Nevertheless, the B_{ext} model satisfies overall 23 XLs (Table S7). B-XL16, B-XL17, B-XL18, and B-XL27 were reproduced between experimental data of both constructs and matched with high frequency and high statistical weight by B_{ext} (Table S7). The outliers B-XL1, B-XL3, B-XL4 can be explained by the high mobility of the KR1 domain. The cross-links B-XL22, B-XL23, B-XL24, and B-XL25 measured across AT and KR2

were also not captured by the model. As for XL7, this can be explained by the attachment of the post-AT linker to the KS-AT core. The outliers might, however, be explained by measurement artifacts due to protein aggregation or reflect a higher mobility of the processing domain as a whole (KR2-ACP2-TE).

Note 11 – XL-MS data can be transferred across constructs *A* and *B*. To evaluate the transferability of the XL-MS data across different constructs, we mapped the highest ranked B_{ext} conformations (1000) on the XL-MS data measured for construct *A*. Interestingly, B_{ext} shows the same pattern of XL interactions documented for A_{ext} (Table S6). In particular, we report high frequencies of interaction across the post-AT linker and KR2 and interaction hot spots of ACP with both KS-AT and KR, XL7 remains as an outlier shared by both A_{ext} and B_{ext}. These results highlight that the same XL-MS data can be mapped onto different DEBS constructs. Such a strategy might help the structural investigation of additional DEBS modules in the future.

Note 12 – Turnover analysis of phage display enriched mutants. The bimodular DEBS-derived system LM+M1+M3-TE, used as a testbed in previous studies [43-44], provides a suitable setting for studying the integrity of ACP1 and catalytic domain recognition by ACP1 (bimodular system built by LM(4), (5)M1(2), and (3)M3-TE; number in brackets denote docking domains used to mediate a weak non-covalent interaction). The sequence identity between the KSs of M2 and M3 is 58% on the protein level (EMBOSS pairwise assembly, www.ebi.ac.uk) [45], so that the ACP1-KS3 interaction, although involving a non-cognate interface, is likely preserved. Previous work demonstrated the catalytic activity of the LM+M1+M3-TE bimodular assembly line and support this assumption. Note that by replacing M2 with DEBS module 3 (M3), yielding LM+M1+M3-TE, the translocation from M1 to M3 is rate-limiting [43-44], and the performance of the ACP1 mutant has direct impact on LM+$M1_{mut}$ -M3-TE turnover.
The phage display approach was performed to evolve the non-cognate ACP1:KS3 interface towards higher affinity, detectable by ELISA screening, thereby studying whether a dedicated sequence motif emerges in response to a single specific binding site at the KS. The turnover rate of the LM+M1+M3-TE assembly line, in which M1 is mutated in ACP1 according to phage display hits, can readout the fitness of the evolved ACP1s as well as the impact of mutations on ACP function (ACP1 mutations were cloned into the full-length M1 construct and the turnover monitored (LM+$M1_{mut}$ -M3-TE)).

Note 13 – Analysis of Second Generation ACP1 Libraries. Similar to the design of library 1, the positions at helix 1 chosen for randomization were not conserved across several ACP domains and located on the surface of the ACP. Using the same protocol established for biopanning of ACP1–Lib1, both libraries were purified and subjected to the biopanning protocol. The theoretical diversity for both libraries was 6.40x107 PFU/mL and the practical diversity exceeded the theoretical diversity with $9.72x10^8$ PFU/mL (ACP1–Lib2) and $3.02x10^8$ PFU/mL (ACP1–Lib3) in both cases. In screening library 2 and 3 mutants (Lib2-Mut"X" and Lib3-Mut"X", respectively) to KS-AT binding by ELISA, we observed increased BSA background. All these ACP1 variants were still enzymatically active, thus overall intact (Fig. S20E). For the additional evaluation of putative hits, we performed an ELISA titration experiment (Fig. S20F). This data was collected with phages purified out of a 45 mL culture using described protocols [1]. Working with purified phages at different conditions allowed a better judgment of affinity than ACP1 wild-type. Just those mutants are included in sequence analysis, which showed higher affinity than ACP1 wild-type (Fig. S20G). Similar to hits from ACP1–Lib1, no consensus sequence was observed among the enriched mutants.

Supporting Figures

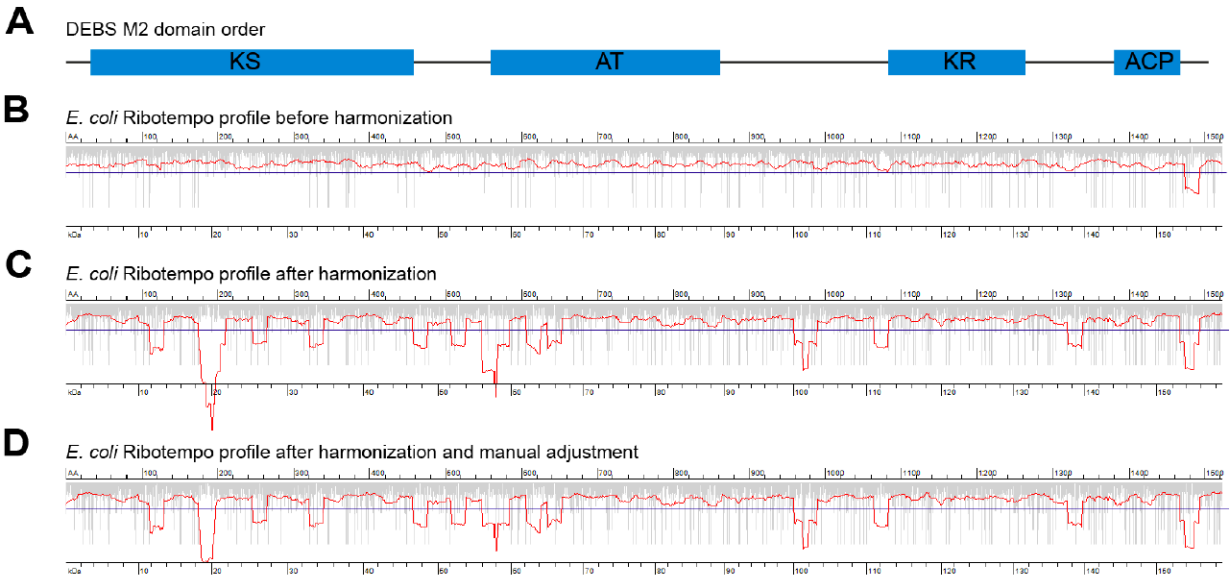


Figure S1. The effect of codon harmonization on the predicted translation rates of M2. (A) M2 domain architecture. Prediction of the translation rate in *E. coli* before harmonization (B), after harmonization (C), and after manual adjustment of the harmonized sequence (D). Translation rates were predicted using the Ribotempo tool [46]. For gene sequences of original and harmonized gene, see Table S2.

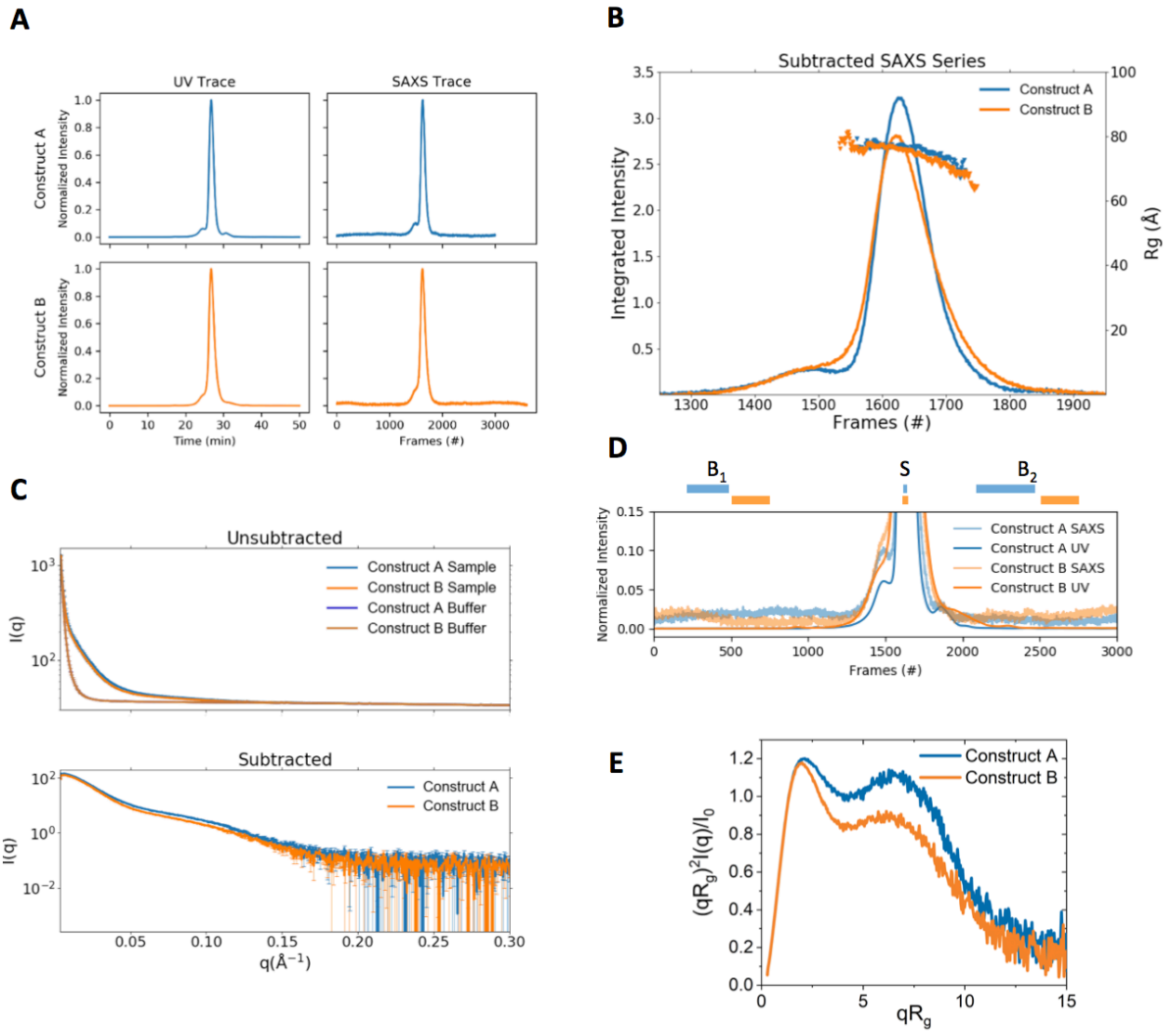
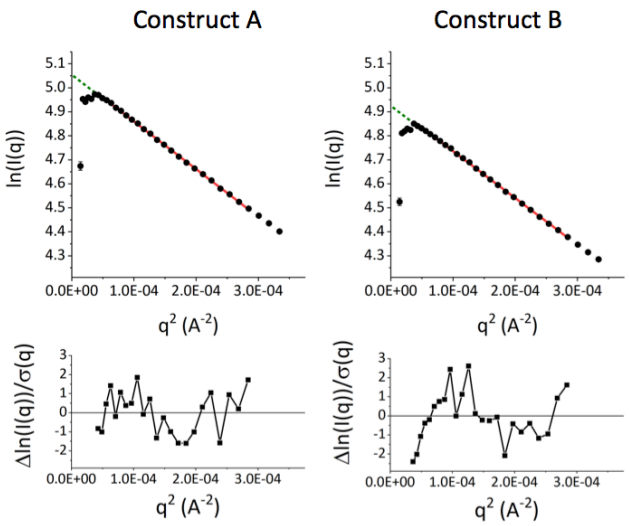


Figure S2: SEC-SAXS Analysis of Constructs *A* and *B*. (A) Normalized UV and SAXS traces of constructs *A* (blue) and *B* (orange). SAXS traces are integrated over the full *q*-range. (B) Subtracted SAXS trace of the peak region. An aggregate-shoulder fronts the main peaks of both constructs. R_g estimates are mostly frame-independent at the peak center, but rise and fall at the left and right edges, respectively. (C) Averaged scattering profiles of the frames selected as buffer and sample (top), and the subtracted intensities (bottom). Error bars indicate experimental errors (standard deviation). (D) Aligned overlay of the SAXS and UV traces, highlighting additional peaks of low intensity only visible in the UV trace. The UV trace was shifted to account for the detector-detector distance within the FPLC system. Colored bars indicate regions averaged for the analysis. B_1, buffer region 1; S, sample region; B_2, buffer region 2 (E) Normalized Kratky-plots of the constructs.

A



B

	A	B	Method
D_{max} (Å)	268 ± 4	278 ± 5	
R_g (Å)	80.2 ± 0.1	79.1 ± 0.1	BIFT
I_0	160.0 ± 1	140 ± 1	
D_{max} (Å)	283	286	
R_g (Å)	80.7 ± 0.1	79.3 ± 0.2	GNOM
I_0	159 ± 1	140 ± 1	
R_g (Å)	77.2 ± 0.2	76.6 ± 0.2	Guinier
I_0	157 ± 1	139 ± 1	
M_W (kDa)	403	484	Chem. Comp.
M_W (kDa)	408	488	Vc
M_W (kDa)	494	558	Vp

C

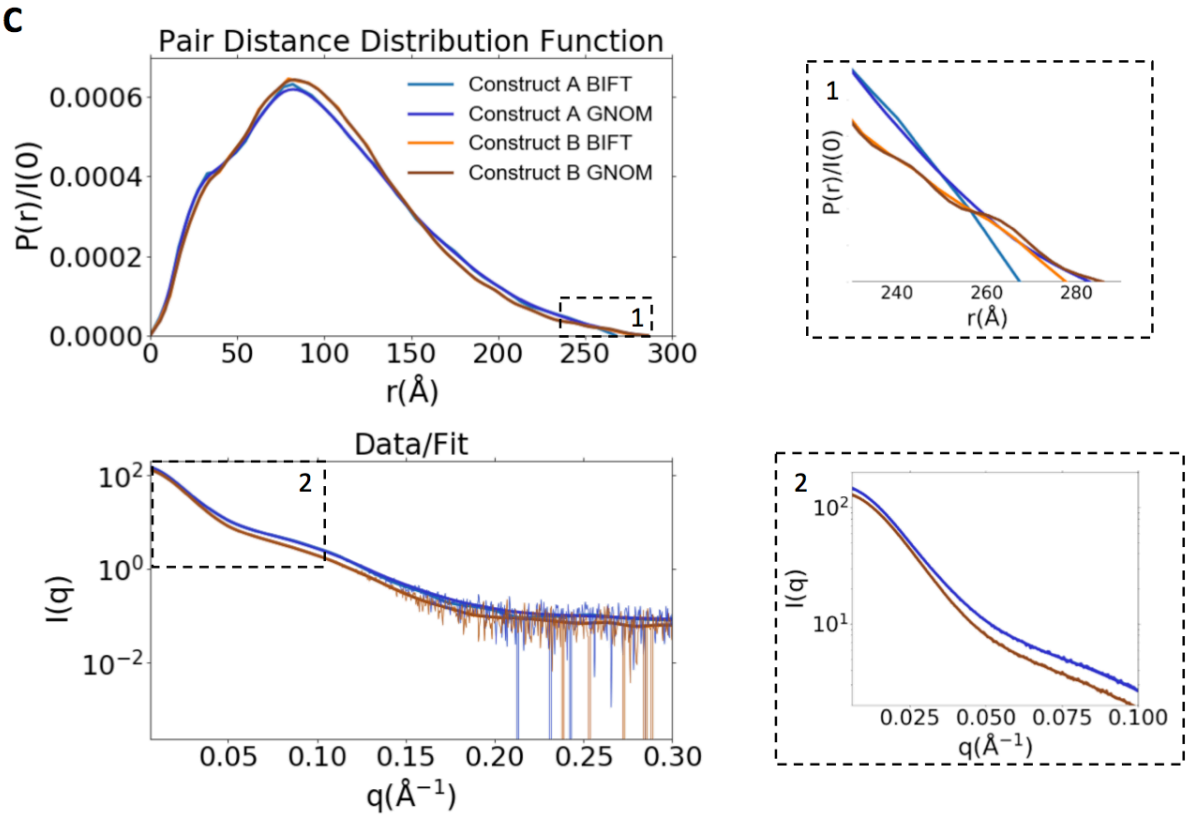


Figure S3. Guinier Fits and Pair Distance Distributions (A) Guinier plots of the subtracted sample averages. The fit is displayed in red with the I_0 approximation as a green dashed line. The distribution of normalized residuals for the regressions are shown below. For construct *B*, the initial downward stretch of the residuals hints toward the presence of mild aggregation (B) Summary of the values extracted from the Guinier analysis and the *P(r)* function together with the corresponding M_W estimates. (C) *P(r)* functions of constructs *A* and *B*, determined both via BIFT and GNOM (top). The curves generated with BIFT show a steep descent at high *r*, whereas the GNOM solution for construct *A* shows an uneven decay (inset 1). All computed *P(r)* distributions fit the scattering data (bottom), especially in the Porod-region (inset 2).

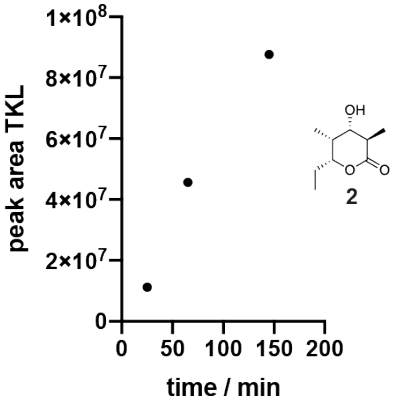


Figure S4. Triketide lactone formation of construct *B* (KR1-ACP1-M2-TE). The activity of construct *B* (in detail, the domains of the M2-TE part) was measured using 4 µM of enzyme and saturating (non-rate-limiting) concentrations of (2*S*,3*R*)-2-methyl-3-hydroxy-diketide-SNAC, methylmalonyl-CoA and NADPH. In this assay, the diketide is directly loaded in the KS domain and then elongated with methylmalonyl. Three reactions were set up in parallel (same protein batch) and quenched at different time points (25, 65 and 145 min, respectively). A constant increase of triketide lactone **2** over a 145 min time course was confirmed by LC-MS. At time points 25, 65 and 145 min, an aliquot of the reaction solution was taken for determining turnover rates at estimated substrates saturating conditions, essentially following a protocol as described previously [6]. Turnover rates were determined to 0.32 min^{-1}, 0.36 min^{-1} and 0.35 min^{-1} at 25, 65 and 145 min, respectively, which is about 10 % of the rates reported in by Li *et al* (ref. 6). A compromised turnover rate can be explained by the ACP1 interfering in substrate elongation. A similar effect of a second ACP interfering in turnover and comprising activity was reported before for murine FAS [47]. We note that construct *A* was not available for this assay, because of the missing TE domain.

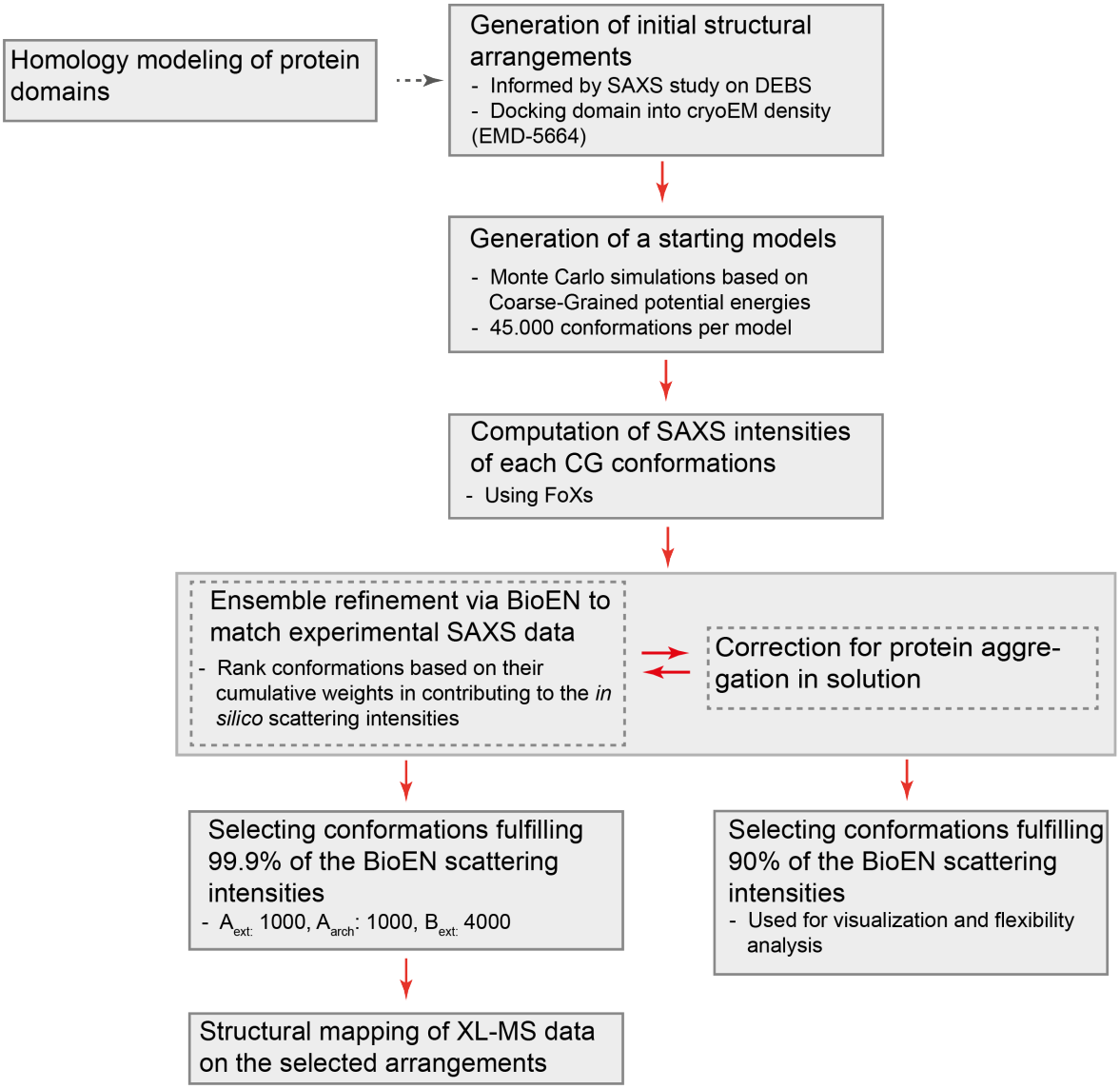


Figure S5. Scheme of the modeling approach. For details see Material & Methods.

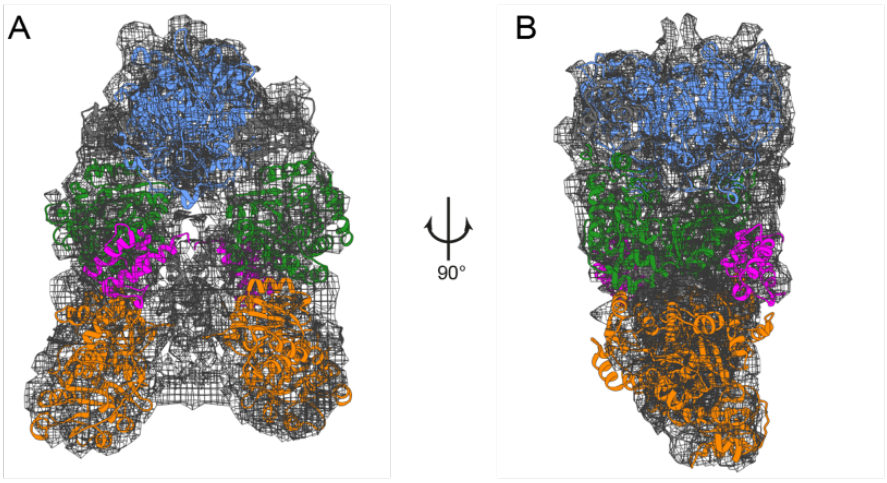


Figure S6. A_{arch} fitted within of the PikAIII complex cryoEM density (EMD-5664). Front (A) and lateral (B) views of the docked protein domains. Domain coloring: Density- light gray, KS- blue, LD- dark gray, AT- green, KR- orange, ACP- magenta, DD- dark gray.

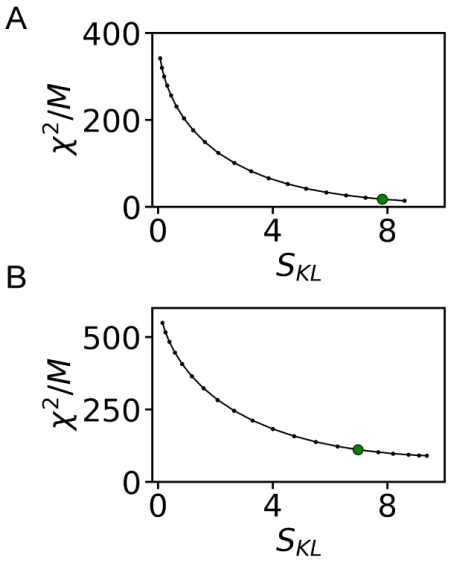


Figure S7. L-curve analysis for the SAXS measurements. Consistency between the simulated and experimental data (reduced chi-squared χ^2) plotted as a function of changes in weights based on relative entropy S_{KL} for A_{ext} (A) and A_{arch} (B). At θ values of $1.58*10^4$ and $3.98*10^4$ for A_{ext} and A_{arch}, respectively, the S_{KL} and reduced χ^2 are equal to 7.8 and 17.6 for A_{ext} and 7.0 and 111.1 for A_{arch}. respectively (green dot).

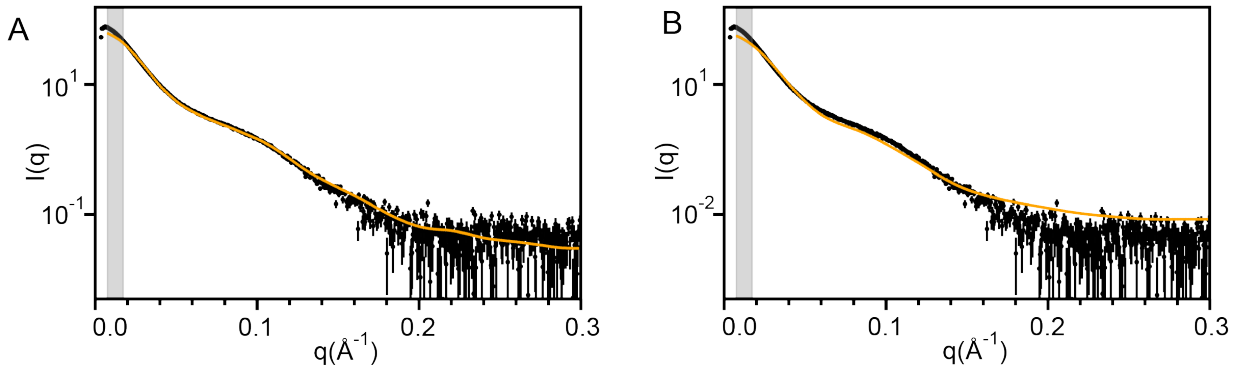


Figure S8. BioEn ensemble refinement without protein aggregation. Experimental (black) and BioEn (orange) scattering intensities for A_{ext} (A) and A_{arch} (B) as $I(q)$ versus q obtained without accounting for protein aggregates. The Guinier region used to compute R_g is highlighted in gray.

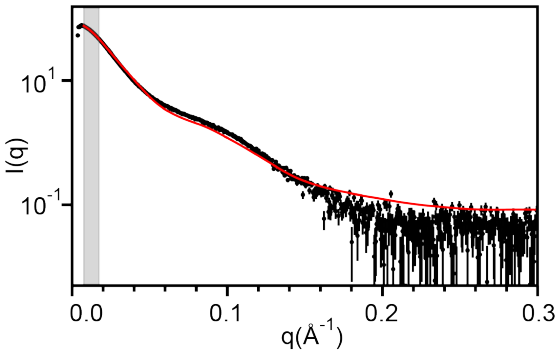


Figure S9. A higher amount of protein aggregation improves the fit for A_{arch} in the Guinier region, but not beyond. Fitting of the BioEn scattering curve as $I(q)$ versus q of A_{arch} (red line) on the experimental data (black points) corrected for protein aggregation in solution equal to 10.1%. The Guinier region used to compute R_g is highlighted in gray.

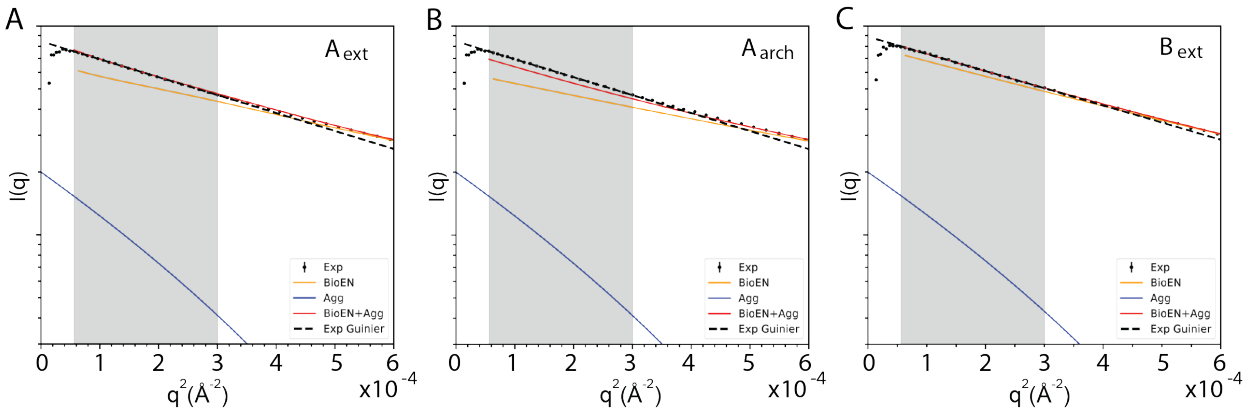


Figure S10. The models with aggregate correction match the experimental intensities at the Guinier region. Experimental (black points), simulated without aggregation (orange), aggregate (blue) and simulated in presence of aggregation (red curve) intensities at the Guinier region as ln(I) vs q^2 for A_{ext} (A), A_{arch} (B) and B_{ext} (C). The Guinier fittings for the experimental intensities of both constructs are shown as a black dashed line.

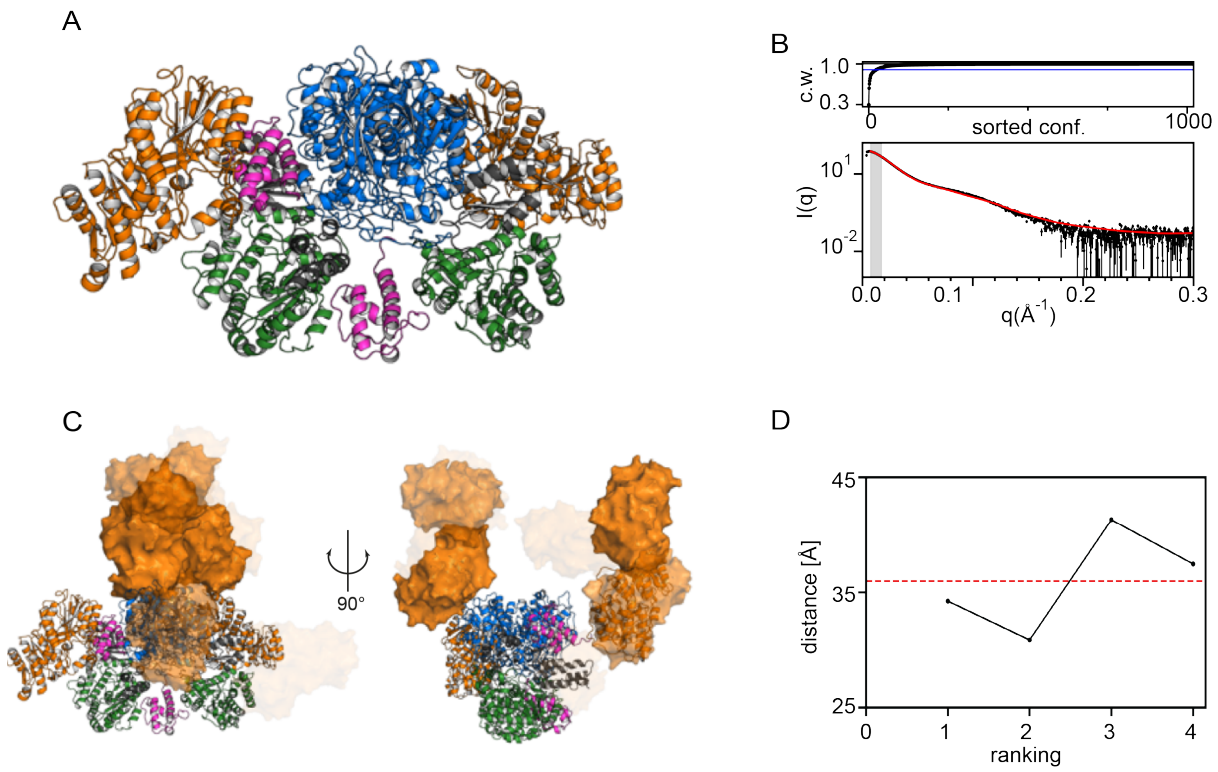


Figure S11. A_{arch}^{598} disrupts the arched architecture. (A) Conformation of A_{arch}^{598} ranked at highest statistical weight; (B) Cumulative weights (c. w.) of the weight-sorted conformations fulfilling 90% (blue line) and 100% (black line) of the fittings and computed (red) scattering intensities shown as $I(q)$ versus q (bottom). The Guinier region used to compute R_g is highlighted in gray (a protein aggregation of 7.1 % is included). (C) Model conformations that explain 90% of the BioEn scattering curve. (D) KR distances (black points) across the 5 ranked conformations that explain 90% of the BioEn scattering curve. The average distance is equal to 36.0 Å (dashed red line). The conformation with highest statistical weight (ranked #0) was used as reference. Also see Note 1&2. Domain coloring: KS2- blue, LD- dark gray, AT- green, KR- orange, ACP- magenta, DD- dark gray.

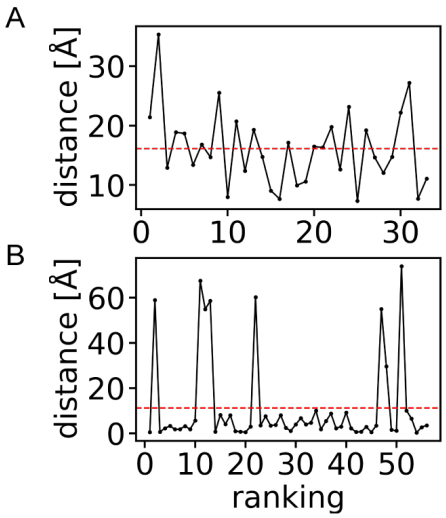


Figure S12. KR flexibility of the conformations ranked at highest weight. KR distances (black points) computed across the 33 and 56 conformations ranked at highest weight and explaining 90% of the fitting of A_{ext} (A) and A_{arch} (B), respectively. Average distances (dashed red line) are equal to 16.1 Å (A_{ext}) and 11.2 Å (A_{arch}). The average distance computed for A_{arch} drops to 3.9 Å when considering the 48 conformations that maintain the fully arched architecture and contribute together to the 71% of the fitting. The conformations with highest statistical weight (ranked #0) were used as reference.

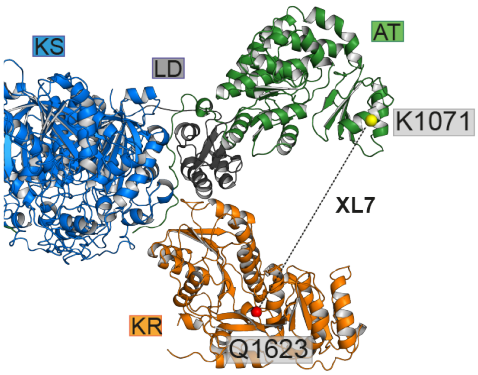


Figure S13. Residues involved in XL7. A_{ext} conformation at highest statistical weight. Depiction of the domains KS-AT_2 (blue and green. respectively) and KR (orange) with the residues in XL7 (AT_{K1071} and KR_{Q1623}) highlighted in yellow and red. respectively.

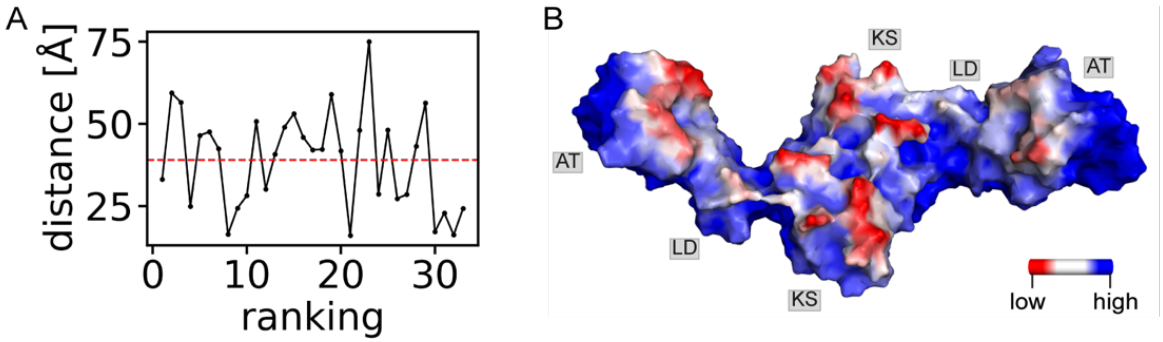


Figure S14. MBP interaction with KS-AT$_2$. (A) Distances computed for MBP (black points) over the 33 conformations ranked at highest weight that explain 90% of the fitting of A_{ext}. Average distance(dashed red line) is equal to 39.0 Å. The conformation with highest statistical weight (ranked #0) was used as reference. (B) Depiction of the KS-AT$_2$ surface with MBP interaction hot spots highlighted. Frequencies of interactions range from low (blue). medium (white). and high frequencies (red).

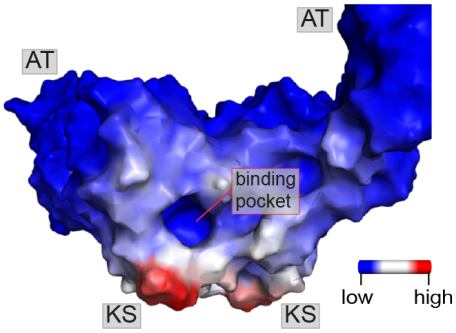


Figure S15. ACP samples the entrance of the KS binding pocket. Depiction of the KS-AT$_2$ surface in a lateral orientation. Spectrum of frequencies of interaction across ACP$_{K1775}$ and ACP$_{K1784}$ and the solvent exposed residues of KS-AT$_2$. The frequencies range from low (blue). medium (white). and high (red) values.

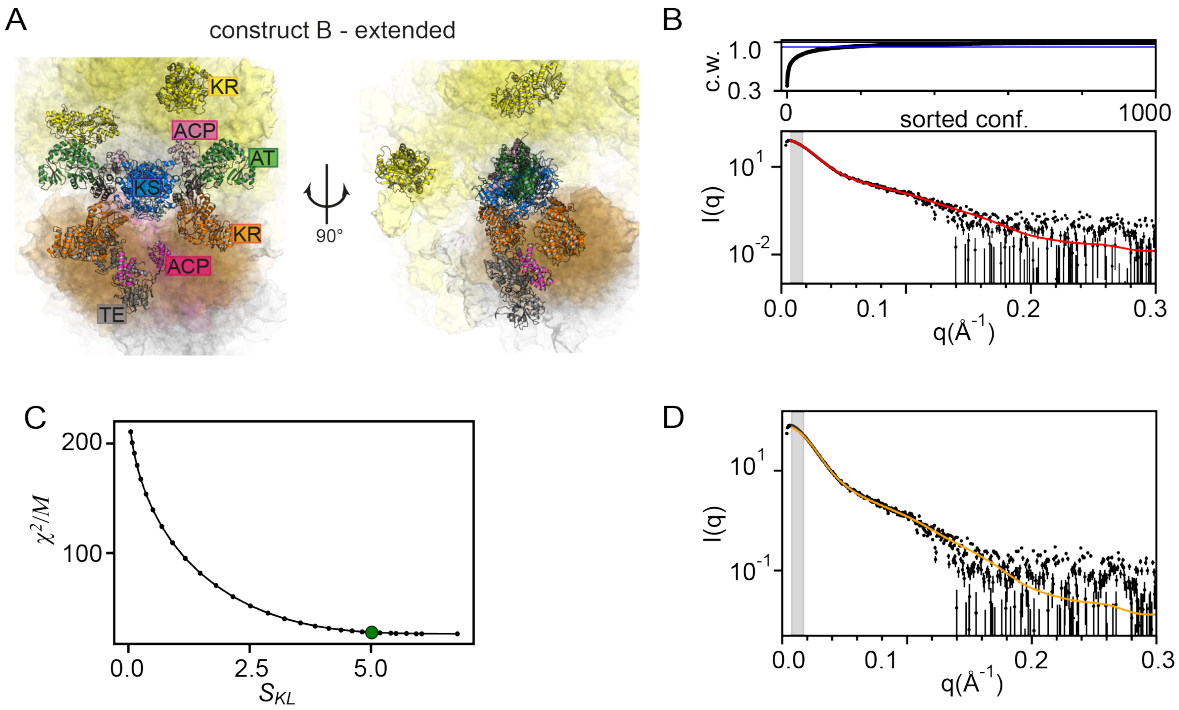


Figure S16. B_{ext} agrees with SAXS experiment. (A) Model conformations at higher statistical weight that explain the scattering fitting of B_{ext}. Domain coloring: KR1- yellow. ACP- light magenta. KS2- blue, AT2- green, KR2- orange, ACP2- magenta, TE- gray. (B) Cumulative weights (c. w.) of the weight-sorted conformations fulfilling 90% (blue line) and 100% (black line) of the fittings. (C) Experimental (black) and computed (red) scattering intensities as $I(q)$ versus q that account for protein aggregation. (D) L-curve analysis for the SAXS measurements. Consistency between the simulated and experimental data (reduced chi-squared χ^2) plotted as a function of changes in weights based on relative entropy S_{KL} for B_{ext}. At θ value of $6.31*10^3$ the S_{KL} and reduced χ^2 are equal to 5.0 and 27.7, respectively (green dot). (E) BioEn ensemble refinement without protein aggregation. Experimental (black) and BioEn (orange) scattering intensities for B_{ext} as $I(q)$ versus q obtained without accounting for protein aggregates. The Guinier region used to compute R_g is highlighted in gray.

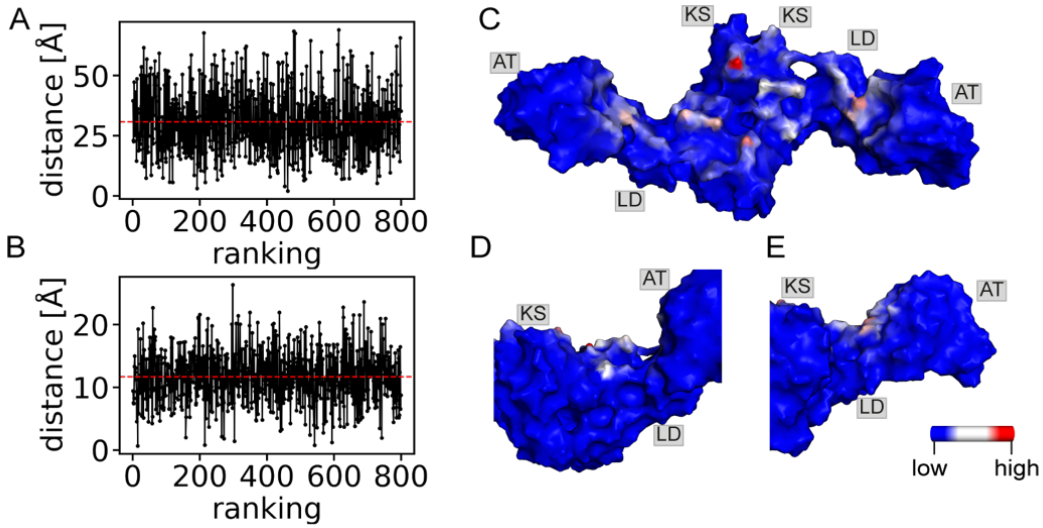


Figure S17. The N-terminal KR1 and ACP1 have different mobility in B_{ext}. Distances (black points) for KR1 (A) and ACP1 (B) computed over the 799 conformations ranked at highest weight that explain 90% of the fitting. Average distances (dashed red line) are equal to 30.7 and 11.7 Å for KR1 and ACP1, respectively. The conformations with highest statistical weight (ranked #0) were used as reference.(C) Depiction of the KS-AT_2 surface with ACP1 interaction hot spots highlighted. Frequencies of interactions monitored between KS and ACP1 (D) and between AT and ACP1 (E). Frequencies of interactions range from low (blue), medium (white), and high frequencies (red).

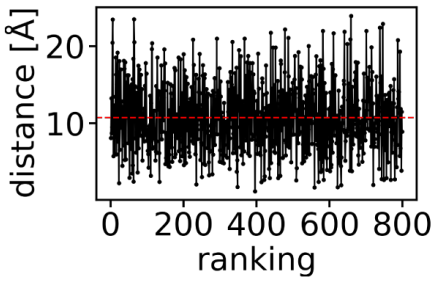


Figure S18. KR2 flexibility of the conformations ranked at highest weight for B_{ext}. Distances for KR2 (black points) computed across the 799 conformations ranked at highest weight and explaining 90% of the fitting of B_{ext}. The average distance (dashed red line) is equal to 10.7 Å.

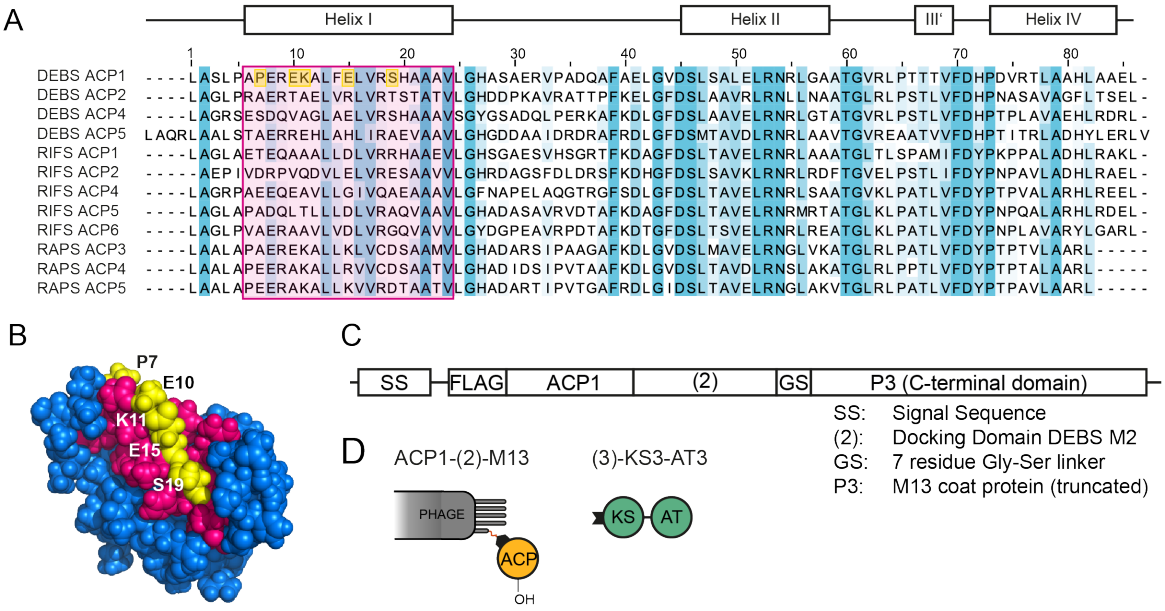


Figure S19. Detailed phage display setup. Design of the first ACP1-phagemid library. (A) Sequence alignment of ACPs. Helix I is highlighted in pink. position randomized for library generation are shown in yellow. Sequence identity highlighted in shades of blue. These residues were chosen as likely candidates to alter domain-domain interaction specificities, because they are located on the surface of ACP in the chain translocation epitope and are not conserved as suggested by the sequence alignment. (B) Homology model of ACP1. Surface residues selected for mutagenesis are highlighted. Color scheme as in (A). (C) Schematic representation of the phagemid architecture used for the library generation. Abbreviations as indicated. (D) Architecture of the ACP1 on phage and its target protein (3)KS3-AT3.

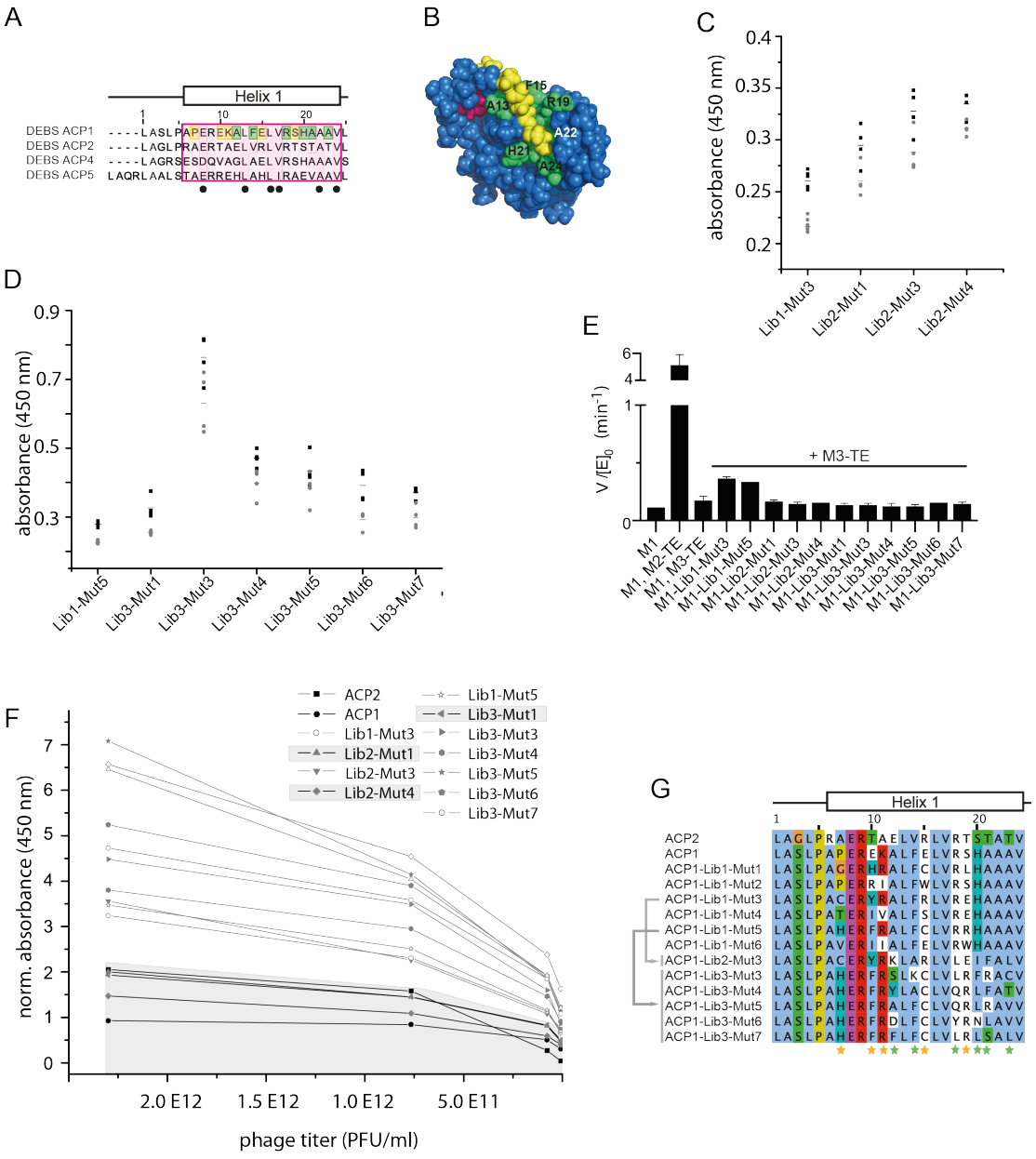


Figure S20. Design and analysis of two second generation ACP1 libraries. (A) Sequence alignment of ACPs highlighting positions selected for mutagenesis. Helix I in pink, position randomized for ACP1–Lib1 in yellow, and positions selected for a second round of directed evolution experiments are shown in green. Conserved positions are indicated with a dot (see alignment Fig. S17). (B) Homology model of ACP1. Surface residues selected for mutagenesis are highlighted. Color scheme as in (A). (C/D) ELISA of ACP1–Lib2 mutants (Lib2–Mut"X") compared to Lib1–Mut3 (C) and ACP1–Lib3 mutants (Lib3–Mut"X") compared to Lib1–Mut5 (D). Signal was obtained in (3)KS3–AT3 coated wells and the degree of unspecific binding was assessed by comparing it to the signal in BSA coated wells. For each ACP of Lib2–Mut"X" and Lib3–Mut"X", four individually grown phage cultures were tested, and the individual data points and data mean are indicated (eight data points and mean given for Lib1-Mut3 (C) and Lib1-Mut5 (D)). (E) Turnover rates of wild–type and chimeric bimodular PKSs comparing all ACP1 mutations. All bimodular PKS consisted of LM(4), (5)M1(2), and (3)Module"X"–TE. All initial rate data was obtained at individual PKS protein concentrations of 4 μM and non–limiting concentrations of propionyl–CoA. methylmalonyl–CoA. and NADPH. Measurements were

performed in triplicate. Either wild–type M1 was used as the first module (M1), or one of the ACP1 mutants enriched in the directed evolution experiments (M1–Lib"X"–Mut"X"). (F) Titration ELISA of wild–type and mutant ACPs presented on the phage surface. Data mean as straight line scatter from two individual plates coated with (3)KS3–AT3 are given (for data points, see Table S8). The ELISA signal was normalized to the amount of applied phage based on OD268 measurement (G) Sequence alignment of obtained mutants compared to wild–type ACP1 and ACP2. Randomized positions are indicated with a yellow (ACP1–Lib1) or green (ACP1–Lib2/3) asterisk.

Supporting Tables

Table S1. Alignment of construct *A* and *B* with DEBS M2 (Uniprot entry Q03131). Numbering of construct *A* and *B* as resulting from protein design. MBP domain for construct *A* and TE of construct *B* not shown.

```
                                  →KR1
construct_B             ------------MDEVSALRYRIEWRPTGAGEPARLDGTWLVAKYAGTADETSTAAREA  47
Q03131|ERYA1_SACER      WLEPKPVARRSTEVDEVSALRYRIEWRPTGAGEPARLDGTWLVAKYAGTADETSTAAREA  1440
construct_A             ------------------------------------------------------------


construct_B             LESAGARVRELVVDARCGRDELAERLRSVGEVAGVLSLLAVDEAEPEEAPLALASLADTL  107
Q03131|ERYA1_SACER      LESAGARVRELVVDARCGRDELAERLRSVGEVAGVLSLLAVDEAEPEEAPLALASLADTL  1500
construct_A             ------------------------------------------------------------


construct_B             SLVQAMVSAELGCPLWTVTESAVATGPFERVRNAAHGALWGVGRVIALENPAVWGGLVDV  167
Q03131|ERYA1_SACER      SLVQAMVSAELGCPLWTVTESAVATGPFERVRNAAHGALWGVGRVIALENPAVWGGLVDV  1560
construct_A             ------------------------------------------------------------


construct_B             PAGSVAELARHLAAVVSGGAGEDQLALRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGG  227
Q03131|ERYA1_SACER      PAGSVAELARHLAAVVSGGAGEDQLALRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGG  1620
construct_A             ------------------------------------------------------------


construct_B             TGGVGGQIARWLARRGAPHLLLVSRSGPDADGAGELVAELEALGARTTVAACDVTDRESV  287
Q03131|ERYA1_SACER      TGGVGGQIARWLARRGAPHLLLVSRSGPDADGAGELVAELEALGARTTVAACDVTDRESV  1680
construct_A             ------------------------------------------------------------


construct_B             RELLGGIGDDVPLSAVFHAAATLDDGTVDTLTGERIERASRAKVLGARNLHELTRELDLT  347
Q03131|ERYA1_SACER      RELLGGIGDDVPLSAVFHAAATLDDGTVDTLTGERIERASRAKVLGARNLHELTRELDLT  1740
construct_A             ------------------------------------------------------------

                                                              KR1←
construct_B             AFVLFSSFASAFGAPGLGGYAPGNAYLDGLAQQRRSDGLPATAVAWGTWAGSGMAEGPVA  407
Q03131|ERYA1_SACER      AFVLFSSFASAFGAPGLGGYAPGNAYLDGLAQQRRSDGLPATAVAWGTWAGSGMAEGAVA  1800
construct_A             ------------------------------------------------------------


construct_B             DRFRRHGVIEMPPETACRALQNALDRAEVCPIVIDVRWDRFLLAYTAQRPTRLFDEIDDA  467
Q03131|ERYA1_SACER      DRFRRHGVIEMPPETACRALQNALDRAEVCPIVIDVRWDRFLLAYTAQRPTRLFDEIDDA  1860
construct_A             ------------------------------------------------------------

                                      →ACP1
construct_B             RRAAPQAAAEPRVGAHMLASLPAPEREKALFELVRSHAAAVLGHASAERVPADQAFAELG  527
Q03131|ERYA1_SACER      RRAAPQAPAEPRVGA--LASLPAPEREEALFELVRSHAAAVLGHASAERVPADQAFAELG  1918
construct_A             ------------------------------------------------------------

                                                      ACP1←
construct_B             VDSLSALELRNRLGAATGVRLPTTTVFDHPDVRTLAAHLAAELGGATGAEQAAPATTAPV  587
Q03131|ERYA1_SACER      VDSLSALELRNRLGAATGVRLPTTTVFDHPDVRTLAAHLAAELGGATGAEQAAPATTAPV  1978
construct_A             ------------------------------------(MBP not shown)AAPATTAPV  400
                                                                         *********
                          →KS2
construct_B             DEPIAIVGMACRLPGEVDSPERLWELITSGRDSAAEVPDDRGWVPDELMASDAAGTRRAH  647
Q03131|ERYA1_SACER      DEPIAIVGMACRLPGEVDSPERLWELITSGRDSAAEVPDDRGWVPDELMASDAAGT-RAH  2037
construct_A             DEPIAIVGMACRLPGEVDSPERLWELITSGRDSAAEVPDDRGWVPDELMASDAAGTRRAH  460
                        ************************************************************ ***

construct_B             GNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSDTGVFVG  707
Q03131|ERYA1_SACER      GNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSDTGVFVG  2097
construct_A             GNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSDTGVFVG  520
                        ************************************************************

construct_B             MSHQGYATGRPRPEDGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTACSSSLVALHT  767
Q03131|ERYA1_SACER      MSHQGYATGRPRPEDGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTACSSSLVALHT  2157
construct_A             MSHQGYATGRPRPEDGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTACSSSLVALHT  580
                        ************************************************************

construct_B             ACGSLRDGDCGLAVAGGVSVMAGPEVFTEFSRQGALSPDGRCKPFSDEADGFGLGEGSAF  827
Q03131|ERYA1_SACER      ACGSLRDGDCGLAVAGGVSVMAGPEVFTEFSRQGALSPDGRCKPFSDEADGFGLGEGSAF  2217
construct_A             ACGSLRDGDCGLAVAGGVSVMAGPEVFTEFSRQGALSPDGRCKPFSDEADGFGLGEGSAF  640
                        ************************************************************

construct_B             VVLQRLSDARREGRRVLGVVAGSAVNQDGASNGLSAPSGVAQQRVIRRAWARAGITGADV  887
Q03131|ERYA1_SACER      VVLQRLSDARREGRRVLGVVAGSAVNQDGASNGLSAPSGVAQQRVIRRAWARAGITGADV  2277
```

```
construct_A             VVLQRLSDARREGRRVLGVVAGSAVNQDGASNGLSAPSGVAQQRVIRRAWARAGITGADV  700
                        ************************************************************

construct_B             AVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQAAAGVAGVIKVL  947
Q03131|ERYA1_SACER      AVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQAAAGVAGVIKVL  2337
construct_A             AVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQAAAGVAGVIKVL  760
                        ************************************************************

construct_B             LGLERGVVPPMLCRGERSGLIDWSSGEIELADGVREWSPAADGVRRAGVSAFGVSGTNAH  1007
Q03131|ERYA1_SACER      LGLERGVVPPMLCRGERSGLIDWSSGEIELADGVREWSPAADGVRRAGVSAFGVSGTNAH  2397
construct_A             LGLERGVVPPMLCRGERSGLIDWSSGEIELADGVREWSPAADGVRRAGVSAFGVSGTNAH  820
                        ************************************************************
                         KS2←→LD2(linker domain)
construct_B             VIIAEPPEPEPVPQPRRMLPATGVVPVVLSARTGAALRAQAGRLADHLAAHPGIAPADVS  1067
Q03131|ERYA1_SACER      VIIAEPPEPEPVPQPRRMLPATGVVPVVLSARTGAALRAQAGRLADHLAAHPGIAPADVS  2457
construct_A             VIIAEPPEPEPVPQPRRMLPATGVVPVVLSARTGAALRAQAGRLADHLAAHPGIAPADVS  880
                        ************************************************************
                                                    LD2(linker domain)←→AT2
construct_B             WTMARARQHFEERAAVLAADTAEAVHRLRAVADGAVVPGVVTGSASDGGSVFVFPGQGAQ  1127
Q03131|ERYA1_SACER      WTMARARQHFEERAAVLAADTAEAVHRLRAVADGAVVPGVVTGSASDGGSVFVFPGQGAQ  2517
construct_A             WTMARARQHFEERAAVLAADTAEAVHRLRAVADGAVVPGVVTGSASDGGSVFVFPGQGAQ  940
                        ************************************************************

construct_B             WEGMARELLPVPVFAESIAECDAVLSEVAGFSVSEVLEPRPDAPSLERVDVVQPVLFAVM  1187
Q03131|ERYA1_SACER      WEGMARELLPVPVFAESIAECDAVLSEVAGFSVSEVLEPRPDAPSLERVDVVQPVLFAVM  2577
construct_A             WEGMARELLPVPVFAESIAECDAVLSEVAGFSVSEVLEPRPDAPSLERVDVVQPVLFAVM  1000
                        ************************************************************

construct_B             VSLARLWRACGAVPSAVIGHSQGEIAAAVVAGALSLEDGMRVVARRSRAVRAVAGRGSML  1247
Q03131|ERYA1_SACER      VSLARLWRACGAVPSAVIGHSQGEIAAAVVAGALSLEDGMRVVARRSRAVRAVAGRGSML  2637
construct_A             VSLARLWRACGAVPSAVIGHSQGEIAAAVVAGALSLEDGMRVVARRSRAVRAVAGRGSML  1060
                        ************************************************************

construct_B             SVRGGRSDVEKLLADDSWTGRLEVAAVNGPDAVVVAGDAQAAREFLEYCEGVGIRARAIP  1307
Q03131|ERYA1_SACER      SVRGGRSDVEKLLADDSWTGRLEVAAVNGPDAVVVAGDAQAAREFLEYCEGVGIRARAIP  2697
construct_A             SVRGGRSDVEKLLADDSWTGRLEVAAVNGPDAVVVAGDAQAAREFLEYCEGVGIRARAIP  1120
                        ************************************************************

construct_B             VDYASHTAHVEPVRDELVQALAGITPRRAEVPFFSTLTGDFLDGTELDAGYWYRNLRHPV  1367
Q03131|ERYA1_SACER      VDYASHTAHVEPVRDELVQALAGITPRRAEVPFFSTLTGDFLDGTELDAGYWYRNLRHPV  2757
construct_A             VDYASHTAHVEPVRDELVQALAGITPRRAEVPFFSTLTGDFLDGTELDAGYWYRNLRHPV  1180
                        ************************************************************

construct_B             EFHSAVQALTDQGYATFIEVSPHPVLASSVQETLDDAESDAAVLGTLERDAGDADRFLTA  1427
Q03131|ERYA1_SACER      EFHSAVQALTDQGYATFIEVSPHPVLASSVQETLDDAESDAAVLGTLERDAGDADRFLTA  2817
construct_A             EFHSAVQALTDQGYATFIEVSPHPVLASSVQETLDDAESDAAVLGTLERDAGDADRFLTA  1240
                        ************************************************************
                                                    AT2←          →KR2
construct_B             LADAHTRGVAVDWEAVLGRAGLVDLPGYPFQGKRFWLLPDRTTPRDELDGWFYRVDWTEV  1487
Q03131|ERYA1_SACER      LADAHTRGVAVDWEAVLGRAGLVDLPGYPFQGKRFWLLPDRTTPRDELDGWFYRVDWTEV  2877
construct_A             LADAHTRGVAVDWEAVLGRAGLVDLPGYPFQGKRFWLLPDRTTPRDELDGWFYRVDWTEV  1300
                        ************************************************************

construct_B             PRSEPAALRGRWLVVVPEGHEEDGWTVEVRSALAEAGAEPEVTRGVGGLVGDCAGVVSLL  1547
Q03131|ERYA1_SACER      PRSEPAALRGRWLVVVPEGHEEDGWTVEVRSALAEAGAEPEVTRGVGGLVGDCAGVVSLL  2937
construct_A             PRSEPAALRGRWLVVVPEGHEEDGWTVEVRSALAEAGAEPEVTRGVGGLVGDCAGVVSLL  1360
                        ************************************************************

construct_B             ALEGDGAVQTLVLVRELDAEGIDAPLWTVTFGAVDAGSPVARPDQAKLWGLGQVASLERG  1607
Q03131|ERYA1_SACER      ALEGDGAVQTLVLVRELDAEGIDAPLWTVTFGAVDAGSPVARPDQAKLWGLGQVASLERG  2997
construct_A             ALEGDGAVQTLVLVRELDAEGIDAPLWTVTFGAVDAGSPVARPDQAKLWGLGQVASLERG  1420
                        ************************************************************

construct_B             PRWTGLVDLPHMPDPELRGRLTAVLAGSEDQVAVRADAVRARRLSPAHVTATSEYAVPGG  1667
Q03131|ERYA1_SACER      PRWTGLVDLPHMPDPELRGRLTAVLAGSEDQVAVRADAVRARRLSPAHVTATSEYAVPGG  3057
construct_A             PRWTGLVDLPHMPDPELRGRLTAVLAGSEDQVAVRADAVRARRLSPAHVTATSEYAVPGG  1480
                        ************************************************************

construct_B             TILVTGGTAGLGAEVARWLAGRGAEHLALVSRRGPDTEGVGDLTAELTRLGARVSVHACD  1727
Q03131|ERYA1_SACER      TILVTGGTAGLGAEVARWLAGRGAEHLALVSRRGPDTEGVGDLTAELTRLGARVSVHACD  3117
construct_A             TILVTGGTAGLGAEVARWLAGRGAEHLALVSRRGPDTEGVGDLTAELTRLGARVSVHACD  1540
                        ************************************************************

construct_B             VSSREPVRELVHGLIEQGDVVRGVVHAAGLPQQVAINDMDEAAFDEVVAAKAGGAVHLDE  1787
Q03131|ERYA1_SACER      VSSREPVRELVHGLIEQGDVVRGVVHAAGLPQQVAINDMDEAAFDEVVAAKAGGAVHLDE  3177
construct_A             VSSREPVRELVHGLIEQGDVVRGVVHAAGLPQQVAINDMDEAAFDEVVAAKAGGAVHLDE  1600
                        ************************************************************

construct_B             LCSDAELFLLFSSGAGVWGSARQGAYAAGNAFLDAFARHRRGRGLPATSVAWGLWAAGGM  1847
Q03131|ERYA1_SACER      LCSDAELFLLFSSGAGVWGSARQGAYAAGNAFLDAFARHRRGRGLPATSVAWGLWAAGGM  3237
construct_A             LCSDAELFLLFSSGAGVWGSARQGAYAAGNAFLDAFARHRRGRGLPATSVAWGLWAAGGM  1660
```

```
                         ************************************************************

construct_B              TGDEEAVSFLRERGVRAMPVPRALAALDRVLASGETAVVVTDVDWPAFAESYTAARPRPL  1907
Q03131|ERYA1_SACER       TGDEEAVSFLRERGVRAMPVPRALAALDRVLASGETAVVVTDVDWPAFAESYTAARPRPL  3297
construct_A              TGDEEAVSFLRERGVRAMPVPRALAALDRVLASGETAVVVTDVDWPAFAESYTAARPRPL  1720
                         ************************************************************
                               KR2←                →ACP2
construct_B              LDRIVTTAPSERAGEPETESLRDRLAGLPRAERTAELVRLVRTSTATVLGHDDPKAVRAT  1967
Q03131|ERYA1_SACER       LDRIVTTAPSERAGEPETESLRDRLAGLPRAERTAELVRLVRTSTATVLGHDDPKAVRAT  3357
construct_A              LDRIVTTAPSERAGEPETESLRDRLAGLPRAERTAELVRLVRTSTATVLGHDDPKAVRAT  1780
                         ************************************************************
                                                                ACP2←→DD2
construct_B              TPFKELGFDSLAAVRLRNLLNAATGLRLPSTLVFDHPNASAVAGFLDAELSGTPAREASS  2027
Q03131|ERYA1_SACER       TPFKELGFDSLAAVRLRNLLNAATGLRLPSTLVFDHPNASAVAGFLDAELGTEVRGEAPS  3417
construct_A              TPFKELGFDSLAAVRLRNLLNAATGLRLPSTLVFDHPNASAVAGFLDAELGTEVRGEAPS  1840
                         ************************************************.     ** *
                                                          DD2←
construct_B              AL(TE6 not shown)-------------------------------------------  2029
Q03131|ERYA1_SACER       ALAGLDALEGALPEVPATEREELVQRLERMLAALRPVAQAADASGTGANPSGDDLGEAGV  3477
construct_A              ALAGLDALEAALPEVPATEREELVQRLERMLAALRPVAQAADASGTGANPSGDDLGEAGV  1900
                         **

construct_B              --------------
Q03131|ERYA1_SACER       DELLEALGRELDGD     3491
construct_A              DELLEALGRELDGD     1914
```

Table S2. Gene sequence of M2 before and after harmonization

Construct	DNA sequence
M2 *S. erythraea*	GAGCCGATCGCGATCGTCGGCATGGCGTGCCGGCTGCCCGGGGAGGTCGACTCCCCGGAGCGGCTGTGGGAGCTGATCACCTCGGGA CGCGACTCCGCGGCGGAGGTCCCCGATGACCGGGGCTGGGTCCCCGACGAGCTGATGGCCTCCGACGCGGCGGGAACCCGCCGCGCC CACGGCAACTTCATGGCGGGCGCCGGTGACTTCGACGCGGCGTTCTTCGGGATCTCGCCGCGCGAGGCGCTGGCGATGGACCCGCAG CAGCGCCAGGCGCTGGAGACGACGTGGGAGGCGCTGGAAAGCGCGGGCATCCCACCGGAGACGTTGCGCGGCAGCGACACCGGCGTG TTCGTCGGCATGTCCCACCAGGGCTACGCGACCGGGCGTCCGCGCCCGGAGGACGGCGTCGACGGGTACCTGCTCACCGGCAACACC GCGAGCGTCGCGTCGGGACGCATCGCCTACGTGCTGGGGCTGGAAGGTCCCGCGCTGACGGTGGACACGGCGTGTTCGTCGTCGTTG GTGGCGTTGCACACGGCGTGTGGGTCGTTGCGTGACGGTGACTGCGGTCTTGCGGTGGCCGGTGGTGTGTCGGTGATGGCGGGTCCG GAGGTGTTCACCGAGTTCTCCCGCCAGGGCGCGCTCTCGCCGGACGGCCGGTGCAAGCCCTTCTCGGACGAGGCCGACGGATTCGGT CTCGGGGAGGGTTCGGCGTTCGTCGTGCTCCAGCGGTTGTCCGACGCCAGGCGGGAGGGCCGCCGCGTGCTCGGCGTGGTGGCCGGG TCCGCGGTGAACCAGGACGGCGCGAGCAACGGGCTCTCCGCTCCGAGCGGCGTCGCGCAGCAGCGGGTCATCCGCCGGGCGTGGGCG CGTGCGGGGATCACGGGCGCGGATGTGGCCGTGGTGGAGGCGCATGGGACCGGTACGCGGCTGGGCGATCCGGTGGAGGCGTCGGCG TTGCTGGCTACTTACGGCAAGTCGCGCGGGTCGTCGGGCCCGGTGCTGCTGGGTTCGGTGAAGTCGAACATCGGTCACGCGCAGGCG GCCGCGGGTGTCGCGGGCGTGATCAAGGTGCTGCTCGGCCTGGAACGCGGTGTGGTGCCCCCGATGCTGTGCCGGGGCGAGAGGTCG GGCCTCATCGACTGGTCCTCCGGCGAGATCGAGCTCGCAGACGGCGTGCGGGAGTGGTCGCCCGCCGCGGACGGGGTGCGCCGGGCA GGTGTGTCGGCGTTCGGGGTGAGCGGGACGAACGCGCACGTGATCATCGCCGAGCCGCCGGAACCGGAGCCCGTGCCGCAACCGCGA CGCATGCTGCCCGCGACCGGGGTGGTGCCGGTCGTGCTGTCGGCCAGGACCGGGGCGGCGTTGCGGGCGCAGGCCGGCAGGCTCGCC GACCACCTCGCCGCGCATCCCGGGATCGCACCGGCCGACGTGAGCTGGACGATGGCGCGGGCCCGCCAGCACTTCGAGGAGCGGGCC GCGGTGCTCGCCGCCGACACCGCCGAGGCCGTGCACCGGTTGCGGGCGGTGGCCGACGGCGCGGTGGTTCCCGGTGTTGTCACCGGC AGTGCCTCCGACGGTGGTTCAGTGTTCGTCTTCCCTGGGCAGGGTGCCCAGTGGGAAGGCATGGCGCGGGAGTTGTTGCCGGTTCCC GTCTTCGCCGAGTCGATCGCCGAGTGCGATGCGGTGTTGTCGGAGGTGGCCGGATTCTCGGTGTCCGAGGTGCTGGAGCCACGTCCG GACGCGCCGTCGCTGGAGCGGGTCGACGTGGTGCAGCCGGTGCTGTTCGCGGTGATGGTGTCGCTGGCGCGGTTGTGGCGTGCCTGC GGTGCCGTTCCTTCGGCCGTCATAGGGCACTCGCAGGGTGAGATCGCCGCCGCGGTGGTGGCGGGAGCGTTGTCGCTGGAGGACGGC ATGCGCGTCGTCGCCCGCCGGTCGAGGGCGGTGCGTGCGGTCGCGGGCCGGGGGAGCATGCTCTCGGTGCGCGGCGGCCGCTCCGAC GTCGAGAAGCTGCTCGCCGACGACAGCTGGACCGGCAGGCTGGAGGTCGCCGCGGTCAACGGCCCCGACGCCGTGGTGGTGGCCGGT GACGCCCAGGCGGCGCGCGAGTTCCTGGAGTACTGCGAGGGCGTGGGCATCCGCGCCCGCGCGATCCCGGTGGACTACGCCTCGCAC ACCGCGCACGTCGAGCCCGTGCGCGACGAACTGGTCCAGGCGCTGGCCGGGATCACCCCGCGACGGGCCGAGGTGCCGTTCTTCTCC ACCCTGACCGGCGACTTCCTCGACGGCACCGAGCTGGACGCGGGCTACTGGTACCGCAACCTGCGTCACCCGGTGGAGTTCCACTCC GCCGTGCAGGCGCTGACCGACCAGGGATACGCGACGTTCATCGAGGTCAGCCCGCACCCGGTGCTGGCGTCGAGCGTCCAGGAGACC CTCGACGACGCCGAGTCGGACGCGGCGGTGCTCGGGACGCTGGAACGCGACGCGGGCGACGCCGACCGCTTCCTCACGGCACTCGCC GACGCGCACACGCGCGGTGTCGCGGTCGACTGGGAAGCGGTGCTCGGCCGGGCCGGACTGGTCGACCTGCCGGGTTATCCTTTCCAG GGCAAGCGGTTCTGGCTGCTGCCGGACCGCACCACCCCTCGTGACGAGCTCGACGGCTGGTTCTACCGGGTCGACTGGACCGAGGTG CCGCGCTCCGAACCTGCCGCGCTGCGCGGCCGTTGGCTCGTGGTGGTGCCCGAGGGGCACGAGGAGGACGGCTGGACCGTCGAGGTG CGGTCCGCGCTCGCCGAGGCCGGCGCCGAACCGGAGGTCACGCGCGGCGTCGGCGGGCTGGTCGGTGACTGCGCGGGCGTGGTGTCG TTGCTCGCCCTCGAGGGCGATGGTGCGGTGCAAACCCTTGTGCTGGTGCGGGAACTCGACGCCGAGGGCATCGACGCGCCACTGTGG ACGGTCACCTTCGGCGCGGTCGACGCGGGCAGTCCGGTGGCCCGCCCGGACCAGGCGAAGCTGTGGGGGCTGGGCCAGGTCGCGTCC CTGGAACGCGGGCCCCGCTGGACCGGCCTCGTCGACCTGCCGCACATGCCGGACCCGGAACTGCGAGGCCGTCTCACCGCGGTGCTG GCCGGCTCGGAGGACCAGGTCGCGGTGCGCGCCGACGCCGTGCGTGCGCGGCGGCTTTCCCCCGCCCACGTCACCGCCACCTCGGAG TACGCGGTGCCGGGCGGCACAATCCTGGTCACCGGTGGCACCGCCGGCCTGGGCGCGGAGGTGGCCCGGTGGCTCGCCGGTCGCGGC GCCGAACACCTCGCGCTGGTCAGCAGGCGAGGCCCGGACACCGAGGGCGTCGGCGACCTGACCGCCGAGCTGACCCGGCTCGGCGCG CGGGTGTCGGTGCACGCGTGCGACGTCAGCAGCCGCGAACCGGTGAGGGAACTCGTGCACGGCCTGATCGAGCAGGGCGACGTCGTC CGCGGTGTGGTGCACGCGGCGGGACTGCCGCAGCAGGTCGCGATCAACGACATGGACGAGGCCGCCTTCGACGAGGTGGTCGCGGCC AAGGCCGGGGGCGCGGTGCACCTGGACGAGCTGTGCTCGGACGCCGAGCTGTTCCTGCTGTTCTCCTCCGGGGCCGGGGTGTGGGGA AGCGCCCGCCAGGGCGCCTACGCCGCGGGCAACGCGTTCCTGGACGCCTTCGCCCGGCACCGCCGGGGCCGCGGCCTGCCCGCCACG TCGGTGGCGTGGGGGCTGTGGGCGGCGGGCGGCATGACCGGCGACGAGGAGGCCGTGTCGTTCCTGCGCGAGCGCGGTGTGCGGGCG ATGCCCGTACCGCGCGCCCTCGCCGCCCTGGACAGGGTGCTGGCCTCCGGGGAGACGGCGGTGGTCGTGACGGACGTGGACTGGCCC GCCTTCGCCGAGTCCTACACCGCCGCCCGGCCCCGGCCGTTGCTCGACCGCATCGTCACGACCGCGCCGAGCGAGCGGGCCGGAGAA CCGGAGACGGAGAGCCTGCGCGACCGGCTGGCGGGTCTGCCGCGTGCCGAGCGGACGGCGGAGCTGGTGCGCCTGGTCCGCACCAGC ACCGCGACCGTGCTGGGCCACGACGACCCGAAGGCGGTGCGCGCGACCACGCCGTTCAAGGAGCTCGGGTTCGACTCGCTGGCGGCC GTCCGGCTGCGCAACCTGCTCAACGCGGCCACCGGGCTCCGCCTGCCGTCGACGCTGGTCTTCGACCACCCGAACGCCTCCGCGGTC GCCGGTTTCCTCGACGCCGAGCTC

Construct	DNA sequence
M2 harmonized	GAACCGATTGCAATTGTTGGTATGGCATGTCGTCTGCCGGGCGAAGTTGATAGCCCGGAACGTCTGTGGGAACTGATTACCAGCGGG CGTGATAGCGCAGCAGAAGTTCCGGACGATCGTGGTTGGGTTCCGGATGAACTGATGGCAAGCGATGCAGCAGGGACCCGTCGTGCA CATGGTAACTTTATGGCAGGTGCAGGGGATTTTGATGCAGCATTTTTTGGCATTAGCCCGCGTGAAGCACTGGCAATGGATCCGCAG CAGCGTCAGGCACTGGAAACAACATGGGAAGCACTGGAGAGCGCAGGTATTCCCCCGGAAACACTACGTGGTAGCGATACCGGTGTT TTTGTTGGTATGAGCCATCAGGGTTATGCAACCGGCAGGCCGCGTCCGGAAGATGGTGTTGATGGCTATCTGCTGACCGGTAACACC GCAAGCGTTGCAAGCGGGCGTATTGCCTATGTTCTGGGCCTGGAGGGGCCGGCACTGACAGTTGATACAGCAtgtAGCAGCAGCCTA GTTGCACTACATACAGCATGCGGCAGCctcAGGGATGGGGATTGTGGGctcGCAGTTGCAGGGGGGGTTAGCGTTATGGCAGGGCCG GAAGTTTTTACCGAATTTAGCCGTCAGGGTGCACTGAGCCCGGATGGTCGTTGTAAGCCGTTTAGCGATGAAGCAGATGGGTTTGGG CTGGGCGAAGGGAGCGCATTTGTTGTTCTGCAGCGTCTAAGCGATGCACGACGTGAAGGTCGTCGTGTTCTGGGTGTTGTTGCAGGC AGCGCAGTTAACCAGGATGGTGCAAGCAACGGCCTGAGCGCTCCGAGCGGTGTTGCACAGCAGCGTGTTATTCGTCGTGCATGGGCA AGGGCAGGCATTACAGGTGCAGACGTTGCAGTTGTTGAAGCACACGGCACCGGGACACGTCTGGGTGACCCGGTTGAAGCAAGCGCA CTACTGGCTACTTATGGTAAGAGCCGTGGCAGCAGCGGTCCGGTTCTGCTGGGGAGCGTTAAGAGCAACATTGGGCATGCACAGGCA GCAGCAGGGGTTGCAGGTGTTATTAAGGTTCTGCTGGGTCTGGAGCGTGGGGTTGTTCCGCCGATGCTGTGTCGTGGTGAACGAAGC GGTCTGATTGATTGGAGCAGCGGTGAAATTGAACTGGCTGATGGTGTTCGTGAATGGAGCCCGGCAGCAGATGGCGTTCGTCGTGCT GGGGTTAGCGCATTTGGCGTTAGCGGCACAAACGCACATGTTATTATTGCAGAACCGCCGGAGCCGGAACCGGTTCCGCAACCGAGG CGTATGCTGCCGGCAACCGGCGTTGTTCCGGTTGTTCTGAGCGCACGAACCGGCGCAGCACTACGTGCACAGGCAGGTCGACTGGCA GATCATCTGGCAGCACACCCGGGCATTGCTCCGGCAGATGTTAGCTGGACAATGGCACGTGCACGTCAGCATTTTGAAGAACGTGCA GCAGTTCTGGCAGCAGATACCGCAGAAGCAGTTCATCGTCTACGTGCAGTTGCAGATGGTGCAGTTGTACCGGGGGTAGTTACCGGT TCGGCAAGCGATGGGGGGTCGGTTTTTGTTTTTCCCGGCCAGGGGGCACAGTGGGAGGGTATGGCACGTGAActcCTACCGGTACCG GTTTTTGCAGAAAGCATTGCAGAATGTGACGCAGTTCTAAGCGAAGTTGCAGGGTTTAGCGTTAGCGAAGTTCTGGAACCCAGGCCG GATGCACCGAGCCTGGAACGTGTTGATGTTGTTCAGCCGGTTCTGTTTGCAGTTATGGTTAGCCTGGCACGTCTATGGAGGGCATGT GGGGCAGTACCCAGCGCAGTTATAGGCCATAGCCAGGGGGAAATTGCAGCAGCAGTTGTTGCAGGGGCACTAAGCCTGGAAGATGGT ATGCGTGTTGTTGCACGTCGTAGCCGAGCAGTTAGGGCAGTTGCAGGTCGTGGCAGCATGCTGAGCGTTCGTGGTGGTCGTAGCGAT GTTGAAAAGCTGCTGGCAGATGATAGCTGGACCGGTCGACTGGAAGTTGCAGCAGTTAACGGTCCGGATGCAGTTGTTGTTGCAGGG GATGCACAGGCAGCACGTGAATTTCTGGAATATTGTGAAGGTGTTGGTATTCGTGCACGTGCAATTCCGGTTGATTATGCAAGCCAT ACCGCACATGTTGAACCGGTTCGTGATGAGCTGGTTCAGGCACTGGCAGGCATTACCCCGAGGCGTGCAGAAGTTCCGTTTTTTAGC ACCCTGACCGGTGATTTTCTGGATGGTACCGAACTGGATGCAGGTTATTGGTATCGTAACCTGAGGCATCCGGTTGAATTTCATAGC GCAGTTCAGGCACTGACCGATCAGGGGTATGCAACATTTATTGAAGTTAGCCCGCATCCGGTTCTGGCAAGCAGCGTTCAGGAAACC CTGGATGATGCAGAAAGCGATGCAGCAGTTCTGGGCACACTGGAGCGTGATGCAGGTGATGCAGATCGTTTTCTGACAGCTCTGGCA GATGCACATACACGTGGGGTTGCAGTTGATTGGGAGGCAGTTCTGGGTCGTGCAGGGCTGGTTGATCTGCCGGGGTACCCCTTTCAG GGTAAGCGTTTTTGGCTGCTGCCGGATCGTACCACCCCCAGGGATGAACTGGATGGTTGGTTTTATCGTGTTGATTGGACCGAAGTT CCGCGTAGCGAGCCCGCAGCACTGCGTGGTAGGTGGCTGGTTGTTGTTCCGGAAGGCCATGAAGAAGATGGTTGGACCGTTGAAGTT CGTAGCGCACTGGCAGAAGCAGGTGCAGAGCCGGAAGTTACACGTGGTGTTGGTGGCCTGGTTGGGGATTGTGCAGGTGTTGTTAGC CTACTGGCACTGGAAGGTGACGGGGCAGTTCAAACCCTAGTTCTGGTTCGTGAGCTGGATGCAGAAGGTATTGATGCACCCCTGTGG ACAGTTACCTTTGGTGCAGTTGATGCAGGTTCGCCGGTTGCACGTCCGGATCAGGCAAAGCTGTGGGGCCTGGGTCAGGTTGCAAGC CTGGAGCGTGGCCCGCGTTGGACCGGTCTGGTTGATCTGCCGCACATGCCGGATCCGGAGCTGAGGGGTAGGCTGACCGCAGTTCTG GCAGGTAGCGAAGATCAGGTTGCAGTTCGTGCAGATGCAGTTAGGGCACGTCGTCTAAGCCCGGCACATGTTACCGCAACCAGCGAA TATGCAGTTCCGGGTGGTACTATTCTGGTTACCGGGGGTACCGCAGGTCTGGGTGCAGAAGTTGCACGTTGGCTGGCAGGGCGTGGT GCAGAGCATCTGGCACTGGTTAGCCGAAGGGGTCCGGATACCGAAGGTGTTGGTGATCTGACCGCAGAACTGACCCGTCTGGGTGCA CGTGTTAGCGTTCATGCATGTGATGTTAGCAGCCGTGAGCCGGTTCGAGAGCTGGTTCATGGTCTGATTGAACAGGGTGATGTTGTT CGTGGGGTTGTTCATGCAGCAGGGCTGCCGCAGCAGGTTGCAATTAACGATATGGATGAAGCAGCATTTGATGAAGTTGTTGCAGCA AAGGCAGGCGGTGCAGTTCATCTGGATGAACTGTGTAGCGATGCAGAACTGTTTCTGCTGTTTAGCAGCGGCGCAGGCGTTTGGGGG AGCGCACGTCAGGGTGCCTATGCAGCAGGTAACGCATTTCTGGATGCATTTGCACGTCATCGTCGTGGTCGTGGTCTGCCGGCAACA AGCGTTGCATGGGGCCTGTGGGCAGCAGGTGGTATGACCGGTGATGAAGAAGCAGTTAGCTTTCTGCGTGAACGTGGGGTTCGTGCA ATGCCGGTACCGCGTGCACTGGCAGCACTGGATCGAGTTCTGGCAAGCGGCGAAACAGCAGTTGTTGTTACAGATGTTGATTGGCCG GCATTTGCAGAAAGCTATACCGCAGCACGTCCGCGTCCGCTACTGGATCGTATTGTTACAACCGCACCGAGCGAACGTGCAGGGGAG CCGGAAACAGAAAGCCTGCGTGATCGTCTGGCAGGGCTGCCGAGGGCAGAACGTACAGCAGAACTGGTTCGTCTGGTTCGTACCAGC ACCGCAACCGTTCTGGGTCATGATGATCCGAAGGCAGTTCGTGCAACCACACCGTTTAAGGAACTGGGCTTTGATAGCCTGGCAGCA GTTCGTCTGCGTAACCTGCTGAACGCAGCAACCGGCCTGCGTCTGCCGAGCACACTGGTTTTTGATCATCCGAACGCAAGCGCAGTT GCAGGGTTTCTGGATGCAGAACTG

Table S3. Experimental scattering data derived parameters for Construct *A* and *B*. Values calculated from frame numbers 1618-1641 for construct *A* and 1619-1646 for construct *B*.

	Construct *A*	**Construct *B***
Protein ID	MK53	ADD01
Domain-Structure	MBP-M2-dd2	KR1-ACP1-M2-TE
Data Collection Parameters		
Beamline	BM29	BM29
Type of experiment	SEC-SAXS	SEC-SAXS
Wavelength (Å)	0.992	0.992
Detector distance (m)	2.87	2.87
q range ($Å^{-1}$)	0.005 - 0.028	0.006- 0.028
Exposure time (s)	1	1
Temperature (°C)	20	20
SEC Parameters		
SEC column	Superose 6 increase 10/300	Superose 6 increase 10/300
Amount loaded (nmol)	2.54	2.02
Flow rate (mL/min)	0.6	0.6
Ve (mL)	44.62	44.58
Structural Parameters		
Mw theoretical - (kDa)	403	484
Mw Vc / Vp - (kDa)	408 / 494	488 / 558
I(0) from Guinier	157 ± 1	139 ± 1
R_g from Guinier - (Å)	77.2 ± 0.2	76.6 ± 0.2
I(0) from P(r) GNOM/BIFT	160 ± 1 / 159 ± 1	140 ± 1 / 140 ± 1
R_g from P(r) - (Å) GNOM/BIFT	80.7 ± 0.2 / 80.2 ± 0.1	79.3 ± 0.2 / 79.1 ± 0.1
D_{max} from P(r) - (Å) GNOM/BIFT	283 / 268	286 / 278
Software		
Data processing	BioXtas RAW, PRIMUS, Python	BioXtas RAW, PRIMUS, Python

Table S4. Dimeric interfaces in construct *A* and *B*. Interfaces of docking domain (DD) and TE were calculated with atomic structures of protein databank files (pdb) as indicated [15, 17]. The interface of the KS dimer was calculated with the modeled KS-AT structure used in this study (see SI Material and Methods). Interfaces were calculated with the PDBePISA [48].

Construct *A*			Construct *B*		
Domain	Coordinates	Interface	Domain	Coordinates	Interface
KS-AT	Modeler	2410.8	KS-AT	Modeler	2410.8
DD*	1pzq	1628	TE	1kez	879.1

Table S5. Distances computed for the cross-links measured within single protein domains.

Domain	Residue 1	Residue 2	Distance
High frequency hits			
KS-AT_2	K727	A693	20.11
KS-AT_2	K727	D699	13.44
KS-AT_2	K727	V703	15.16
KS-AT_2	K727	S778	19.41
KS-AT_2	K727	D782	11.87
KS-AT_2	K727	S784	12.65
KS-AT_2	K727	S785	9.77
KS-AT_2	K727	G786	6.94
KS-AT_2	K727	E787	6.69
KS-AT_2	K727	E796	23.56
KS-AT_2	K727	G816	25.42
KS-AT_2	K741	V702	15.69
KS-AT_2	K1071	E1104	15.09
KS-AT_2	K1071	L1106	14.06
KS-AT_2	K1071	E1107	15.8
KS-AT_2	K1071	Y1108	12.94
KS-AT_2	K1273	V1251	34.67
KS-AT_2	K1273	E1254	35.47
KR2	K1407	A1396	16.87
KR2	K1407	G1397	18.59
KR2	K1407	P1399	16.05
KR2	K1407	A1401	12.97
KR2	K1407	I1576	13.64
KR2	K1407	D1578	16.79
KR2	K1407	M1579	13.21
KR2	K1407	D1580	12.87
KR2	K1407	A1583	13.26
KR2	K1407	F1584	9.77

Domain	Residue 1	Residue 2	Distance
KR2	K1407	D1585	10.15
KR2	K1407	E1586	13.62
KR2	K1407	V1587	13.3
ACP2	K1775	R1759	10.74
ACP2	K1784	L1808	18.39
ACP2	K1784	L1812	11.77
ACP2	K1784	F1814	8.37
ACP2	K1784	D1815	11.29
ACP2	K1784	H1816	9.89
ACP2	K1784	N1818	9.08
ACP2	K1784	A1819	11.16
ACP2	K1784	S1820	14.26
ACP2	K1784	A1821	13.75
ACP2	K1784	V1822	13.17
ACP2	K1784	E1829	22.1
Medium frequency hits			
$KS\text{-}AT_2$	K623	G525	28.21
$KS\text{-}AT_2$	K623	Y526	29.51
$KS\text{-}AT_2$	K623	G539	42.72
$KS\text{-}AT_2$	K623	Y540	40.51
$KS\text{-}AT_2$	K623	V598	22.74
$KS\text{-}AT_2$	K623	S599	20.82
$KS\text{-}AT_2$	K623	P624	3.75
$KS\text{-}AT_2$	K727	T696	12.36
$KS\text{-}AT_2$	K727	G697	9.57
$KS\text{-}AT_2$	K727	A698	9.85
$KS\text{-}AT_2$	K727	V700	13.44
$KS\text{-}AT_2$	K727	I781	14
$KS\text{-}AT_2$	K727	S815	28.86
$KS\text{-}AT_2$	K727	I822	19.02
$KS\text{-}AT_2$	K727	I823	21.53

Domain	Residue 1	Residue 2	Distance
KS-AT_2	K741	D699	24.82
KS-AT_2	K741	V700	21.01
KS-AT_2	K741	V703	13.16
KS-AT_2	K741	L712	16.07
KS-AT_2	K741	V716	17.03
KS-AT_2	K741	E717	14.44
KS-AT_2	K741	L721	17.8
KS-AT_2	K741	L722	20.17
KS-AT_2	K758	V702	11.53
KS-AT_2	K758	V703	10.75
KS-AT_2	K758	E704	8.08
KS-AT_2	K1071	Q940	32.78
KS-AT_2	K1071	W941	35.62
KS-AT_2	K1071	L1082	12.49
KS-AT_2	K1071	V1084	11.19
KS-AT_2	K1071	Y1123	23.73
KS-AT_2	K1071	A1124	20.3
KS-AT_2	K1071	H1126	18.25
KS-AT_2	K1071	A1128	17.56
KS-AT_2	K1071	H1129	20.57
KS-AT_2	K1273	D1252	35.64
KR2	K1407	W1387	14.98
KR2	K1407	T1388	11.92
KR2	K1407	V1389	12.5
KR2	K1407	T1390	10.51
KR2	K1407	V1394	10.69
KR2	K1407	S1398	17.99
KR2	K1407	L1426	11.4
KR2	K1407	V1427	13.25
KR2	K1407	E1474	32.63
KR2	K1407	Y1475	31.81

Domain	Residue 1	Residue 2	Distance
KR2	K1407	N1577	16.65
KR2	K1407	E1581	10.36
KR2	K1407	A1582	13.06
KR2	K1407	Q1623	12.34
KR2	K1407	G1624	9.89
KR2	K1407	Y1626	9.67
KR2	K1407	A1627	7.01
ACP2	K1784	V1813	8.2
ACP2	K1784	P1817	8.45
ACP2	K1784	A1823	13.75
ACP2	K1784	G1824	13.17
ACP2	K1784	F1825	16.55
ACP2	K1784	L1826	18.15
ACP2	K1784	A1828	18.62

Table S6. Number of conformations ranked at highest weight that match the cross-links measured within construct *A*. All experimental data was obtained with construct *A* of which experiments (Exp. 1 – Exp. 3) were performed with apo- and experiment Exp. 4 with phosphopantetheinylated holo-protein. Conformation show the total number (tot) of conformations that capture the cross-links and their BioEn cumulative weight (c. w.). The number of conformations was calculated based on the observed number of cross-linked peptide spectrum matches (CSMs) supporting a specific cross-link measured for construct *A* only and captured by A_{ext}, A_{arch}, and B_{ext}. Hits were sorted into "high frequency" hits (appearing more than twice in multiple experiments) or "medium frequency" hits (appearing once and/or twice in multiple experiments). Domain nomenclature: KS- ketosynthase, AT- acyltransferases, ACP- acyl carrier protein, KR – ketoreductase, LD – linker domain, DD- docking domain, and MBP- maltose binding protein.

Cross-linked Residues					Experimental #				Conformation #					
XL #	**Residue 1**	**Domain 1**	**Residue 2**	**Domain 2**	**Exp. 1**	**Exp. 2**	**Exp. 3**	**Exp. 4^{holo}**	**A_{ext} (tot)**	**A_{ext} (c.w.)**	**A_{arch} (tot)**	**A_{arch} (c.w.)**	**B_{ext} (tot)**	**B_{ext} (c.w.)**
High frequency hits														
XL1	K727	KS	E279	MBP	3	4	2	8	17	4.59E-03	7	3.52E-07		
XL2	K727	KS	L281	MBP	1	6			9	3.87E-03	3	2.37E-07		
XL3	K727	KS	L285	MBP		4	3	2	1	8.32E-05	0			
XL4	K727	KS	L286	MBP		11	4	1	3	2.25E-03	2	6.40E-06		
XL5	K727	KS	T1470	KR2		3	1		1	3.11E-05	8	2.73E-05	0	
XL6	K1071	AT	F48	MBP		5	3	1	1	2.45E-05	0			
XL7	K1071	AT	Q1623	KR2		5	4	2	0		0		0	
XL8	K1273	post-AT linker	E1474	KR2	1	6			393	1.55E-01			383	4.21E-01
XL9	K1273	post-AT linker	Y1475	KR2	1	2	2		432	2.15E-01			424	5.05E-01
XL10	K1273	post-AT linker	A1476	KR2	1	2	1	2	327	2.02E-01			350	4.05E-01
XL11	K1273	post-AT linker	I1482	KR2	3	2			85	1.36E-02			127	2.19E-01
XL12	K1273	post-AT linker	G1503	KR2	5	4			379	3.02E-01			425	3.91E-01
XL13	K1273	post-AT linker	A1683	KR2		1	1	7	628	3.18E-01			548	3.50E-01
XL14	K1775	ACP2	V1295	KR2			7	1	64	8.35E-03	231	6.60E-03	76	4.18E-02
XL15	K1775	ACP2	S1303	KR2	3	4	2	3	103	3.40E-02	169	6.50E-03	78	6.72E-02
XL16	K1775	ACP2	E1304	KR2			3	1	68	2.94E-02	87	2.16E-03	50	5.63E-02
XL17	K1775	ACP2	E1861	DD			4	2	165	5.14E-01	407	1.61E-01		

Cross-linked Residues					Experimental #				Conformation #					
XL #	**Residue 1**	**Domain 1**	**Residue 2**	**Domain 2**	**Exp. 1**	**Exp. 2**	**Exp. 3**	**Exp. 4^{holo}**	**A_{ext} (tot)**	**A_{ext} (c.w.)**	**A_{arch} (tot)**	**A_{arch} (c.w.)**	**B_{ext} (tot)**	**B_{ext} (c.w.)**
XL18	K1784	ACP2	G539	KS			2	2	32	4.55E-03	0		25	2.62E-02
XL19	K1784	ACP2	I554	KS	1	2	7		0		0		0	
XL20	K1784	ACP2	Y556	KS	6	5			6	2.67E-02	0		4	8.01E-04
XL21	K1784	ACP2	L560	KS	7	1	1	1	1	1.75E-04	0		0	
XL22	K1784	ACP2	A894	LD	6		1		1	3.97E-05	0		3	7.44E-04
XL23	K1784	ACP2	S1303	KR2	3	3			90	6.48E-02	70	1.39E-03	77	4.94E-02
XL24	K1784	ACP2	H1468	KR2	4	3	1		74	3.95E-02	48	5.48E-02	83	7.00E-02
XL25	K1784	ACP2	V1469	KR2	1	1	6		73	2.08E-02	183	5.11E-01	92	8.03E-02
XL26	K1784	ACP2	S1473	KR2	1	1	2		108	3.09E-02	825	8.15E-01	119	8.96E-02
XL27	K1784	ACP2	E1474	KR2		2	2		86	2.91E-02	958	6.11E-01	100	7.73E-02
XL28	K1784	ACP2	L1845	DD	2	4			323	2.33E-01	123	5.51E-02		
XL29	K47	MBP	L1408	KR2			3	3	0		0			
XL30	K143	MBP	L1408	KR2			2		0		0			
XL31	K145	MBP	L1408	KR2	1		1	5	0		0			
XL32	K180	MBP	Q667	KS		3			13	9.81E-03	13	3.22E-07		
XL33	K180	MBP	D668	KS		3			22	1.04E-02	14	3.56E-07		
XL34	K180	MBP	N672	KS		5	1		19	1.15E-02	6	2.10E-07		
XL35	K180	MBP	N1088	AT		5	2		1	6.30E-06	0			
XL36	K180	MBP	E1104	AT		2	2		0		0			
Medium frequency hits														
XL37	K1273	post-AT linker	L1464	KR2	1	1			117	1.12E-01			79	4.66E-02
XL38	K1273	post-AT linker	S1465	KR2	1	2			193	1.29E-01			149	7.25E-02
XL39	K1273	post-AT linker	A1471	KR2	1		1		525	6.53E-01			478	4.90E-01
XL40	K1273	post-AT linker	S1473	KR2		2			469	1.66E-01			451	5.05E-01
XL41	K1273	post-AT linker	V1484	KR2		3			44	4.56E-02			60	4.11E-02

Cross-linked Residues					Experimental #				Conformation #					
XL #	**Residue 1**	**Domain 1**	**Residue 2**	**Domain 2**	**Exp. 1**	**Exp. 2**	**Exp. 3**	**Exp. 4^{holo}**	**A_{ext} (tot)**	**A_{ext} (c.w.)**	**A_{arch} (tot)**	**A_{arch} (c.w.)**	**B_{ext} (tot)**	**B_{ext} (c.w.)**
XL42	K1775	ACP2	D538	KS			1		188	1.72E-01	0		211	1.87E-01
XL43	K1775	ACP2	G539	KS			2		136	1.53E-01	1	5.68E-09	140	1.29E-01
XL44	K1775	ACP2	Y540	KS			6		111	1.87E-01	3	2.63E-08	112	1.50E-01
XL45	K1775	ACP2	I554	KS			1		22	2.09E-03	0		7	6.85E-03
XL46	K1775	ACP2	Y556	KS	1		7		54	9.06E-03	3	6.13E-08	34	4.94E-02
XL47	K1784	ACP2	Y540	KS			3		29	4.93E-03	0		19	2.14E-02
XL48	K1784	ACP2	V598	KS		4			0		0		0	
XL49	K1784	ACP2	S599	KS		3			0		0		0	
XL50	K1784	ACP2	V600	KS		4			0		0		3	1.32E-03
XL51	K1784	ACP2	A1476	KR2		1	1		60	1.58E-02	1069	8.29E-01	84	6.91E-02
XL52	K1784	ACP2	P1853	DD	2		1		248	2.13E-01	67	1.13E-02		
XL53	K1784	ACP2	E1854	DD	3	1			364	2.05E-01	159	4.81E-01		
XL54	K84	MBP	T496	KS		9			0		1	2.04E-07		
XL55	K84	MBP	T1701	KR2		6			0		0			
XL56	K84	MBP	V1703	KR2		3			0		0			
XL57	K203	MBP	D668	KS		3			15	1.52E-03	10	3.64E-07		
XL58	K203	MBP	N672	KS		4			18	2.32E-03	8	1.06E-07		
XL59	K240	MBP	S522	KS				2	0		0			
XL60	K240	MBP	G525	KS				2	0		0			
XL61	K240	MBP	P531	KS				2	0		0			
XL62	K240	MBP	R532	KS				2	0		0			

* XL19. 48. 49. and 50 were satisfied 42. 2. 12. and 178 times when mapping the XL-MS data over all 45000 sampled CG conformations.

Table S7. Number of conformations ranked at highest weight that match the cross-links measured within construct *B*. All experimental data was obtained with construct *B* as phosphopantetheinylated holo-protein Conformation show the total number (tot) of conformations that capture the cross-links and their BioEn cumulative weight (c. w.). The number of conformations was calculated based on the observed number of cross-linked peptide spectrum matches (CSMs) supporting a specific cross-link measured for construct *B* and captured by B_{ext}. Although. the general reproducibility of CSMs of construct *B* was low, some XLs were reproduced between experimental data on construct *A* and *B* (Table S6). Domain nomenclature: KS- ketosynthase, AT- acyltransferase, ACP- acyl carrier protein, KR- ketoreductase, LD- linker domain, and TE- thioesterase.

Cross-linked residues					**Experimental #**			**Conformation #**	
XL #	**Residue 1**	**Domain 1**	**Residue 2**	**Domain**	**Exp. 1^{holo}**	**Exp. 2^{holo}**	**XL# from construct *A***	**B_{ext} (tot)**	**B_{ext} (c.w.)**
B-XL1	K217	KR1	V1274	AT		1	Domain absent in *A*	0	
B-XL2	K217	KR1	N1275	AT		2		1	5.16E-06
B-XL3	K217	KR1	W1484	KR2	1			0	
B-XL4	K217	KR1	T1485	KR2	1			0	
B-XL5	K810	KS	G259	KR1	1			4	2.91E-05
B-XL6	K810	KS	G261	KR1	1			7	2.25E-04
B-XL7	K810	KS	L263	KR1	3			4	2.91E-05
B-XL8	K914	KS	V391	KR1		1		5	1.93E-04
B-XL9	K914	KS	W393	KR1		4		20	3.51E-04
B-XL10	K914	KS	T395	KR1	2			38	7.37E-04
B-XL11	K914	KS	W396	KR1	4			44	7.33E-04
B-XL12	K1258	AT	G394	KR1	1			32	2.04E-03
B-XL13	K1258	AT	T395	KR1	1			32	2.00E-03
B-XL14	K1258	AT	W396	KR1	1			31	2.46E-03
B-XL15	K1460	post-AT linker	S1660	KR2		1	-	2315	5.35E-01
B-XL16	K1460	post-AT linker	E1661	KR2	2		XL8	1941	4.47E-01
B-XL17	K1460	post-AT linker	Y1662	KR2	2		XL9	2149	5.33E-01
B-XL18	K1460	post-AT linker	A1663	KR2	6		XL10	1671	4.27E-01
B-XL19	K1460	post-AT linker	V1664	KR2	5		-	1609	3.90E-01
B-XL20	K1460	post-AT linker	S1790	KR2	1		-	501	2.26E-01
B-XL21	K1460	post-AT linker	E1793	KR2	1		-	233	4.47E-02
B-XL22	K1594	KR2	V1120	AT	1		-	0	
B-XL23	K1594	KR2	F1121	AT	2		-	0	
B-XL24	K1594	KR2	E1143	AT		1	-	0	
B-XL25	K1594	KR2	I1145	AT		1	-	0	
B-XL26	K1962	ACP2	G186	KR1		1		2	2.53E-05

Cross-linked residues					Experimental #			Conformation #	
XL #	**Residue 1**	**Domain 1**	**Residue 2**	**Domain**	**Exp. 1^{holo}**	**Exp. 2^{holo}**	**XL# from construct *A***	**B_{ext} (tot)**	**B_{ext} (c.w.)**
B-XL27	K1962	ACP2	S1490	KR2	1	2	XL15	398	7.29E-02
B-XL28	K1971	ACP2	E2119	TE		1		170	1.90E-02
B-XL29	K1971	ACP2	P2122	TE	1			879	1.20E-01
B-XL30	K1971	ACP2	L2123	TE	1			1356	2.80E-01

Table S8. ELISA titration data. Data and data mean from two individual plates coated with (3)KS3–AT3. Data means are plotted in Fig. S20.

	Phage amount	ACP2-2	ACP1-2	Lib1-Mut3	Lib1-Mut5	Lib2-Mut1	Lib2-Mut3	Lib2-Mut4	Lib3-Mut1	Lib3-Mut3	Lib3-Mut4	Lib3-Mut5	Lib3-Mut6	Lib3-Mut7
Plate1	2.3E+12	1.81	0.79	3.01	3.14	1.77	2.86	1.30	1.73	3.77	3.38	7.38	5.47	4.35
	7.65E+11	1.42	0.72	2.12	2.32	1.29	2.01	0.97	1.28	2.94	2.66	4.28	3.80	3.54
	7.65E+10	0.26	0.49	1.11	1.18	0.78	1.04	0.58	0.78	1.54	1.44	1.95	1.90	1.88
	7.65E+9	0.04	0.30	0.66	0.68	0.47	0.72	0.39	0.49	0.85	0.94	1.12	1.20	1.07
Plate 2	2.3E+12	2.29	1.08	3.48	3.8	2.23	4.27	1.64	2.12	5.19	4.22	6.79	5	5.11
	7.65E+11	1.73	0.96	2.5	2.7	1.6	2.51	1.21	1.62	4.03	3.25	4.01	3.99	3.62
	7.65E+10	0.27	0.53	1.15	1.13	0.83	1.12	0.61	0.88	1.67	1.47	1.76	1.92	1.82
	7.65E+9	0.04	0.3	0.65	0.71	0.48	0.78	0.4	0.51	0.87	0.89	1.14	1.23	1.09
Mean	2.3E+12	2.05	0.93	3.24	3.47	2	3.56	1.47	1.93	4.48	3.8	7.08	5.24	4.73
	7.65E+11	1.58	0.84	2.31	2.51	1.45	2.26	1.09	1.45	3.49	2.95	4.15	3.9	3.58
	7.65E+10	0.27	0.51	1.13	1.16	0.81	1.08	0.59	0.83	1.6	1.46	1.86	1.91	1.85
	7.65E+9	0.04	0.3	0.66	0.69	0.47	0.75	0.4	0.5	0.86	0.91	1.13	1.22	1.08

Table S9. Amino acid sequences and alignment of proteins used in this study and the length of the modeled domains. Length of the modeled domains is indicated by the color coding: MBP (light gray), KR1 (yellow), ACP1 (pink on gray background), KS (blue), LD (dark gray), AT (green), KR2 (orange), ACP2 (magenta), and DD or TE (black).

Construct	Amino acid sequence
A. MBP-M2(2) (A_{ext})	MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQ DKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYEN GKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVG VLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAV INAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGPGAAHYGSAAPATTAPVDEPIAIVGMACRLPGEVDSPERLWELITSGRDSAA EVPDDRGWVPDELMASDAAGTRRAHGNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSDTGVFVGMS HQGYATGRPRPEDGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTACSSSLVALHTACGSLRDGDCGLAVAGGVSVMAGPEVFTE FSRQGALSPDGRCKPFSDEADGFGLGEGSAFVVLQRLSDARREGRRVLGVVAGSAVNQDGASNGLSAPSGVAQQRVIRRAWARAGIT GADVAVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQAAAGVAGVIKVLLGLERGVVPPMLCRGERSGLIDW SSGEIELADGVREWSPAADGVRRAGVSAFGVSGTNAHVIIAEPPEPEPVPQPRRMLPATGVVPVVLSARTGAALRAQAGRLADHLAA HPGIAPADVSWTMARARQHFEERAAVLAADTAEAVHRLRAVADGAVVPGVVTGSASDGGSVFVFPGQGAQWEGMARELLPVPVFAES IAECDAVLSEVAGFSVSEVLEPRPDAPSLERVDVVQPVLFAVMVSLARLWRACGAVPSAVIGHSQGEIAAAVVAGALSLEDGMRVVA RRSRAVRAVAGRGSMLSVRGGRSDVEKLLADDSWTGRLEVAAVNGPDAVVVAGDAQAAREFLEYCEGVGIRARAIPVDYASHTAHVE PVRDELVQALAGITPRRAEVPFFSTLTGDFLDGTELDAGYWYRNLRHPVEFHSAVQALTDQGYATFIEVSPHPVLASSVQETLDDAE SDAAVLGTLERDAGDADRFLTALADAHTRGVAVDWEAVLGRAGLVDLPGYPFQGKRFWLLPDRTTPRDELDGWFYRVDWTEVPRSEP AALRGRWLVVVPEGHEEDGWTVEVRSALAEAGAEPEVTRGVGGLVGDCAGVVSLLALEGDGAVQTLVLVRELDAEGIDAPLWTVTFG AVDAGSPVARPDQAKLWGLGQVASLERGPRWTGLVDLPHMPDPELRGRLTAVLAGSEDQVAVRADAVRARRLSPAHVTATSEYAVPG GTILVTGGTAGLGAEVARWLAGRGAEHLALVSRRGPDTEGVGDLTAELTRLGARVSVHACDVSSREPVRELVHGLIEQGDVVRGVVH AAGLPQQVAINDMDEAAFDEVVAAKAGGAVHLDELCSDAELFLLFSSGAGVWGSARQGAYAAGNAFLDAFARHRRGRGLPATSVAWG LWAAGGMTGDEEAVSFLRERGVRAMPVPRALAALDRVLASGETAVVVTDVDWPAFAESYTAARPRPLLDRIVTTAPSERAGEPETES LRDRLAGLPRAERTAELVRLVRTSTATVLGHDDPKAVRATTPFKELGFDSLAAVRLRNLLNAATGLRLPSTLVFDHPNASAVAGFLD AELGTEVRGEAPSALAGLDALEAALPEVPATEREELVQRLERMLAALRPVAQAADASGTGANPSGDDLGEAGVDELLEALGRELDGD
A. MBP-M2(2) (A_{arch})	MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQ DKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYEN GKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVG VLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAV INAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGPGAAHYGSAAPATTAPVDEPIAIVGMACRLPGEVDSPERLWELITSGRDSAA EVPDDRGWVPDELMASDAAGTRRAHGNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSDTGVFVGMS HQGYATGRPRPEDGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTACSSSLVALHTACGSLRDGDCGLAVAGGVSVMAGPEVFTE FSRQGALSPDGRCKPFSDEADGFGLGEGSAFVVLQRLSDARREGRRVLGVVAGSAVNQDGASNGLSAPSGVAQQRVIRRAWARAGIT GADVAVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQAAAGVAGVIKVLLGLERGVVPPMLCRGERSGLIDW SSGEIELADGVREWSPAADGVRRAGVSAFGVSGTNAHVIIAEPPEPEPVPQPRRMLPATGVVPVVLSARTGAALRAQAGRLADHLAA HPGIAPADVSWTMARARQHFEERAAVLAADTAEAVHRLRAVADGAVVPGVVTGSASDGGSVFVFPGQGAQWEGMARELLPVPVFAES IAECDAVLSEVAGFSVSEVLEPRPDAPSLERVDVVQPVLFAVMVSLARLWRACGAVPSAVIGHSQGEIAAAVVAGALSLEDGMRVVA RRSRAVRAVAGRGSMLSVRGGRSDVEKLLADDSWTGRLEVAAVNGPDAVVVAGDAQAAREFLEYCEGVGIRARAIPVDYASHTAHVE PVRDELVQALAGITPRRAEVPFFSTLTGDFLDGTELDAGYWYRNLRHPVEFHSAVQALTDQGYATFIEVSPHPVLASSVQETLDDAE SDAAVLGTLERDAGDADRFLTALADAHTRGVAVDWEAVLGRAGLVDLPGYPFQGKRFWLLPDRTTPRDELDGWFYRVDWTEVPRSEP AALRGRWLVVVPEGHEEDGWTVEVRSALAEAGAEPEVTRGVGGLVGDCAGVVSLLALEGDGAVQTLVLVRELDAEGIDAPLWTVTFG AVDAGSPVARPDQAKLWGLGQVASLERGPRWTGLVDLPHMPDPELRGRLTAVLAGSEDQVAVRADAVRARRLSPAHVTATSEYAVPG GTILVTGGTAGLGAEVARWLAGRGAEHLALVSRRGPDTEGVGDLTAELTRLGARVSVHACDVSSREPVRELVHGLIEQGDVVRGVVH AAGLPQQVAINDMDEAAFDEVVAAKAGGAVHLDELCSDAELFLLFSSGAGVWGSARQGAYAAGNAFLDAFARHRRGRGLPATSVAWG LWAAGGMTGDEEAVSFLRERGVRAMPVPRALAALDRVLASGETAVVVTDVDWPAFAESYTAARPRPLLDRIVTTAPSERAGEPETES LRDRLAGLPRAERTAELVRLVRTSTATVLGHDDPKAVRATTPFKELGFDSLAAVRLRNLLNAATGLRLPSTLVFDHPNASAVAGFLD AELGTEVRGEAPSALAGLDALEAALPEVPATEREELVQRLERMLAALRPVAQAADASGTGANPSGDDLGEAGVDELLEALGRELDGD

B. KR1-ACP1-M2-TE (B$_{ext}$)	MDEVSALRYRIEWRPTGAGEPARLDGTWLVAKYAGTADETSTAAREALESAGARVRELVVDARCGRDELAERLRSVGEVAGVLSLLAVD EAEPEEAPLALASLADTLSLVQAMVSAELGCPLWTVTESAVATGPFERVRNAAHGALWGVGRVIALENPAVWGGLVDVPAGSVAELARH LAAVVSGGAGEDQLALRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGGTGGVGGQIARWLARRGAPHLLLVSRSGPDADGAGELVAEL EALGARTTVAACDVTDRESVRELLGGIGDDVPLSAVFHAAATLDDGTVDTLTGERIERASRAKVLGARNLHELTRELDLTAFVLFSSFA SAFGAPGLGGYAPGNAYLDGLAQQRRSDGLPATAVAWGTWAGSGMAEGPVADRFRRHGVIEMPPETACRALQNALDRAEVCPIVIDVRW DRFLLAYTAQRPTRLFDEIDDARRAAPQAAAEPRVGAHMGASLPAPEREKALFELVRSHAAAVLGHASAERVPADQAFAELGVDSLSAL ELRNRLGAATGVRLPTTTVFDHPDVRTLAAHLAAELGGATGAEQAAPATTAPVDEPIAIVGMACRLPGEVDSPERLWELITSGRDSAAE VPDDRGWVPDELMASDAAGTRRAHGNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSDTGVFVGMSHQG YATGRPRPEDGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTACSSSLVALHTACGSLRDGDCGLAVAGGVSVMAGPEVFTEFSRQG ALSPDGRCKPFSDEADGFGLGEGSAFVVLQRLSDARREGRRVLGVVAGSAVNQDGASNGLSAPSGVAQQRVIRRAWARAGITGADVAVV EAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQAAAGVAGVIKVLLGLERGVVPPMLCRGERSGLIDWSSGEIELAD GVREWSPAADGVRRAGVSAFGVSGTNAHVIIAEPPEPEPVPQPRRMLPATGVVPVVLSARTGAALRAQAGRLADHLAAHPGIAPADVSW TMARARQHFEERAAVLAADTAEAVHRLRAVADGAVVPGVVTGSASDGGSVFVFPGQGAQWEGMARELLPVPVFAESIAECDAVLSEVAG FSVSEVLEPRPDAPSLERVDVVQPVLFAVMVSLARLWRACGAVPSAVIGHSQGEIAAAVVAGALSLEDGMRVVARRSRAVRAVAGRGSM LSVRGGRSDVEKLLADDSWTGRLEVAAVNGPDAVVVAGDAQAAREFLEYCEGVGIRARAIPVDYASHTAHVEPVRDELVQALAGITPRR AEVPFFSTLTGDFLDGTELDAGYWYRNLRHPVEFHSAVQALTDQGYATFIEVSPHPVLASSVQETLDDAESDAAVLGTLERDAGDADRF LTALADAHTRGVAVDWEAVLGRAGLVDLPGYPFQGKRFWLLPDRTTPRDELDGWFYRVDWTEVPRSEPAALRGRWLVVVPEGHEEDGWT VEVRSALAEAGAEPEVTRGVGGLVGDCAGVVSLLALEGDGAVQTLVLVRELDAEGIDAPLWTVTFGAVDAGSPVARPDQAKLWGLGQVA SLERGPRWTGLVDLPHMPDPELRGRLTAVLAGSEDQVAVRADAVRARRLSPAHVTATSEYAVPGGTILVTGGTAGLGAEVARWLAGRGA EHLALVSRRGPDTEGVGDLTAELTRLGARVSVHACDVSSREPVRELVHGLIEQGDVVRGVVHAAGLPQQVAINDMDEAAFDEVVAAKAG GAVHLDELCSDAELFLLFSSGAGVWGSARQGAYAAGNAFLDAFARHRRGRGLPATSVAWGLWAAGGMTGDEEAVSFLRERGVRAMPVPR ALAALDRVLASGETAVVVTDVDWPAFAESYTAARPRPLLDRIVTTAPSERAGEPETESLRDRLAGLPRAERTAELVRLVRTSTATVLGH DDPKAVRATTPFKELGFDSLAAVRLRNLLNAATGLRLPSTLVFDHPNASAVAGFLDAELSGTPAREASSALRDGYRQAGVSGRVRSYLD LLAGLSDFREHFDGSDGFSLDLVDMADGPGEVTVICCAGTAAISGPHEFTRLAGALRGIAPVRAVPQPGYEEGEPLPSSMAAVAAVQAD AVIRTQGDKPFVVAGHSAGALMAYALATELLDRGHPPRGVVLIDVYPPGHQDAMNAWLEELTATLFDRETVRMDDTRLTALGAYDRLTG QWRPRETGLPTLLVSAGEPMGPWPDDSWKPTWPFEHDTVAVPGDHFTMVQEHADAIARHIDAWLGGGNSSSVDKLAAALEHHHHHH

Table S10. Plasmids and primer used in this study.

Plasmid	Cloning Method	Fragments	Primer Name	Primer Sequence 5'-3'	Template
pMK53. MBP-harmM2(2)	In-Fusion	Insert	P-MK141	GCAGCTCCGGCAACCACA	pMK39
			P-MK146	GGATCCGGCGCCTTTTTCGAA	
		Vector	P-MK175	GGTTGCCGGAGCTGCGGATCCGTAGTGTGCCGCA	pMK05
			P-MK143	GAACTGGATGGTGACTAGGATCCGGCTGCTAACAAAG	
pMK79. ACP1-harmM2+TE-H6 (used as a template for pADD01)	In-Fusion	Vector1	P-MK238	GGTGCGACCGGAGCCGAGCAGGCAGCTCCGGCAACCACA	pMK62
			P-MK240	GGGCGGGAGTCCCGCT	
		Vector2	P-MK153	AGCGGGACTCCCGCCC	pMK62
			P-MK239	CATATGTATATCTCCTTCTTAAAGTTAAACAAAATTA	
		Insert	P-MK236	AAGAAGGAGATATACATATGCTGGCGTCGCTGCCCG	pMK59
			P-MK237	CTCGGCTCCGGTCGCACCGCCGAGTTCGGCGGCCAGGT	
pADD01. KR1-ACP1-harmM2-TE	In-Fusion and Restriction Digest	KR-Insert	P-MK276	AAGGAGATATACATATGGACGAGGTTTCCGCGCTG	pBL13
			P-MK277	AGCGACGCCAGCATATGCGCGCCCACCCGCGGTTC	
		Vector	pMK79 linearized with NdeI		
pMK45. (3)K3-AT3-Strep	In-Fusion	Vector	P-MK163	GCGGTAGGCCAGCTCGT	pRSG34
			P-MK159	TGAGATCCGGTAACAAAGCCC	
		Insert	P-MK161	TGTTACCGGATCTCAGGATCCGGCGCCTTTTTCGAA	pAR10 (any Strep containing vector)
			P-MK164	GAGCTGGCCTACCGCAGCGCTTGGAGCCATCCAC	

Plasmids pMK39 (TwinStrep-**harm_DEBSM2**(2)), pMK05 (**MBP**_DEBSM2**(2)-H8**), pMK62 ((3)**harmDEBSM2+TE-H6**), pMK59 (TwinStrp-**DEBSACP1**(2)), and pBL13[2] ((5)KS1-AT1-**KR1**-ACP1(2)) were used as templates to yield the fragments indicated in bold. For phage display analysis, the following plasmids were used: pBL12 (LLD(4)), pBL13 ((5)M1(2)), pBL16 (all plasmids, see ref. [49]) ((3)-M2-TE), RSG34 (ref. [50]) ((3)-M3-TE). Phagemid pMK22 (ACP1-(2)-phagemid) was assembled via In-Fusion cloning from pBL13 (ACP1 and docking domain (2) and the phagemid pNE.

Supporting Movie:

SI Movie 1: A_{ext} relaxed conformations show high flexibility of construct *A* in solution. Trajectory illustrating the 33 configurations of model A_{ext} ranked to explain 90% of the BioEn fitting. Domain coloring: KS- blue, LD- dark gray, AT- green, KR- orange, ACP-magenta, DD- light gray.

Supporting References:

1. Tonikian, R.; Zhang, Y.; Boone, C.; Sidhu, S. S., (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nat Protoc 2*, 1368-1386.
2. Rittner, A.; Paithankar, K. S.; Drexler, D. J.; Himmler, A.; Grininger, M., (2019) Probing the modularity of megasynthases by rational engineering of a fatty acid synthase Type I. *Protein Sci 28*, 414-428.
3. Hughes, A. J.; Keatinge-Clay, A., (2011) Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of Streptomyces coelicolor MatB. *Chem Biol 18*, 165-176.
4. Sharma, K. K.; Boddy, C. N., (2007) The thioesterase domain from the pimaricin and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates. *Bioorg Med Chem Lett 17*, 3034-3037.
5. Peter, D. M.; Schada von Borzyskowski, L.; Kiefer, P.; Christen, P.; Vorholt, J. A.; Erb, T. J., (2015) Screening and Engineering the Synthetic Potential of Carboxylating Reductases from Central Metabolism and Polyketide Biosynthesis. *Angewandte Chemie International Edition 54*, 13457-13461.
6. Li, X.; Sevillano, N.; La Greca, F.; Deis, L.; Liu, Y. C.; Deller, M. C.; Mathews, II; Matsui, T.; Cane, D. E.; Craik, C. S.; Khosla, C., (2018) Structure-Function Analysis of the Extended Conformation of a Polyketide Synthase Module. *J Am Chem Soc 140*, 6518-6521.
7. Peter, D. M.; Schada von Borzyskowski, L.; Kiefer, P.; Christen, P.; Vorholt, J. A.; Erb, T. J., (2015) Screening and Engineering the Synthetic Potential of Carboxylating Reductases from Central Metabolism and Polyketide Biosynthesis. *Angew Chem Int Ed Engl 54*, 13457-13461.
8. Pernot, P.; Round, A.; Barrett, R.; De Maria Antolinos, A.; Gobbo, A.; Gordon, E.; Huet, J.; Kieffer, J.; Lentini, M.; Mattenet, M.; Morawe, C.; Mueller-Dieckmann, C.; Ohlsson, S.; Schmid, W.; Surr, J.; Theveneau, P.; Zerrad, L.; McSweeney, S., (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat 20*, 660-664.
9. Yang, B.; Wu, Y. J.; Zhu, M.; Fan, S. B.; Lin, J.; Zhang, K.; Li, S.; Chi, H.; Li, Y. X.; Chen, H. F.; Luo, S. K.; Ding, Y. H.; Wang, L. H.; Hao, Z.; Xiu, L. Y.; Chen, S.; Ye, K.; He, S. M.; Dong, M. Q., (2012) Identification of cross-linked peptides from complex samples. *Nat Methods 9*, 904-906.
10. Sali, A.; Blundell, T. L., (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol 234*, 779-815.
11. Tang, Y.; Chen, A. Y.; Kim, C. Y.; Cane, D. E.; Khosla, C., (2007) Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem Biol 14*, 931-943.
12. Tang, Y.; Kim, C.-Y.; Mathews, I. I.; Cane, D. E.; Khosla, C., (2006) The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc Natl Acad Sci U S A. 103*, 11124-11129.
13. Keatinge-Clay, A. T.; Stroud, R. M., (2006) The structure of a ketoreductase determines the organization of the β-carbon processing enzymes of modular polyketide synthases. *Structure 14*, 737-748.
14. Alekseyev, V. Y.; Liu, C. W.; Cane, D. E.; Puglisi, J. D.; Khosla, C., (2007) Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Sci 16*, 2093-2107.
15. Broadhurst, R. W.; Nietlispach, D.; Wheatcroft, M. P.; Leadlay, P. F.; Weissman, K. J., (2003) The structure of docking domains in modular polyketide synthases. *Chem Biol 10*, 723-731.
16. Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E., (1997) Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure 5*, 997-1015.

17. Tsai, S. C.; Miercke, L. J.; Krucinski, J.; Gokhale, R.; Chen, J. C.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M., (2001) Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. *Proc Natl Acad Sci U S A 98*, 14808-14813.
18. DeLano, W. L., (2002) Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography 40*, 82-92.
19. Edwards, A. L.; Matsui, T.; Weiss, T. M.; Khosla, C., (2014) Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J Mol Biol 426*, 2229-2245.
20. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem 25*, 1605-1612.
21. Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G., (2014) Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature 510*, 560-564.
22. Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G., (2014) Structure of a modular polyketide synthase. *Nature 510*, 512-517.
23. Kim, Y. C.; Hummer, G., (2008) Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. *J Mol Biol 375*, 1416-1433.
24. Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A., (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J 105*, 962-974.
25. Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A., (2016) FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res 44*, W424-429.
26. Różycki, B.; Kim, Y. C.; Hummer, G., (2011) SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure 19*, 109-116.
27. Köfinger, J.; Stelzl, L. S.; Reuter, K.; Allande, C.; Reichel, K.; Hummer, G., (2019) Efficient Ensemble Refinement by Reweighting. *J Chem Theory Comput 15*, 3390-3401.
28. Hummer, G.; Kofinger, J., (2015) Bayesian ensemble refinement by replica simulations and reweighting. *J Chem Phys 143*, 243150.
29. Svergun, D., (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography 25*, 495-503.
30. Oliphant, T. E., *A guide to NumPy. vol. 1* Tregol Publishing USA: 2006.
31. Jones, E.; Oliphant, T. E.; Peterson, P., *SciPy: Open source scientific tools for Python.* 2001.
32. McKinney, W.; and others. . in Proceedings of the 9th Python in Science Conference 51--56 (Austin, T., 2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference 51--56* Austin, TX, 2010.
33. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Dotson, D. L.; Domanski, J.; Buchoux, S.; Kenney, I. M.; Beckstein, O., (2016) MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations. *Proceedings of the 15th Python in Science Conference*, 102-109.
34. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O., (2011) MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem 32*, 2319-2327.
35. Hunter, J. D., (2007) A 2D Graphics Environment. *Computing in Science & Engineering 9.*, 90-95.
36. Humphrey, W.; Dalke, A.; Schulten, K., (1996) VMD: visual molecular dynamics. *J Mol Graph 14*, 33-38, 27-38.
37. Grishaev, A., (2012) Sample preparation, data collection, and preliminary data analysis in biomolecular solution X-ray scattering. *Curr Protoc Protein Sci Chapter 17*, Unit17.14.
38. Herbst, D. A.; Jakob, R. P.; Zahringer, F.; Maier, T., (2016) Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature 531*, 533-537.
39. Rambo, R. P.; Tainer, J. A., (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature 496*, 477-481.
40. Piiadov, V.; Ares de Araújo, E.; Oliveira Neto, M.; Craievich, A. F.; Polikarpov, I., (2019) SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. *Protein Science 28*, 454-463.
41. Hansen, S., (2000) Bayesian estimation of hyperparameters for indirect Fourier transformation in small-angle scattering. *J Appl Crystallogr 33*, 1415-1421.

42. Hopkins, J. B.; Gillilan, R. E.; Skou, S., (2017) BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J Appl Crystallogr 50*, 1545-1553.
43. Klaus, M.; D'Souza, A. D.; Nivina, A.; Khosla, C.; Grininger, M., (2019) Engineering of Chimeric Polyketide Synthases Using SYNZIP Docking Domains. *ACS Chem Biol 14*, 426-433.
44. Klaus, M.; Buyachuihan, L.; Grininger, M., (2020) Ketosynthase Domain Constrains the Design of Polyketide Synthases. *ACS Chem Biol 15*, 2422-2432.
45. Needleman, S. B.; Wunsch, C. D., (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol 48*, 443-453.
46. Zhang, G.; Ignatova, Z., (2009) Generic algorithm to predict the speed of translational elongation: implications for protein biogenesis. *PLoS One 4*, e5036.
47. Rittner, A.; Paithankar, K. S.; Huu, K. V.; Grininger, M., (2018) Characterization of the Polyspecific Transferase of Murine Type I Fatty Acid Synthase (FAS) and Implications for Polyketide Synthase (PKS) Engineering. *ACS Chem Biol 13*, 723-732.
48. Krissinel, E.; Henrick, K., (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol 372*, 774-797.
49. Lowry, B.; Robbins, T.; Weng, C. H.; O'Brien, R. V.; Cane, D. E.; Khosla, C., (2013) In vitro reconstitution and analysis of the 6-deoxyerythronolide B synthase. *J Am Chem Soc 135*, 16809-16812.
50. Gokhale, R.; Tsuji, S.; Cane, D.; Khosla, C., (1999) Dissecting and exploiting intermodular communication in polyketide synthases. *Science 284*, 482-485.