Supplementary Materials for

Mapping the catalytic conformations of an assembly-line polyketide synthase module

Dillon P. Cogan *et al.*

Corresponding authors: Wah Chiu, wahc@stanford.edu; Chaitan Khosla, khosla@stanford.edu

Science **374**, 729 (2021)
DOI: 10.1126/science.abi8358

The PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S24
Tables S1 to S3
References

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S3
MDAR Reproducibility Checklist

Materials and Methods

Materials

Chemicals, resins, plasmid isolation kits, and bacterial growth media were from MilliPore Sigma or Thermo Fisher Scientific. Isopropyl β-D-1-thiogalactopyranoside (IPTG) and antibiotics were from Gold Biotechnology. Enzymes for PCR and Gibson Assembly were from Takara Bio USA, Inc., and oligonucleotides and DNA sequencing data were from Elim Biopharmaceutical Inc.

General Methods

The following proteins were expressed and purified for use in bimodule and trimodule enzyme assays, as described previously(*10*, *14*): LDD(4), (5)M1(2), (3)M2(2), and (3)M2TE from the 6-deoxyerythronolide B synthase (DEBS; parenthetical numbers refer to native docking domains and the modules from which they derive); auxiliary proteins MatB, PrpE, and methylmalonyl-CoA epimerase (SCME). Protein concentrations were measured using the Bradford assay with bovine serum albumin standards (Thermo Fisher Scientific). Absorbance was measured at 595 nm in 1 cm cuvettes using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific)(*41*).

Protein Sequences

Key:

KS = ketosynthase
AT = acyltransferase
KR = ketoreductase
ACP = acyl carrier protein
TE = thioesterase
DEBS M1
DEBS M2 Docking Domain
DEBS M3
DEBS M3 Docking Domain
DEBS TE

DEBS M3/1 with a C-terminal TE domain (pTED23) **| (3)**-KS3-AT3-KR1-ACP1-TE

MASTDSEKVAEYLRRATLDLRAARQRIRELESDPIAIVSMACRLPGGVNTPQRLWELLREGGET
LSGFPTDRGWDLARLHHPDPDNPGTSYVDKGGFLDDAAGFDAEFFGVSPREAAAMDPQQRLLLE
TSWELVENAGIDPHSLRGTATGVFLGVAKFGYGEDTAAAEDVEGYSVTGVAPAVASGRISYTMG
LEGPSISVDTACSSSLVALHLAVESLRKGESSMAVVGGAAVMATPGVFVDFSRQRALAADGRSK
AFGAGADGFGFSEGVTLVLLERLSEARRNGHEVLAVVRGSALNQDGASNGLSAPSGPAQRRVIR
QALESCGLEPGDVDAVEAHGTGTALGDPIEANALLDTYGRDRDADRPLWLGSVKSNIGHTQAAA
GVTGLLKVVLALRNGELPATLHVEEPTPHVDWSSGGVALLAGNQPWRRGERTRRARVSAFGISG
TNAHVIVEEAPEREHRETTAHDGRPVPLVVSARTTAALRAQAAQIAELLERPDADLAGVGLGLA
TTRARHEHRAAVVASTREEAVRGLREIAAGAATADAVVEGVTEVDGRNVVFLFPGQGSQWAGMG
AELLSSSPVFAGKIRACDESMAPMQDWKVSDVLRQAPGAPGLDRVDVVQPVLFAVMVSLAELWR
SYGVEPAAVVGHSQGEIAAAHVAGALTLEDAAKLVVGRSRLMRSLSGEGGMAAVALGEAAVRER
LRPWQDRLSVAAVNGPRSVVVSGEPGALRAFSEDCAAEGIRVRDIDVDYASHSPQIERVREELL
ETTGDIAPRPARVTFHSTVESRSMDGTELDARYWYRNLRETVRFADAVTRLAESGYDAFIEVSP
HPVVVQAVEEAVEEADGAEDAVVVGSLHRDGGDLSAFLRSMATAHVSGVDIRWDVALPGAAPFA
LPTYPFQRKRYWLQPAAPAAASDELAYRIEWRPTGAGEPARLDGTWLVAKYAGTADETSTAARE
ALESAGARVRELVVDARCGRDELAERLRSVGEVAGVLSLLAVDEAEPEEAPLALASLADTLSLV
QAMVSAELGCPLWTVTESAVATGPFERVRNAAHGALWGVGRVIALENPAVWGGLVDVPAGSVAE

LARHLAAVVSGGAGEDQLALRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGGTGGVGGQIARW
LARRGAPHLLLVSRSGPDADGAGELVAELEALGARTTVAACDVTDRESVRELLGGIGDDVPLSA
VFHAAATLDDGTVDTLTGERIERASRAKVLGARNLHELTRELDLTAFVLFSSFASAFGAPGLGG
YAPGNAYLDGLAQQRRSDGLPATAVAWGTWAGSGMAEGPVADRFRRHGVIEMPPETACRALQNA
LDRAEVCPIVIDVRWDRFLLAYTAQRPTRLFDEIDDARRAAPQAAAEPRVGALASLPAPEREKA
LFELVRSHAAAVLGHASAERVPADQAFAELGVDSLSALELRNRLGAATGVRLPTTTVFDHPDVR
TLAAHLAAELGSGTPAREASSALRDGYRQAGVSGRVRSYLDLLAGLSDFREHFDGSDGFSLDLV
DMADGPGEVTVICCAGTAAISGPHEFTRLAGALRGIAPVRAVPQPGYEEGEPLPSSMAAVAAVQ
ADAVIRTQGDKPFVVAGHSAGALMAYALATELLDRGHPPRGVVLIDVYPPGHQDAMNAWLEELT
ATLFDRETVRMDDTRLTALGAYDRLTGQWRPRETGLPTLLVSAGEPMGPWPDDSWKPTWPFEHD
TVAVPGDHFTMVQEHADAIARHIDAWLGGGNSSSVDKLAAALEHHHHHH

DEBS M1 with a C-terminal TE domain; also "M1TE" (pDC1) | **(3)**-KS1-AT1-KR1-ACP1-TE
MASTDSEKVAEYLRRATLDLRAARQRIRELEGEPVAVVAMACRLPGGVSTPEEFWELLSEGRDA
VAGLPTDRGWDLDSLFHPDPTRSGTAHQRGGGFLTEATAFDPAFFGMSPREALAVDPQQRLMLE
LSWEVLERAGIPPTSLQASPTGVFVGLIPQEYGPRLAEGGEGVEGYLMTGTTTSVASGRIAYTL
GLEGPAISVDTACSSSLVAVHLACQSLRRGESSLAMAGGVTVMPTPGMLVDFSRMNSLAPDGRC
KAFSAGANGFGMAEGAGMLLLERLSDARRNGHPVLAVLRGTAVNSDGASNGLSAPNGRAQVRVI
QQALAESGLGPADIDAVEAHGTGTRLGDPIEARALFEAYGRDREQPLHLGSVKSNLGHTQAAAG
VAGVIKMVLAMRAGTLPRTLHASERSKEIDWSSGAISLLDEPEPWPAGARPRRAGVSSFGISGT
NAHAIIEEAPQVVEGERVEAGDVVAPWVLSASSAEGLRAQAARLAAHLREHPGQDPRDIAYSLA
TGRAALPHRAAFAPVDESAALRVLDGLATGNADGAAVGTSRAQQRAVFVFPGQGWQWAGMAVDL
LDTSPVFAAALRECADALEPHLDFEVIPFLRAEAARREQDAALSTERVDVVQPVMFAVMVSLAS
MWRAHGVEPAAVIGHSQGEIAAACVAGALSLDDAARVVALRSRVIATMPGNKGMASIAAPAGEV
RARIGDRVEIAAVNGPRSVVVAGDSDELDRLVASCTTECIRAKRLAVDYASHSSHVETIRDALH
AELGEDFHPLPGFVPFFSTVTGRWTQPDELDAGYWYRNLRRTVRFADAVRALAEQGYRTFLEVS
AHPILTAAIEEIGDGSGADLSAIHSLRRGDGSLADFGEALSRAFAAGVAVDWESVHLGTGARRV
PLPTYPFQRERVWLEPKPVARRSTEVDEVSALRYRIEWRPTGAGEPARLDGTWLVAKYAGTADE
TSTAAREALESAGARVRELVVDARCGRDELAERLRSVGEVAGVLSLLAVDEAEPEEAPLALASL
ADTLSLVQAMVSAELGCPLWTVTESAVATGPFERVRNAAHGALWGVGRVIALENPAVWGGLVDV
PAGSVAELARHLAAVVSGGAGEDQLALRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGGTGGV
GGQIARWLARRGAPHLLLVSRSGPDADGAGELVAELEALGARTTVAACDVTDRESVRELLGGIG
DDVPLSAVFHAAATLDDGTVDTLTGERIERASRAKVLGARNLHELTRELDLTAFVLFSSFASAF
GAPGLGGYAPGNAYLDGLAQQRRSDGLPATAVAWGTWAGSGMAEGPVADRFRRHGVIEMPPETA
CRALQNALDRAEVCPIVIDVRWDRFLLAYTAQRPTRLFDEIDDARRAAPQAAAEPRVGALASLP
APEREKALFELVRSHAAAVLGHASAERVPADQAFAELGVDSLSALELRNRLGAATGVRLPTTTV
FDHPDVRTLAAHLTSELGSGTPAREASSALRDGYRQAGVSGRVRSYLDLLAGLSDFREHFDGSD
GFSLDLVDMADGPGEVTVICCAGTAAISGPHEFTRLAGALRGIAPVRAVPQPGYEEGEPLPSSM
AAVAAVQADAVIRTQGDKPFVVAGHSAGALMAYALATELLDRGHPPRGVVLIDVYPPGHQDAMN
AWLEELTATLFDRETVRMDDTRLTALGAYDRLTGQWRPRETGLPTLLVSAGEPMGPWPDDSWKP
TWPFEHDTVAVPGDHFTMVQEHADAIARHIDAWLGGGNSSSVDKLAAALEHHHHHH

DEBS M1 with a C-terminal M2 docking domain (pDC7) | **(3)**-KS1-AT1-KR1-ACP1-**(2)**
MASTDSEKVAEYLRRATLDLRAARQRIRELEGEPVAVVAMACRLPGGVSTPEEFWELLSEGRDA
VAGLPTDRGWDLDSLFHPDPTRSGTAHQRGGGFLTEATAFDPAFFGMSPREALAVDPQQRLMLE
LSWEVLERAGIPPTSLQASPTGVFVGLIPQEYGPRLAEGGEGVEGYLMTGTTTSVASGRIAYTL
GLEGPAISVDTACSSSLVAVHLACQSLRRGESSLAMAGGVTVMPTPGMLVDFSRMNSLAPDGRC

KAFSAGANGFGMAEGAGMLLLERLSDARRNGHPVLAVLRGTAVNSDGASNGLSAPNGRAQVRVI
QQALAESGLGPADIDAVEAHGTGTRLGDPIEARALFEAYGRDREQPLHLGSVKSNLGHTQAAAG
VAGVIKMVLAMRAGTLPRTLHASERSKEIDWSSGAISLLDEPEPWPAGARPRRAGVSSFGISGT
NAHAIIEEAPQVVEGERVEAGDVVAPWVLSASSAEGLRAQAARLAAHLREHPGQDPRDIAYSLA
TGRAALPHRAAFAPVDESAALRVLDGLATGNADGAAVGTSRAQQRAVFVFPGQGWQWAGMAVDL
LDTSPVFAAALRECADALEPHLDFEVIPFLRAEAARREQDAALSTERVDVVQPVMFAVMVSLAS
MWRAHGVEPAAVIGHSQGEIAAACVAGALSLDDAARVVALRSRVIATMPGNKGMASIAAPAGEV
RARIGDRVEIAAVNGPRSVVVAGDSDELDRLVASCTTECIRAKRLAVDYASHSSHVETIRDALH
AELGEDFHPLPGFVPFFSTVTGRWTQPDELDAGYWYRNLRRTVRFADAVRALAEQGYRTFLEVS
AHPILTAAIEEIGDGSGADLSAIHSLRRGDGSLADFGEALSRAFAAGVAVDWESVHLGTGARRV
PLPTYPFQRERVWLEPKPVARRSTEVDEVSALRYRIEWRPTGAGEPARLDGTWLVAKYAGTADE
TSTAAREALESAGARVRELVVDARCGRDELAERLRSVGEVAGVLSLLAVDEAEPEEAPLALASL
ADTLSLVQAMVSAELGCPLWTVTESAVATGPFERVRNAAHGALWGVGRVIALENPAVWGGLVDV
PAGSVAELARHLAAVVSGGAGEDQLALRADGVYGRRWVRAAAPATDDEWKPTGTVLVTGGTGGV
GGQIARWLARRGAPHLLLVSRSGPDADGAGELVAELEALGARTTVAACDVTDRESVRELLGGIG
DDVPLSAVFHAAATLDDGTVDTLTGERIERASRAKVLGARNLHELTRELDLTAFVLFSSFASAF
GAPGLGGYAPGNAYLDGLAQQRRSDGLPATAVAWGTWAGSGMAEGPVADRFRRHGVIEMPPETA
CRALQNALDRAEVCPIVIDVRWDRFLLAYTAQRPTRLFDEIDDARRAAPQAAAEPRVGALASLP
APEREKALFELVRSHAAAVLGHASAERVPADQAFAELGVDSLSALELRNRLGAATGVRLPTTTV
FDHPDVRTLAAHLAAELG**TEVRGEAPSALAGLDALEAALPEVPATEREELVQRLERMLAALRPV**
AQAADASGTGANPSGDDLGEAGVDELLEALGRELDGDPNSSSVDKLAAALEHHHHHH

F_{ab} 1B2 (heavy chain)
MAEVQLVQSGGGLVQPGRSLRLSCTASGFTFGDYAMSWVRQAPGKGLEWVGFIRSKAYGGTTEY
AASVKGRFTISRDDSKSIAYLQMNSLKTEDTAVYYCTRGGTLFDYWGQGTLVTVSSASTKGPSV
FPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVP
SSSLGTQTYICNVNHKPSNTKVDKKVEPKSCAALVPRGSAHHHHHHAADYKDDDDKA

F_{ab} 1B2 (light chain)
LFAIPLVVPFYSHSALDVVMTQSPLSLPVTPGEPASISCRSSQSLLHSNGYNYLDWYLQKPGQS
PQLLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQSLQTPRLTFGPGTKVDI
KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPRGAKVQWKVDNALQSGNSQESVTEQDSKD
STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Plasmids

Table S3 contains a comprehensive list of the plasmids used in this study. New constructs, pDC1 and pDC7, encode for C-terminally His$_6$-tagged DEBS Module 1 (M1) containing a C-terminal TE domain or the C-terminal docking domain from DEBS M2, respectively. The cryo-EM structures of DEBS M1 reported in Figures 2 and 3 were solved using protein expressed from pDC1, whereas the structure of M1 reported in Figure 4 was solved using protein expressed from pDC7. Both plasmids were constructed via PCR amplification of the following fragment pairs:

> pDC1: (i) pBL13-derived(*14*) DEBS M1 region spanning Gly557 – Leu2010 (5′ – 3′; AAV51820.1) and (ii) pRSG34-derived(*12*) fragment containing the DEBS TE domain starting at Ser2895 (5′; AAV39552.1) and spanning to the DEBS M3 docking domain up to Glu29 (3′; AAV39551.1).

pDC7: (i) pDC1-derived DEBS (3)M1 region spanning Met1 (5′; AAV39551.1) – Gly2015 (3′; AAV51820.1) and (ii) pBL36-derived(*14*) fragment containing the DEBS M2 docking domain starting at Thr2463 (5′; AAV39551.1) and spanning to the DEBS M3 docking domain up to Met1 (3′; AAV39551.1).

PCR amplicons were ligated via Gibson Assembly(*42*, *43*), and the resultant circularized products were used to transform *E. coli* DH5α competent cells. Plasmids were verified by sequencing. For a list of oligonucleotide PCR primers, see table S2.

Protein expression and purification

DEBS modules and the KS-AT didomain of Module 3: Expression plasmids pRSG34, pTED23, pDC1, or pDC7 were used to transform *E. coli* BAP1 competent cells(*12*, *21*) whereas expression plasmid pAYC02 was used to transform *E. coli* BL21(DE3) competent cells(*13*). Transformants from a single colony were used to inoculate 20 mL Luria-Bertani (LB) media supplemented with 100 mg/L carbenicillin and grown overnight at 37 °C (220 rpm). Each 1 L flask of antibiotic-containing LB medium was inoculated with 2 mL of overnight seed culture and grown at 37 °C (220 rpm) until an optical density at 600 nm (OD_{600}) of ~0.4–0.8 was achieved. Cells were then cooled for 20 min in an ice bath before adding 0.2 mM IPTG and continuing growth at 18 °C (220 rpm) for 18–20 h. Cell harvesting, lysis, and Ni-NTA affinity purification was performed as described previously(*21*). The resulting 50 mL of Ni-NTA purified protein eluent was injected onto a 5 mL HiTrap Q HP anion exchange column (GE Healthcare) equilibrated with buffer A (50 mM citric acid, 5 mM HEPES, 10% glycerol, pH 7.6 [NaOH]) at a flow rate of 2 mL/min using an ÄKTA Pure protein purification FPLC system (GE Healthcare). The column was washed with 50 mL of buffer A before employing a linear ramp of 0–60% increasing buffer B (1 M NaCl, 50 mM citric acid, 5 mM HEPES, 10% glycerol, pH 7.6 [NaOH]) over 90 mL while collecting 3 mL fractions at a flow rate of 3 mL/min. Eluted fractions were evaluated for purity by SDS-PAGE before pooling and concentrating the protein using Amicon Ultra Centrifugal Filters (50 kDa MWCO for intact modules and 30 kDa MWCO for the KS-AT didomain; Millipore Sigma) and flash-freezing in liquid N_2 for storage at -80 °C.

F_{ab} 1B2: Expression, periplasmic extraction, and purification of F_{ab} 1B2 by Ni-NTA affinity chromatography was performed as described previously(*21*). Protein eluant was concentrated to >5 mg/mL using Amicon Ultra Centrifugal Filters (10 kDa MWCO; Millipore Sigma), supplemented with 5% glycerol, and flash-frozen in liquid N_2 for storage at -80 °C.

Selecting a buffer for structural analysis of intact DEBS modules

Following a report describing the positive correlation between ideal two-state unfolding and sample quality for macromolecular structure determination(*11*), we set out to identify optimal buffer conditions by protein melt-curve analysis of intact DEBS modules. Previous analyses of DEBS and other PKS modules have highlighted an activating influence of certain polyvalent anions; most notably phosphate and citrate(*14*, *44*). Reasoning that catalytic activation might be rooted in a thermostabilizing effect, we tested the anion concentration dependence of the melting temperature (T_m) of M1 + TE (M1TE; pDC1) and M3 + TE (M3TE; pRSG34) using differential scanning fluorimetry (DSF)(*45*). Addition of both phosphate and citrate up to 0.9 M conferred a marked increase in the T_m of M3TE by ~16 °C, and comparative analysis revealed citrate as the more potent thermostabilizer (figs. S3-S4). Interestingly, M1TE exhibited two unfolding transitions (i.e., non-ideal two-state unfolding) in all cases (fig. S2). Whereas the T_m of one of the two M1TE transitions (~59 °C) does not appear to depend on phosphate or citrate concentration, the other transition (at a lower T_m) exhibits a dose-dependent increase in T_m (~ 39–48 °C from 7–450 mM phosphate/citrate). The relative peak intensities of the lower T_m transition are diminished in the case of citrate compared to phosphate. To relate these DSF observations to catalytic activity, we reconstituted a DEBS bimodule and measured its reaction velocities under a range of phosphate and citrate concentrations. In both cases, enzymatic activity was enhanced in a concentration dependent manner, with citrate again outperforming phosphate in potency (fig. S5). A similar result was obtained after analysis of a DEBS trimodule featuring M3TE (fig. S5). Perhaps signaling at a basis for the anion activating effect, we observed that citrate, and to a lesser degree phosphate, induces dimerization of the DEBS KS3-AT3 didomain(*13*) by size-exclusion chromatography (SEC) analysis (figs. S6).

Thermal melt-curve analysis of DEBS Module 1 + TE (M1TE) and Module 3 + TE (M3TE) by differential scanning fluorimetry (DSF)

DEBS M1TE and M3TE purified by anion-exchange chromatography were further purified by size-exclusion chromatography (SEC, below) before DSF analysis. To optimize for protein and SYPRO Orange (Thermo Fisher Scientific, Cat. # S6650) concentration, a preliminary grid of conditions was assessed in which M1TE and M3TE concentration varied from 0.1 – 0.5 mg/mL, and SYPRO Orange concentration varied from 6× – 12× (relative to a 5,000× concentrated stock, as purchased). Dilutions were prepared in SEC buffer (100 mM citric acid, 100 mM NaCl, 10 mM HEPES, pH 7.2 [NaOH]) with final well volumes of 15 μL in 96-well Hard-Shell PCR plates (Bio-Rad, Cat. # HSP9665). The protein and dye mixtures were allowed to equilibrate at room temperature for 15 min before a thermal melt protocol was initiated using a Bio-Rad CFX96 RT-PCR instrument. The "FRET" channel was selected to satisfy the excitation and emission wavelengths of SYPRO Orange (λ_{ex} 470 nm /λ_{em} 570 nm). The melt sequence involved first maintaining the sample at 20 °C for 5 min before linearly increasing the temperature at 2 °C/min with fluorescence measurements every 30 sec. Judging by the relative fluorescence units (RFU) versus temperature plots, the most optimal unfolding melt curves (displaying clearly resolved unfolding transitions(*11*)) were obtained with 0.1 mg/mL M1TE/M3TE and 9× SYPRO Orange. The same protein and dye concentrations were therefore used in subsequent DSF experiments in which phosphate and citrate concentrations were varied. The following three solutions were prepared:

1: 0.1 mg/mL M1TE or M3TE, 9× SYPRO Orange, 3 mM citric acid, 8.3 mM HEPES, pH 7.2 (NaOH)
2: 0.1 mg/mL M1TE or M3TE, 9× SYPRO Orange, 900 mM NaH_2PO_4, 8.3 mM HEPES, pH 7.2 (NaOH)
3: 0.1 mg/mL M1TE or M3TE, 9× SYPRO Orange, 900 mM citric acid, 8.3 mM HEPES, pH 7.2 (NaOH)

Solution **2** was used to serially dilute into solution **1**, and solution **3** was used to serially dilute into solution **1** to obtain a range of 900 – 7 mM phosphate and citrate conditions, respectively. Serial dilutions were caried out in triplicate and wells containing 15 μL of the serially diluted contents were used in thermal melt-curve analysis, as above. Data corresponding to the melt curves and first derivative plots were exported from the Bio-Rad CFX Manager 3.1 Data Analysis tool and subsequently plotted in OriginPro 2016 (figs. S2-S4).

Catalytic activity measurements of a DEBS bimodule and trimodule as a function of phosphate and citrate concentration

DEBS bimodule | LDD(4) + (5)M1(2) + (3)M2TE (i = phosphate titration; ii = citrate titration): To a 13.0 μL mixture of 77 mM *NaH_2PO_4 (pH 7.2), 7.7 mM tris(2-carboxyethyl)phosphine (TCEP), 15.4 mM $MgCl_2$, 2.3 mM CoASH, 9.2 mM ATP, 3.1 μM PrpE, 3.1 μM MatB, 6.2 μM SCME, 6.2 μM LDD(4), and 6.2 μM (5)M1(2) was added 6.5 μL of the following: (i) 1538 – 153.8 mM *NaH_2PO_4 (pH 7.2) + 9.2 μM (3)M2TE or (ii) 1538 – 153.8 mM *citric acid + 9.2 μM (3)M2TE. The reactions were initiated by addition of 0.5 μL of a mixture containing 35 mM sodium propionate, 35 mM methylmalonic acid, and 28.6 mM NADPH to arrive at a final reaction volume of 20 μL and the following final reaction concentrations:

> (i) 500 – 50 mM *NaH_2PO_4 (pH 7.2; 6 different concentrations), 5 mM TCEP, 10 mM $MgCl_2$, 1.5 mM CoASH, 6 mM ATP, 1 mM sodium propionate, 1 mM methylmalonic acid, 0.8 mM NADPH, 2 μM PrpE, 2 μM MatB, 4 μM SCME, 4 μM LDD(4), 4 μM (5)M1(2), and 3 μM (3)M2TE.
>
> (ii) [†]450 – 50 mM *citric acid (pH 7.2; 5 different concentrations), 50 mM *NaH_2PO_4 (pH 7.2), 5 mM TCEP, 10 mM $MgCl_2$, 1.5 mM CoASH, 6 mM ATP, 1 mM sodium propionate, 1 mM methylmalonic acid, 0.8 mM NADPH, 2 μM PrpE, 2 μM MatB, 4 μM SCME, 4 μM LDD(4), 4 μM (5)M1(2), and 3 μM (3)M2TE.

Similar single-turnover control reactions in which (3)M2TE was omitted were also prepared at a single NaH_2PO_4 or citric acid concentration of 250 mM. The 20 μL reaction contents were transferred to wells of a clear-bottom 384-well plate (Corning, product # 3765) to measure NADPH depletion by monitoring absorbance at 340 nm every 10 s using a BioTek Synergy HT plate reader at ambient temperature (20–22 °C; fig. S5).

DEBS trimodule | LDD(4) + (5)M1(2) + (3)M2(2) + M3TE: DEBS trimodule reactions were carried out identically to the bimodule reactions with the exception of substituting (3)M2TE for (3)M2(2) and M3TE. A final enzyme concentration of 6 μM was used for (5)M1(2) and (3)M2(2), 4 μM was used for M3TE, and all other reaction components were maintained at the same levels as for the bimodule above (fig. S5).

* NaH_2PO_4 and citric acid solutions were derived from 2 M and 1.67 M stock solutions, respectively, that were supplemented with 8.3 mM HEPES and pH-adjusted to 7.2 with NaOH.

[†] It was observed that 50 mM NaH_2PO_4 (pH 7.2) was required to prevent protein aggregation during preparation of the 13 μL mixture at low citrate concentrations (i.e., < 100 mM).

Phosphate and citrate effect on DEBS KS3-AT3 didomain oligomerization by SEC

The anion-exchange purified DEBS KS3-AT3 didomain was injected (200 μL; 5 nmol) onto a 24 mL Superdex 200 pg 10/300 column (GE Healthcare) at a flow rate of 1 mL/min using an ÄKTA Pure protein purification FPLC system (GE Healthcare). Independent, isocratic elutions were carried out with the following running buffers (pH-adjusted with NaOH), and in the same sequential order as the listed buffers (1-5):

1. 50 mM NaH_2PO_4, 50 mM NaCl, 10 mM HEPES, pH 7.2
2. 100 mM NaH_2PO_4, 10 mM HEPES, pH 7.2
3. 100 mM citric acid, 100 mM NaCl, 10 mM HEPES, pH 7.2 (SEC buffer)
4. 500 mM citric acid, 10 mM HEPES, pH 7.2
5. 300 mM NaCl, 10 mM tris, pH 7.2

A 30 mL column pre-equilibration step was included prior to each elution. A final, sixth injection was performed with buffer 1 and its corresponding chromatogram superposed well with that from the first run (fig. S6A). Gel-filtration protein standards (Bio-Rad, Cat. # 1511901) were injected onto the same column pre-equilibrated and eluted with SEC buffer to correlate the experimental peak elution volumes with their oligomeric states (fig. S6B).

Purification of DEBS modules, F_{ab} 1B2, and the module-F_{ab} complexes by size-exclusion chromatography (SEC) for single-particle cryo-EM analysis

The cryo-EM structures of DEBS M1 reported in Figures 2 and 3 were solved using protein expressed from pDC1 (i.e., this protein harbored a C-terminal TE domain, even though the TE was not structurally resolved). In contrast, the structure of M1 reported in Figure 4 was solved using protein expressed from pDC7 (which lacked a TE domain). The cryo-EM structure of the chimeric module (M3/1) reported in Figure 2 was solved using protein expressed from pTED23 (which also harbored a C-terminal TE domain).

All DEBS modules and F_{ab} 1B2 used in cryo-EM experiments were individually purified via SEC prior to preparation and re-purification of the module-F_{ab} complexes. For SEC purification, protein samples ≤ 2 mL were injected onto a 120 mL Superdex 200 pg 16/600 column (GE Healthcare) at a flow rate of 1.5 mL/min and fractionated into 3 mL isocratically using an ÄKTA Pure protein purification FPLC system (GE Healthcare). For DEBS modules, fractions corresponding to the homodimeric species were pooled and concentrated using Amicon Ultra Centrifugal Filters (50 kDa MWCO; Millipore Sigma). For F_{ab} 1B2, the heterodimeric peak was pooled and concentrated using Amicon Ultra Centrifugal Filters (10 kDa MWCO; Millipore Sigma). While ensuring that an overall volume of 2 mL was not exceeded, to 11–16 nmol of concentrated module (calculated based on monomeric mass) was added 1.2-fold excess amounts of concentrated F_{ab} 1B2 heterodimer (i.e., 1:1.2 monomer to F_{ab} 1B2 heterodimer, in accordance with the binding stoichiometry). The protein mixture was allowed to equilibrate for 30 min on ice before purification by SEC using the same column, flow rate, and running buffer as above. The fractions corresponding to the module-F_{ab} complex were collected manually to minimize the levels of co-purified, higher-order oligomeric species. Pooled fractions were concentrated using Amicon Ultra Centrifugal Filters (50 kDa MWCO; Millipore Sigma). In cases where the peak profile was sufficiently devoid of higher molecular weight impurities, the samples were concentrated to ≥ 8 mg/mL and directly mounted onto cryo-EM grids for downstream single-particle analysis. Samples deemed too impure for cryo-EM analysis were re-injected onto the SEC column for a repeated purification, followed by concentration (as above), grid mounting, and single-particle cryo-EM analysis (fig. S7).

Measuring the catalytic activity of M1 in a reconstituted DEBS bimodule system

To 20.5 μL of an enzyme pre-mix of 880 mM *NaH_2PO_4 (pH 7.2), 7.3 mM TCEP, 14.7 mM $MgCl_2$, 2.2 mM CoASH, 8.8 mM ATP, 1.5 mM propionyl-CoA, 2.9 μM MatB, and 5.9 μM SCME was added 9 μL of either (i) 13.3 μM M1TE (pDC1) and 13.3 μM (3)M2TE, (ii) 13.3 μM M1 (pDC7) and 13.3 μM (3)M2TE, (iii) 13.3 μM (3)M2TE, or (iv) 13.3 μM M1 (pDC7). To initiate the reactions, 0.5 μL of a mixture of 55.5 mM methylmalonic acid and 44.4 mM NADPH was added to each of the enzyme mixtures (i-iv) to achieve final reaction concentrations of: 600 mM *NaH_2PO_4 (pH 7.2), 5 mM TCEP, 10 mM $MgCl_2$, 1.5 mM CoASH, 6 mM ATP, 1 mM propionyl-CoA, 1 mM methylmalonic acid, 0.8 mM NADPH, 2 μM MatB, 4 μM SCME, and 4 μM of each module. The 30 μL reaction contents were transferred to wells of a clear-bottom 384-well plate (Corning, product # 3765) to measure NADPH depletion by monitoring absorbance at 340 nm every 10 s using a BioTek Synergy HT plate reader at ambient temperature (20–22 °C; fig. S20).

* NaH_2PO_4 was derived from a 2 M stock solution that was supplemented with 8.3 mM HEPES and pH-adjusted to 7.2 with NaOH.

Preparation of the turnstile-closed M1 for single-particle cryo-EM analysis

The complex between M1 (pDC7) and F_{ab} 1B2 was purified in the same manner as the other module-F_{ab} complexes above. The M1-F_{ab} complex was concentrated to 11 mg/mL before preparing a 20 μL reaction containing 38.5 μM M1-F_{ab}, 4 mM 2(*R/S*)-methylmalonyl-CoA, 2 mM propionyl-CoA, 2 mM NADPH, and 1 mM TCEP. The reaction contents were incubated for 1 h at ambient temperature (20–22 °C) before promptly mounting onto cryo-EM grids by plunge freezing. Notably, particle quality could only be maintained when 1 mM TCEP was added with substrates, although TCEP was not a requirement in unacylated (*holo* state) module preparations.

Cryo-EM data acquisition

All DEBS-F_{ab} complexes isolated by SEC (above) were diluted to 10 mg/mL with SEC buffer. Addition of 0.03% nonyl phenoxypolyethoxylethanol (NP-40) was required to prevent sample degradation before vitrification. Three microliters of the sample were then applied onto glow-discharged 200-mesh R2/1 Quantifoil copper grids. The grids were blotted for 4 s and rapidly cryocooled in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific) at 4 °C and 100% humidity. The samples were imaged in a Titan Krios cryo-electron microscope (Thermo Fisher Scientific) at a magnification of 75,000× (corresponding to a calibrated sampling of 1.0 Å per pixel). Micrographs were recorded by EPU software (Thermo Fisher Scientific) with a Falcon 4 detector, where each image was composed of 40 individual frames in gain-normalized .mrc format with an exposure time of 8.5 s and an exposure rate of 8.28 electrons per second per $Å^2$. A total of 3,974 movie stacks for M3/1-F_{ab}, 10,234 movies stacks for M1-F_{ab} (pDC1), 8,174 movie stacks for M1-F_{ab} (pDC7), and 5,977 movies stacks for diketide-M1-F_{ab} (pDC7) were collected.

Single-particle image processing and 3D reconstruction

All micrographs were first imported into Relion(*46*) for image processing. The motion correction was performed using MotionCor2(*47*) and the contrast transfer function (CTF) was determined using CTFFIND4(*48*). All particles were auto-picked using the NeuralNet option in EMAN2(*49*), then particle coordinates were imported into Relion where the poor 2D class averages were removed by 2D classification. The selected particles after 2D classification were transferred to cryoSPARC(*50*) for ab-initio map generation. Classes with apparent KS-AT and F_{ab} 1B2 regions were subjected to non-uniform refinement to obtain 3D maps. Heterogeneous refinement was performed to further classify the particles. For the M1-F_{ab} dataset (pDC1), three different conformations were obtained (fig. S10C). Another round of non-uniform refinement was then conducted to achieve a 3.2 Å final map for *State 1*, a 4.1 Å final map for *State 2*, and a 4.1 Å final map for the *State 1′* state (figs. S10C, S11). For the M3/1-F_{ab} (pTED23), M1-F_{ab} (pDC7), and diketide-M1-F_{ab} (pDC7) datasets, heterogeneous refinement failed to further classify the particles; therefore, the 3D maps resulting from the first rounds of non-uniform refinement were used as the final maps with resolutions of 3.2 Å, 3.4 Å, and 4.3 Å, respectively (figs. S8, S13, and S21). Resolution for the final maps was estimated with the 0.143 criterion of the Fourier shell correlation curve. Resolution maps were calculated in cryoSPARC using the "Local Resolution Estimation" option (Fig. 4B and figs. S8, S11, S13). The figures were prepared using UCSF Chimera(*51*). See table S1 for a summary of data processing statistics.

Model building

To model the M1-F_{ab} (pDC1) sample in different conformations, the crystal structures of F_{ab} 1B2 (PDB ID: 6C9U) and DEBS KR1 (PDB ID: 2FR0) were selected. KS-AT and ACP fragments of Module 1 were modeled using SWISS-MODEL(*52*). The models were rigidly fitted into the final maps, followed by manual adjustments in Coot(*53*) and automated refinement with phenix.real_space_refine(*54*). In some cases, the AT and KR domains could not be well fit into the cryo-EM density by rigid-body approaches; thus, molecular dynamics flexible fitting (MDFF) was used(*31*). In particular, the flexed AT from subunit A (cyan; Fig. 2C and fig. S17) in *State 2* and the KR from subunit B (magenta; fig. S19) in *State 1′* were fit using MDFF. The MDFF was completed in two runs, with each run including 10^4 minimization steps and 10^5 molecular dynamics steps. Given that the diketide-M1-F_{ab} (pDC7) "*turnstile-closed*" map also possessed flexed ATs but at reduced local resolution compared to the *State 2* flexed AT, the resultant flexed AT coordinates from MDFF using the *State 2* map were fit as rigid bodies into the AT densities of the *turnstile-closed* map and subsequently refined. All of the resulting models were refined using both Coot(*53*) and phenix.real_space_refine(*54*), and the final models were evaluated by MolProbity(*55*). Statistics for the map reconstruction and model building are summarized in table S1. All figures were prepared using UCSF Chimera(*51*).

Supplementary Text

Comparison of the M1-F_{ab} (pDC1) and M1-F_{ab} (pDC7) cryo-EM maps and resolution of the KR subunit attachment

The two M1 modules are of identical composition with the exception of harboring either a C-terminal TE (pDC1) or DEBS M2 docking domain (pDC7). Single-particle cryo-EM analysis of their complexes with F_{ab} 1B2 resulted in comparable 3D classes, suggesting that none of the ambiguous densities corresponded to either the TE or M2 docking domains. Due to the limited resolution in the region spanning the AT and KR domains in the *State 1* map, we looked to the corresponding region in the analogous class average derived from M1-F_{ab} (pDC7, closely resembling the *State 1* map; fig. S13). In this case, a continuous region of density separating the AT and KR from subunit A in the M1-F_{ab} map could be clearly resolved (figs. S13 and S14A); thus, we inferred that the same module subunit connectivity exists in *State 1*.

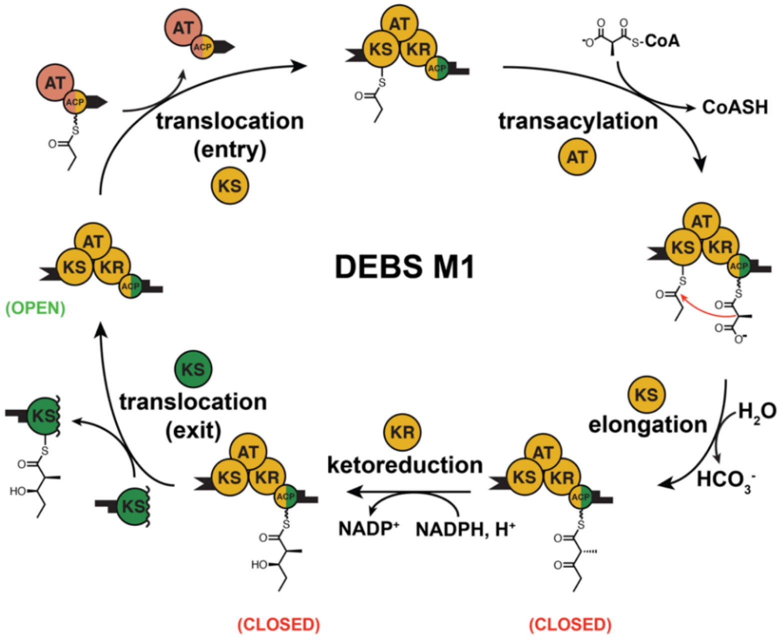
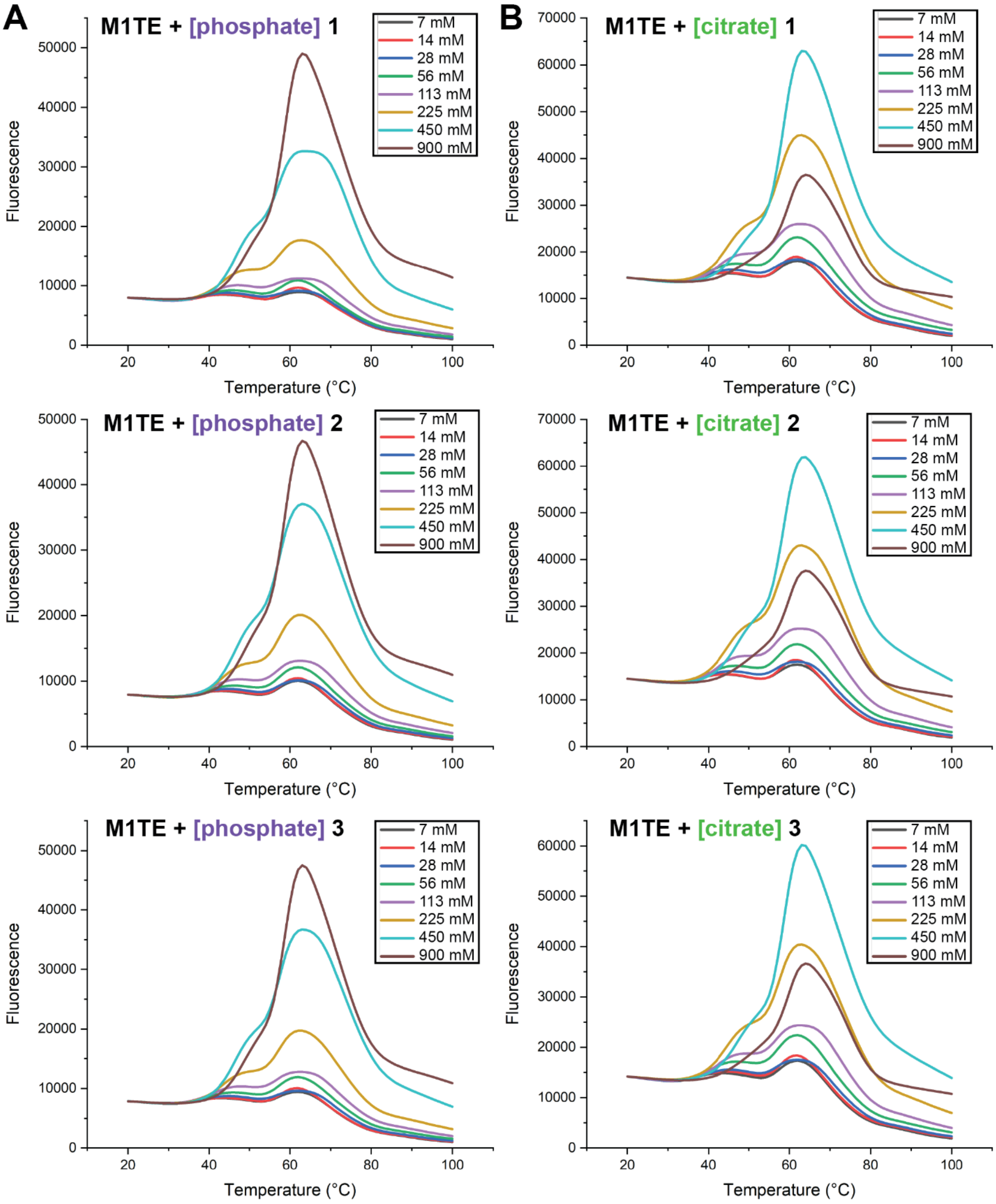


Fig. S1.

The catalytic cycle of Module 1 (M1, mostly amber) of DEBS is initiated by the translocation of a propionyl group from the acyl carrier protein (ACP) domain of the upstream module (LM, mostly salmon) to the active site Cys residue of the M1 ketosynthase (KS) (hereafter referred to as the "entry translocation" step). Meanwhile, an acyltransferase (AT) domain transfers a (2*S*)-methylmalonyl group from methylmalonyl-CoA to its ACP domain ("transacylation"). The bis-acylated M1 module catalyzes decarboxylative condensation between the ACP-bound extender unit and KS-bound polyketide to form an ACP-bound β-ketoacyl thioester product ("elongation"). The exergonicity of this step is harnessed to induce a structural change that prevents polyketide back-transfer (as with iterative PKSs)(*56*) or another entry translocation step, thereby ensuring that only a single intermediate is bound to each modular subunit at a given time. Thus, assembly-line PKS modules appear to harbor molecular turnstiles to process intermediates in a controlled, vectorial manner(*10*). Following elongation, additional domains such as a ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) may modify the β-ketoacyl thioester with stereocontrol at the α- and β-carbons. The product is transferred ("exit translocation") to the next module (M2, mostly green) with concomitant turnstile reopening. The mechanism by which the post-elongation KS is shielded is unknown, but presumably relies on conformational or chemical changes to limit access to its freshly vacated active site Cys(*3*). (Such a mechanism would not be expected for iterative PKSs or vertebrate fatty acid synthases, since they evolved to favor back-transfer of intermediates to the KS for repeated catalytic cycles.) The ACPs are bicolor to emphasize that each ACP harbors at least two distinct recognition elements that impart KS specificity during intermodular translocation and intramodular elongation(*29*, *30*). Shape-complementary black tabs represent native DEBS docking domains that enhance the specificity of intermodular polyketide translocation(*40*). Native M1 lacks any such docking domains, but engineered variants harboring docking domains have been constructed to reconstitute its activity as a free-standing module (Ref. (*14*) and fig. S20). The closed and open states of the turnstile are labeled in parentheses(*10*).


A
B
M1TE + [phosphate] 1
M1TE + [phosphate] 2
M1TE + [phosphate] 3
M1TE + [citrate] 1
M1TE + [citrate] 2
M1TE + [citrate] 3
7 mM
14 mM
28 mM
56 mM
113 mM
225 mM
450 mM
900 mM
Fluorescence
Temperature (°C)
0
10000
20000
30000
40000
50000
60000
70000
20
40
60
80
100

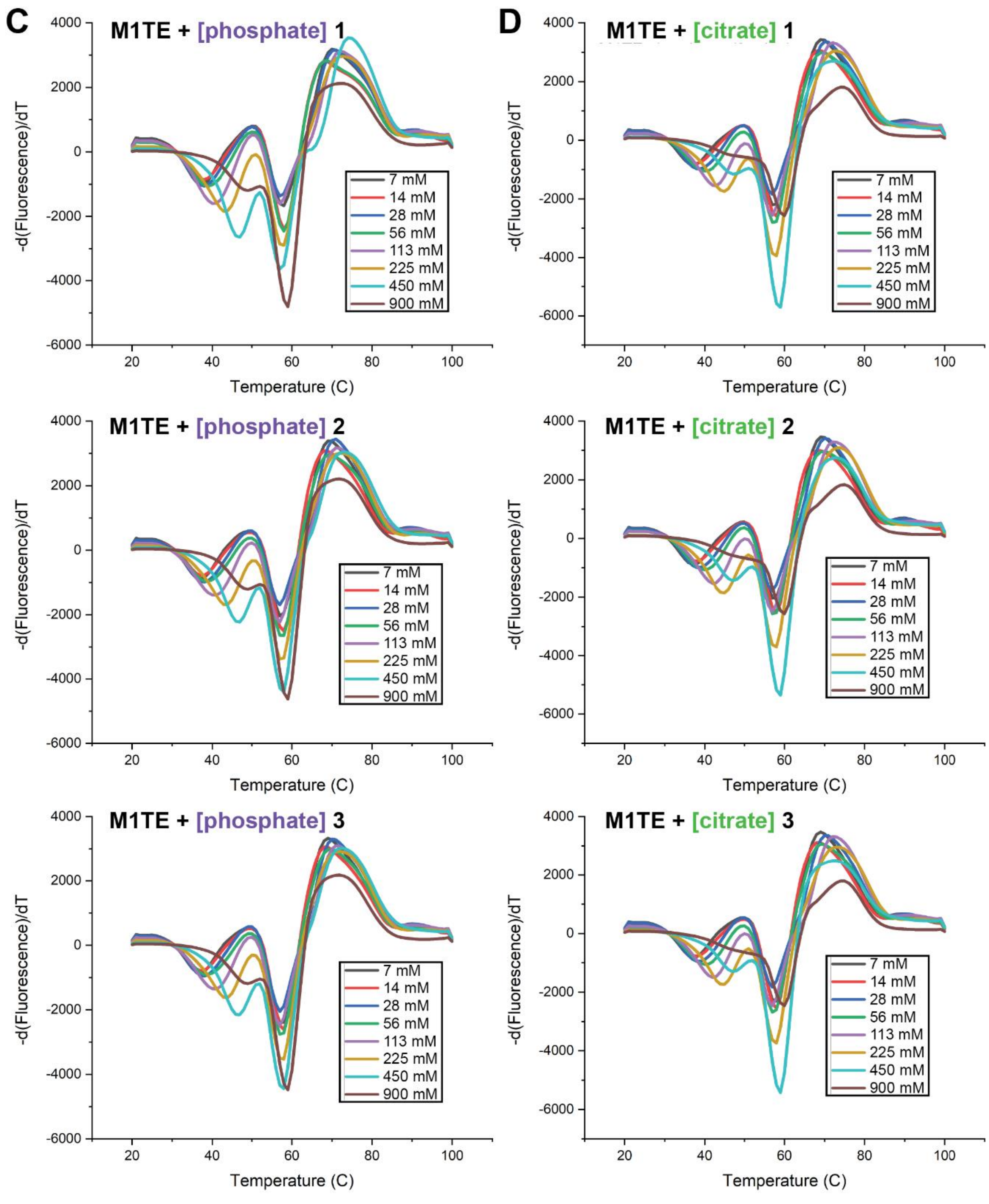
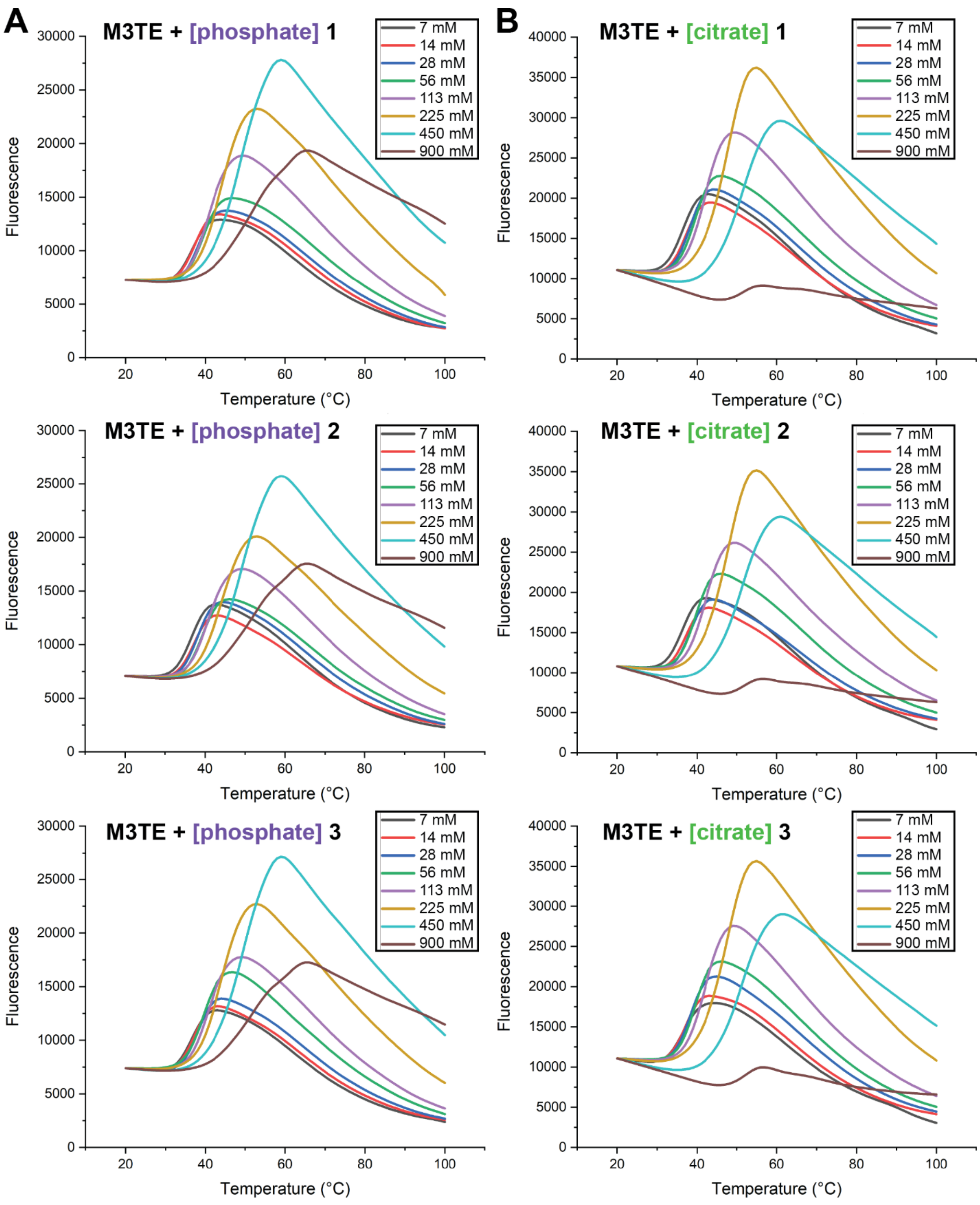


Fig. S2.

Differential scanning fluorimetry analysis of DEBS M1 (pDC1; labeled as "M1TE" here for consistency with Fig. S3). (**A,B**) Thermal unfolding melt curves displayed as fluorescence versus temperature over varying (**A**) phosphate or (**B**) citrate concentrations (7 – 900 mM). Three replicates are shown. (**C,D**) The same thermal unfolding melt curves are displayed as -d(Fluorescence)/dT versus temperature.


A
B
M3TE + [phosphate] 1
M3TE + [phosphate] 2
M3TE + [phosphate] 3
M3TE + [citrate] 1
M3TE + [citrate] 2
M3TE + [citrate] 3
7 mM
14 mM
28 mM
56 mM
113 mM
225 mM
450 mM
900 mM
Fluorescence
Temperature (°C)
0
5000
10000
15000
20000
25000
30000
35000
40000
20
40
60
80
100

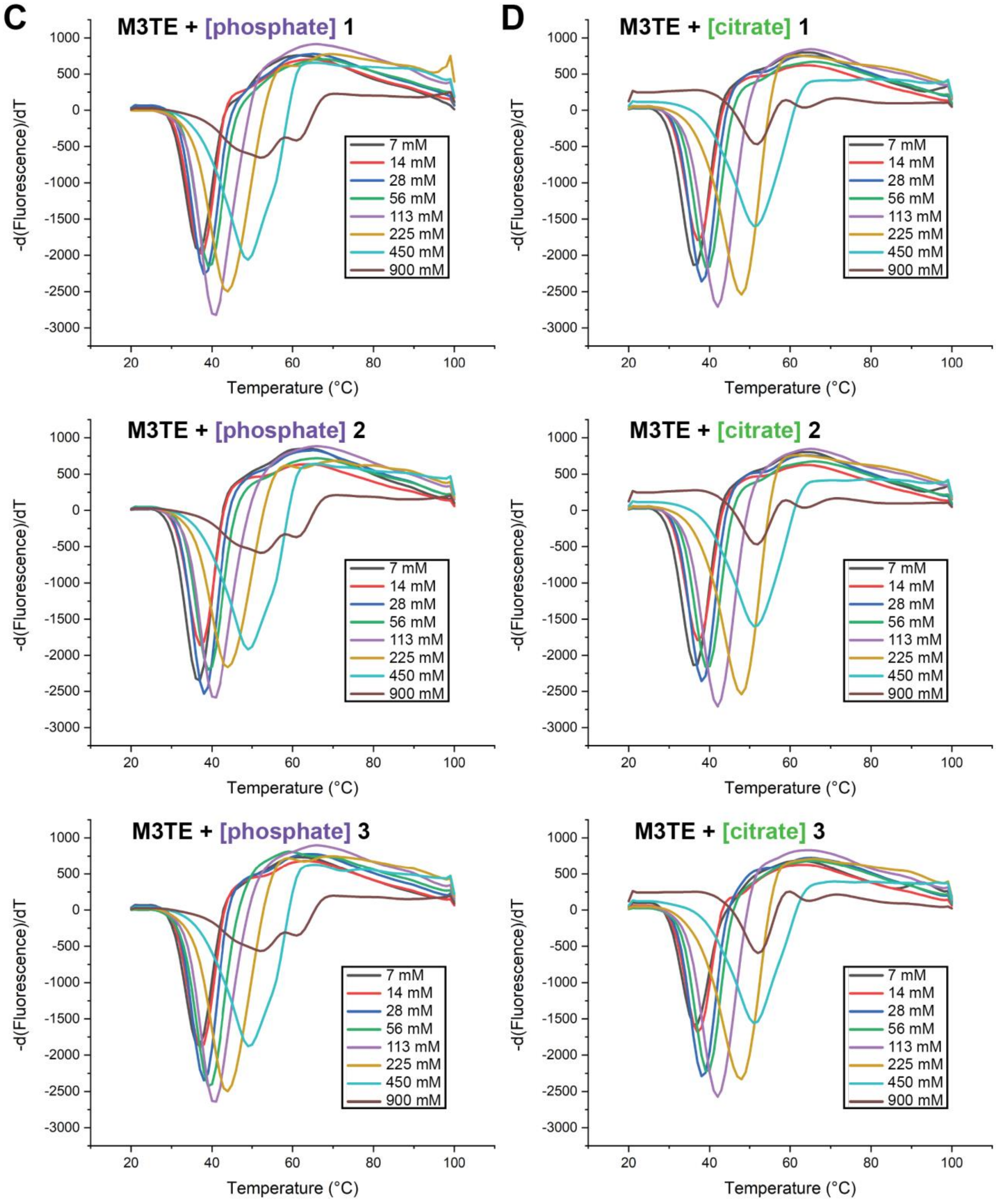


Fig. S3.

Differential scanning fluorimetry analysis of DEBS M3TE(*12*). (**A,B**) Thermal unfolding melt curves displayed as fluorescence versus temperature over varying (**A**) phosphate or (**B**) citrate concentrations (7 – 900 mM). Three replicates are shown. (**C,D**) The same thermal unfolding melt curves are displayed as -d(Fluorescence)/dT versus temperature. From the minima of these curves the average melting temperatures were calculated and used to plot figure S4.

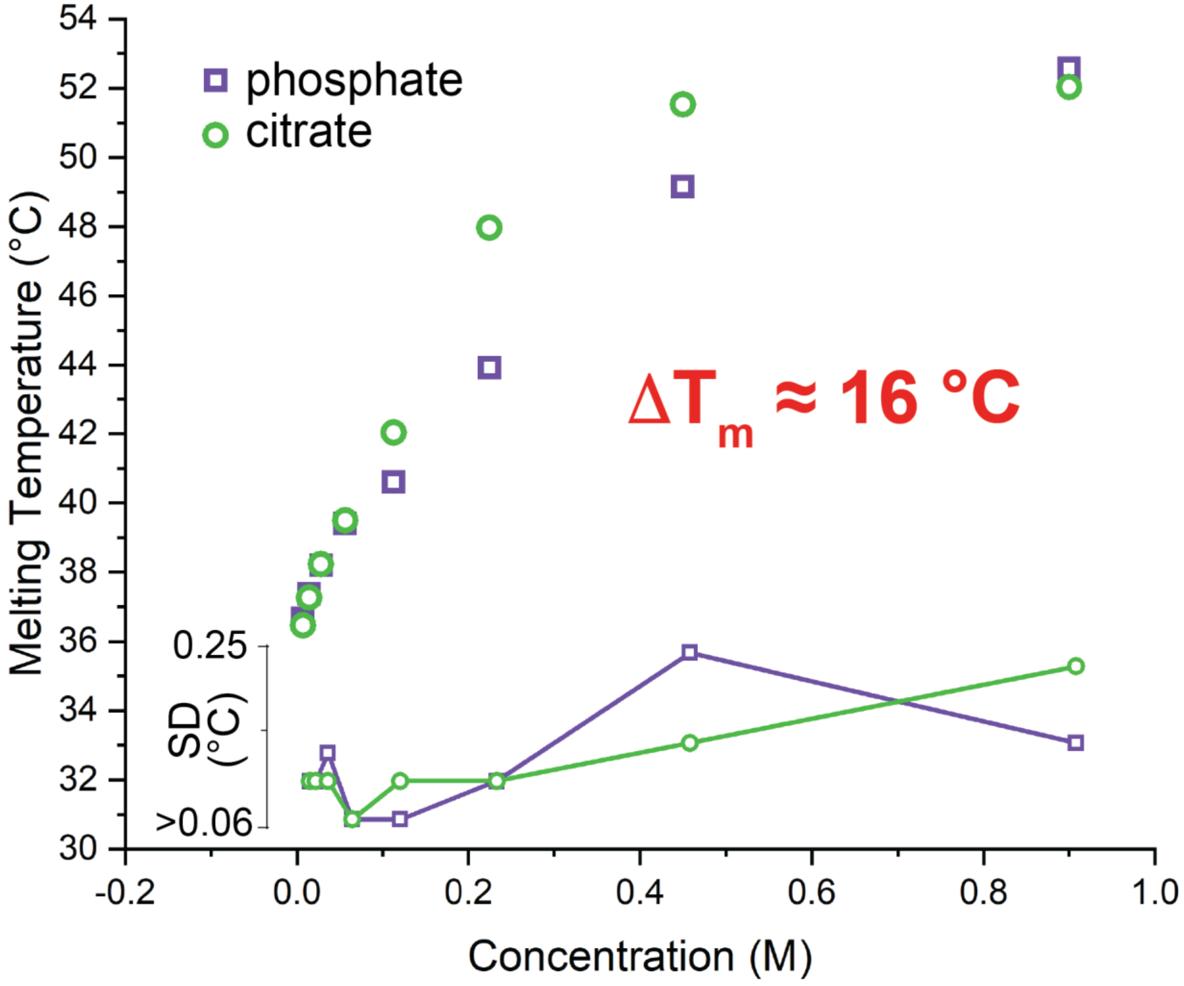


Fig. S4.
Differential scanning fluorimetry analysis of the M3TE derivative of DEBS in the presence of phosphate and citrate. The melting temperature (T_m) of a single, two-step transition is increased as a function of phosphate and citrate concentration (n=3; standard deviations are plotted below; fig. S3).

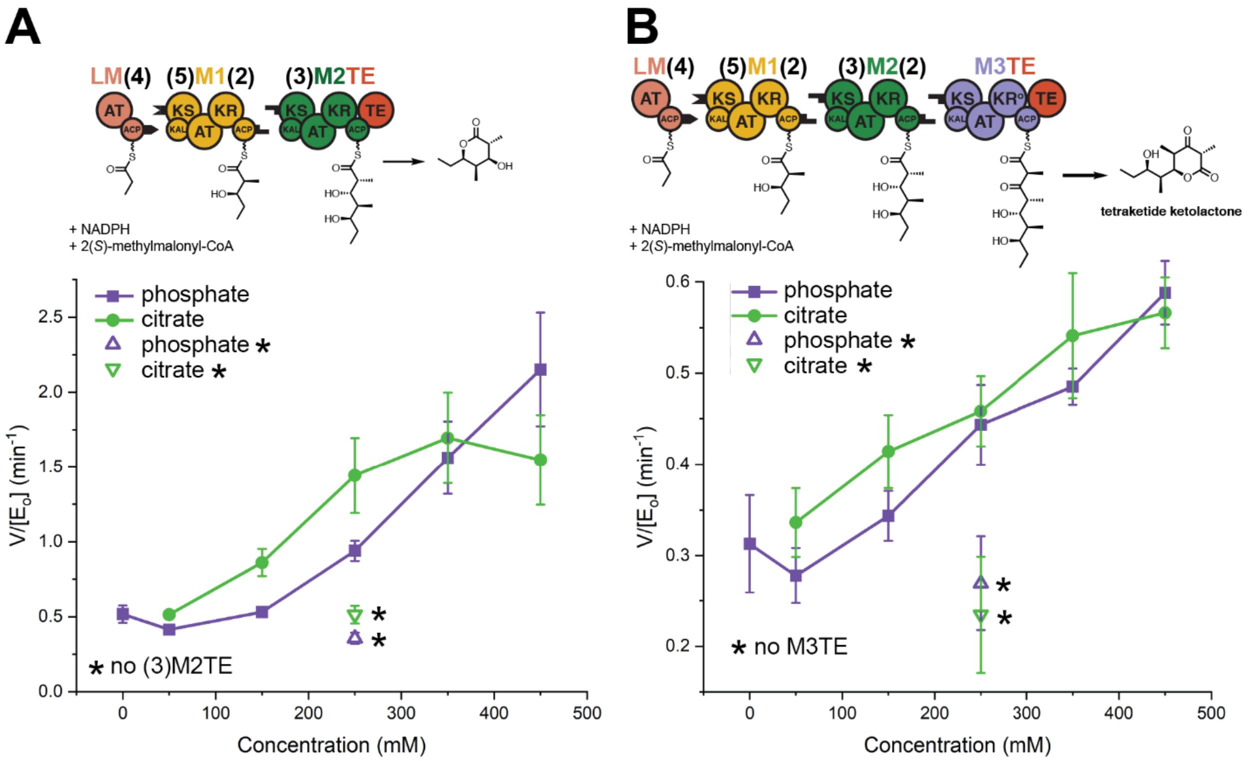


Fig. S5.
Reaction velocities of (**A**) a DEBS bimodule or (**B**) DEBS trimodule as a function of increasing phosphate and citrate concentration. Concentrations indicate the amount of sodium phosphate and sodium citrate above a baseline of 50 mM sodium phosphate (pH 7.2). The data points marked with asterisks correspond to analogous reactions in which the (3)M2TE or M3TE modules were omitted in (A) and (B), respectively (n=3; error bars = standard deviations). KAL = KS-AT linker. In panel A, citrate activation appears to be maximal around 350 mM followed by a diminution in activation at 450 mM, whereas phosphate activation is maximal at 450 mM. While it appears that phosphate is a better activator at higher concentrations (>400 mM), a considerable amount of protein precipitation was observed in reactions containing ≥350 mM citrate but not in reactions containing ≥350 mM phosphate. It remains unclear whether the diminution in activity at 450 mM citrate is due to citrate-specific effects or due to non-specific effects of protein precipitation. Parenthetical numbers correspond to the native DEBS docking domains that facilitate intermodular polyketide translocation, represented as shape-complementary black tabs(*40*).

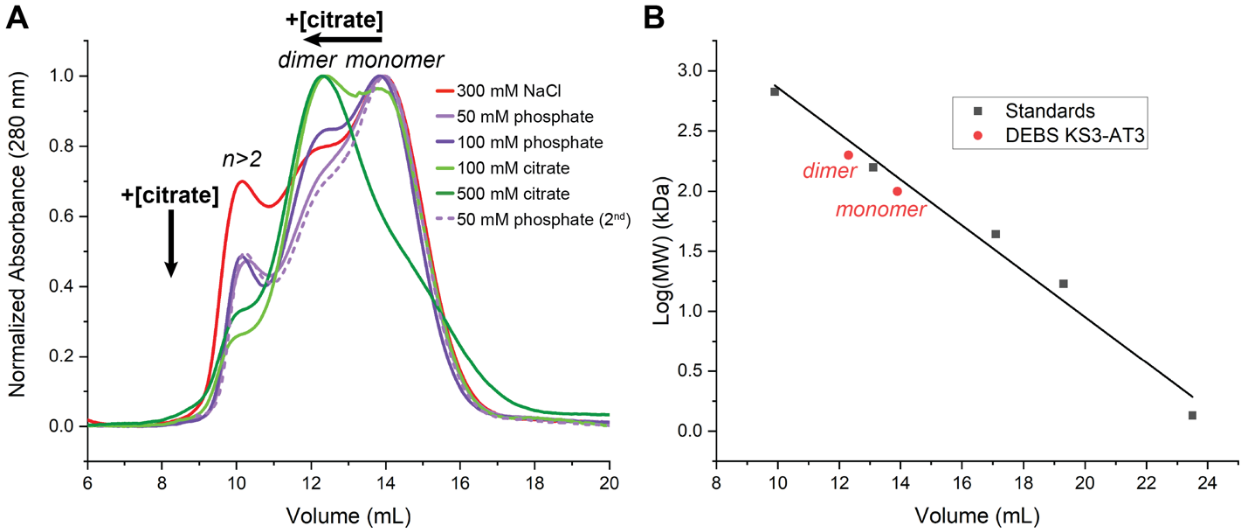


Fig. S6.
Size-exclusion chromatography (SEC) analysis of the KS-AT didomain of DEBS Module 3 (KS3-AT3)(*13*). (**A**) This KS-AT is shifted from its monomeric to dimeric state as a function of increasing phosphate and, more potently, citrate. Nearly all protein is dimeric at 500 mM citrate (for buffer details, see the Methods). KS3-AT3 was analyzed by SEC in the buffers listed from top to bottom in the same sequential order. (**B**) Semi-logarithmic plot of the molecular weights versus retention volumes for gel-filtration protein standards (Bio-Rad, Cat. # 1511901) and KS3-AT3. Gel-filtration protein standards were analyzed in a similar manner to KS3-AT3 to generate a standard curve (adjusted $R^2 = 0.98$) and to assign the oligomeric states of KS3-AT3 using the chromatograms in (A).

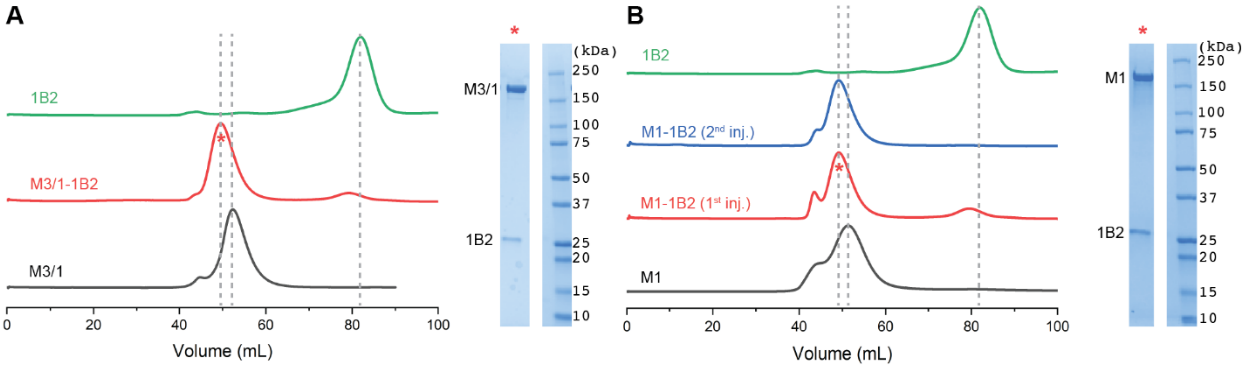


Fig. S7.
Size-exclusion chromatography (SEC) analysis of the DEBS module-F_{ab} complexes used for single-particle cryo-EM. (**A**) M3/1 (pTED23) with (red) or without (black) F_{ab} (green). (**B**) M1 (pDC1) with (red; 1st injection) or without (black) F_{ab} 1B2 (green). The major peak corresponding to the M1-F_{ab}(1B2) complex (red asterisk) was subjected to another round of SEC purification (blue; 2nd injection). The complex underwent no apparent dissociation under this time frame. (**A,B**) Peaks marked with a red asterisk were analyzed by SDS-PAGE under reducing conditions (+5 mM dithiothreitol). See the Methods for details.

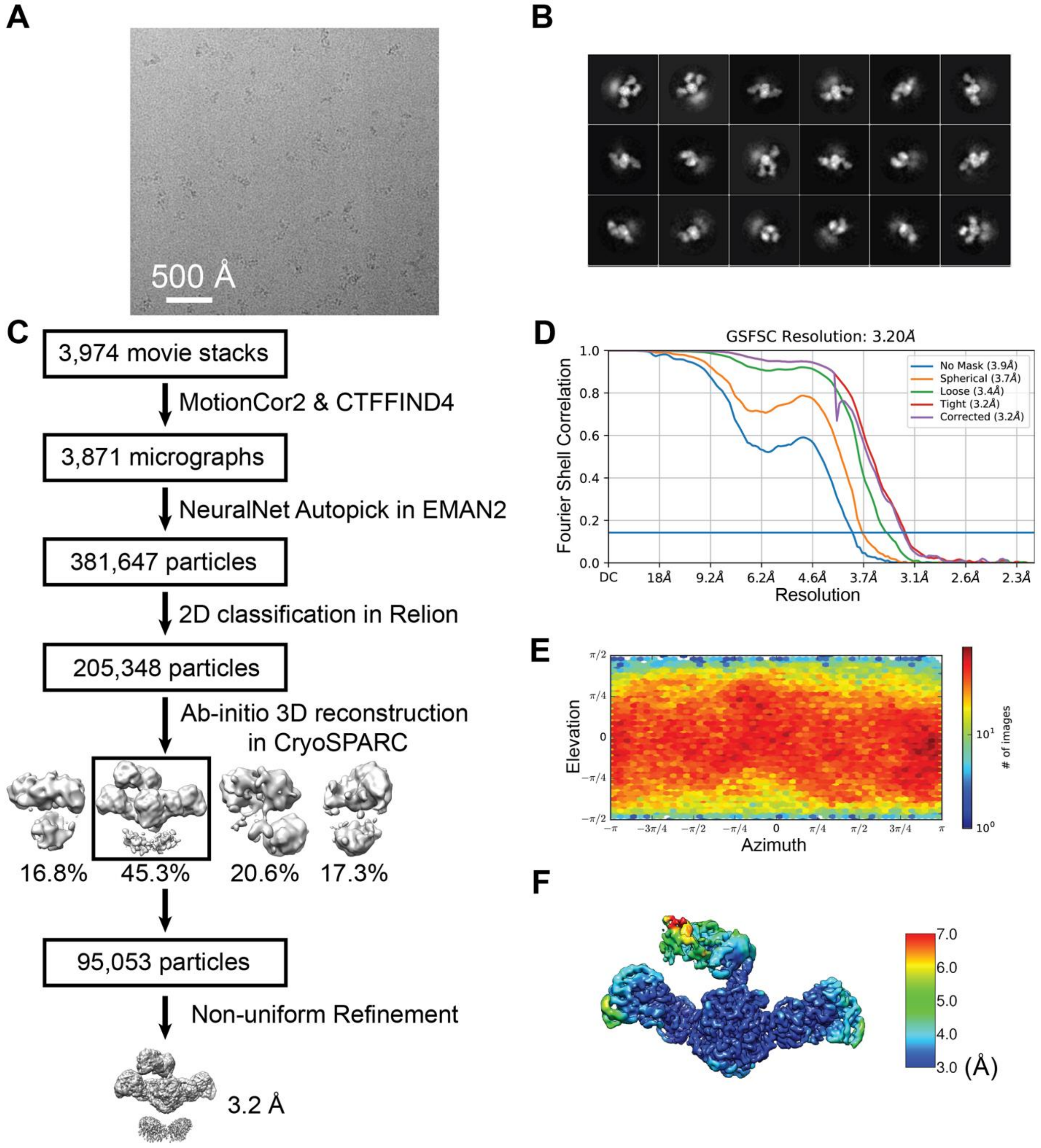


Fig. S8.
Single-particle cryo-EM analysis of DEBS M3/1-F$_{ab}$. (**A**) Representative motion-corrected cryo-EM micrograph. (**B**) Reference-free 2D class averages. (**C**) Data processing workflow (PDB 7M7E/EMD-23710)(*46*, *47*, *49*, *50*, *57*). (**D**) Gold standard FSC plot for the 3D reconstruction, calculated in cryoSPARC(*50*). (**E**) Euler angle distribution of the particle images. (**F**) Resolution map of the final 3D reconstruction.

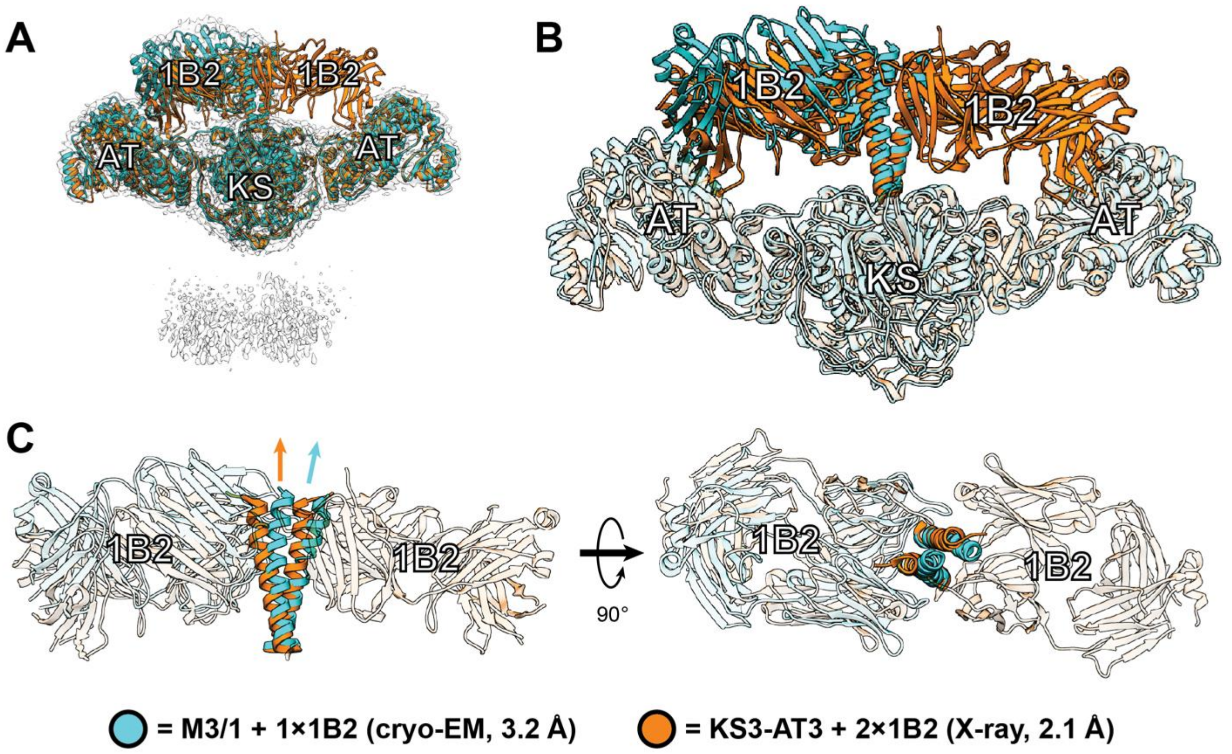


Fig. S9.

Comparison of the X-ray crystal structure of F_{ab}(1B2)-bound KS-AT from DEBS M3 (KS3-AT3; PDB 6C9U, orange)(*15*) and the cryo-EM structure of M3/1-F_{ab} (PDB 7M7E, cyan). Only one copy of the F_{ab} heterodimer is bound to the KS-AT didomain in the cryo-EM structure, whereas two symmetry-related copies of the F_{ab} are observed in the crystal structure. (**A**) Superposition of the two structures fitted into the 3.2 Å cryo-EM map. The model alignment was carried out using MatchMaker in UCSF Chimera following specification of the KS-AT regions as the reference and match chains to minimize competing alignment between the F_{ab}s (RMSD = 1.281 Å across 1469 atom pairs)(*51*, *58*). (**B**) A closer view of the superposition in panel A, with the cryo-EM map omitted for clarity. A relatively poor alignment between the F_{ab}s appears to be the result of differentially oriented coiled-coil docking domains. (**C**) Singling out the coiled-coil docking domains highlights significant differences between the two structures. Whereas the average coiled-coil helical axis is nearly parallel with the C_2 axis of symmetry in the crystal structure (orange arrow), the cryo-EM structure reveals a roughly 12° deviation from the C_2 axis (blue arrow). The 12° bend would require placement of the second, symmetry-related F_{ab} in a sub-optimal position that would clash with the AT (right-sided F_{ab} binding site in panel B). Thus, single F_{ab} occupancy is observed in the cryo-EM structure of M3/1. Consistent with this proposal, all subsequent M1-F_{ab} (pDC1) cryo-EM structures contained two copies of the F_{ab} and lacked any significant bend in their coiled-coils. The basis for the differential coiled-coil bending is unclear.

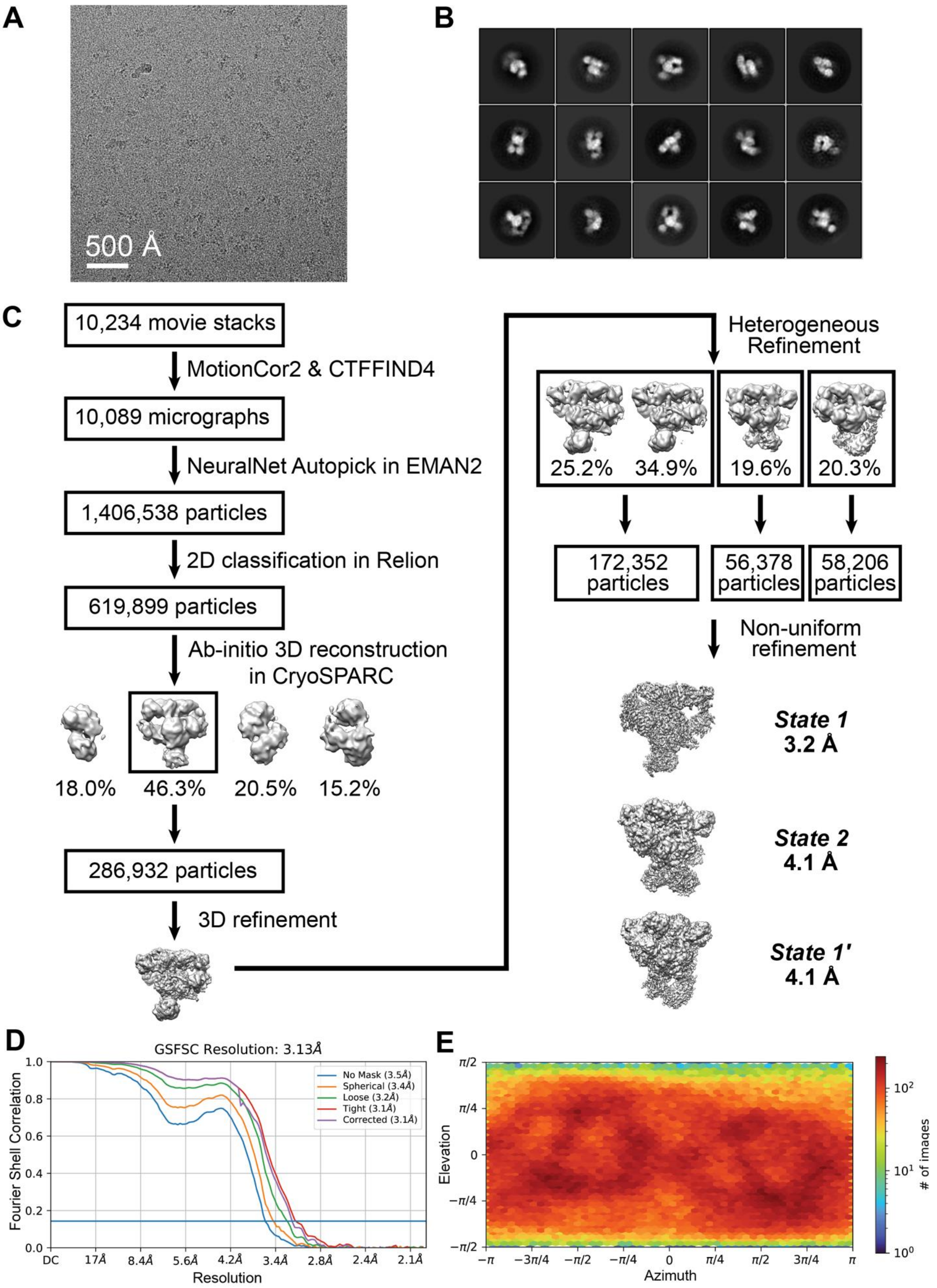

A
500 Å
B
C
10,234 movie stacks
MotionCor2 & CTFFIND4
10,089 micrographs
NeuralNet Autopick in EMAN2
1,406,538 particles
2D classification in Relion
619,899 particles
Ab-initio 3D reconstruction in CryoSPARC
18.0%
46.3%
20.5%
15.2%
286,932 particles
3D refinement
Heterogeneous Refinement
25.2%
34.9%
19.6%
20.3%
172,352 particles
56,378 particles
58,206 particles
Non-uniform refinement
State 1
3.2 Å
State 2
4.1 Å
State 1'
4.1 Å
D
GSFSC Resolution: 3.13Å
Fourier Shell Correlation
1.0
0.8
0.6
0.4
0.2
0.0
DC
17Å
8.4Å
5.6Å
4.2Å
3.4Å
2.8Å
2.4Å
2.1Å
Resolution
No Mask (3.5Å)
Spherical (3.4Å)
Loose (3.2Å)
Tight (3.1Å)
Corrected (3.1Å)
E
Elevation
π/2
π/4
0
−π/4
−π/2
−π
−3π/4
−π/2
−π/4
0
π/4
π/2
3π/4
π
Azimuth
# of images
10^0
10^1
10^2

Fig. S10.

Single-particle cryo-EM analysis of M1-F$_{ab}$ (pDC1) (**A**) Representative motion-corrected cryo-EM micrograph. (**B**) Reference-free 2D class averages. (**C**) Data processing workflow resulted in 3D maps for three sub-classes (PDB 7M7F/EMD-23711, PDB 7M7G/EMD-23712, and PDB 7M7H/EMD-23713)(*46*, *47*, *49*, *50*, *57*). (**D**) Gold standard Fourier shell correlation (FSC) plot for the 3D reconstruction of the full particle set, calculated in cryoSPARC(*50*). (**E**) Euler angle distribution of the particle images.

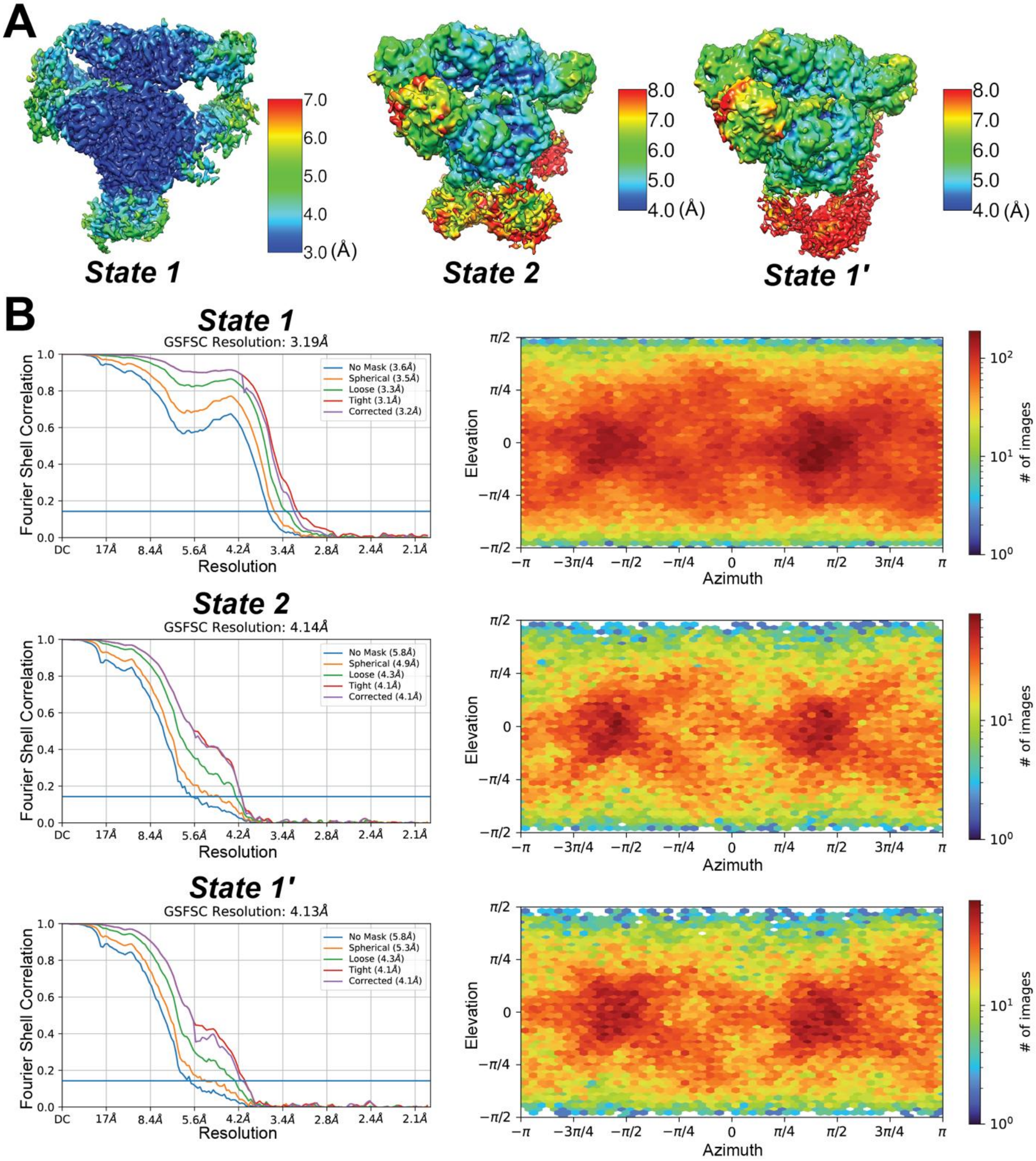
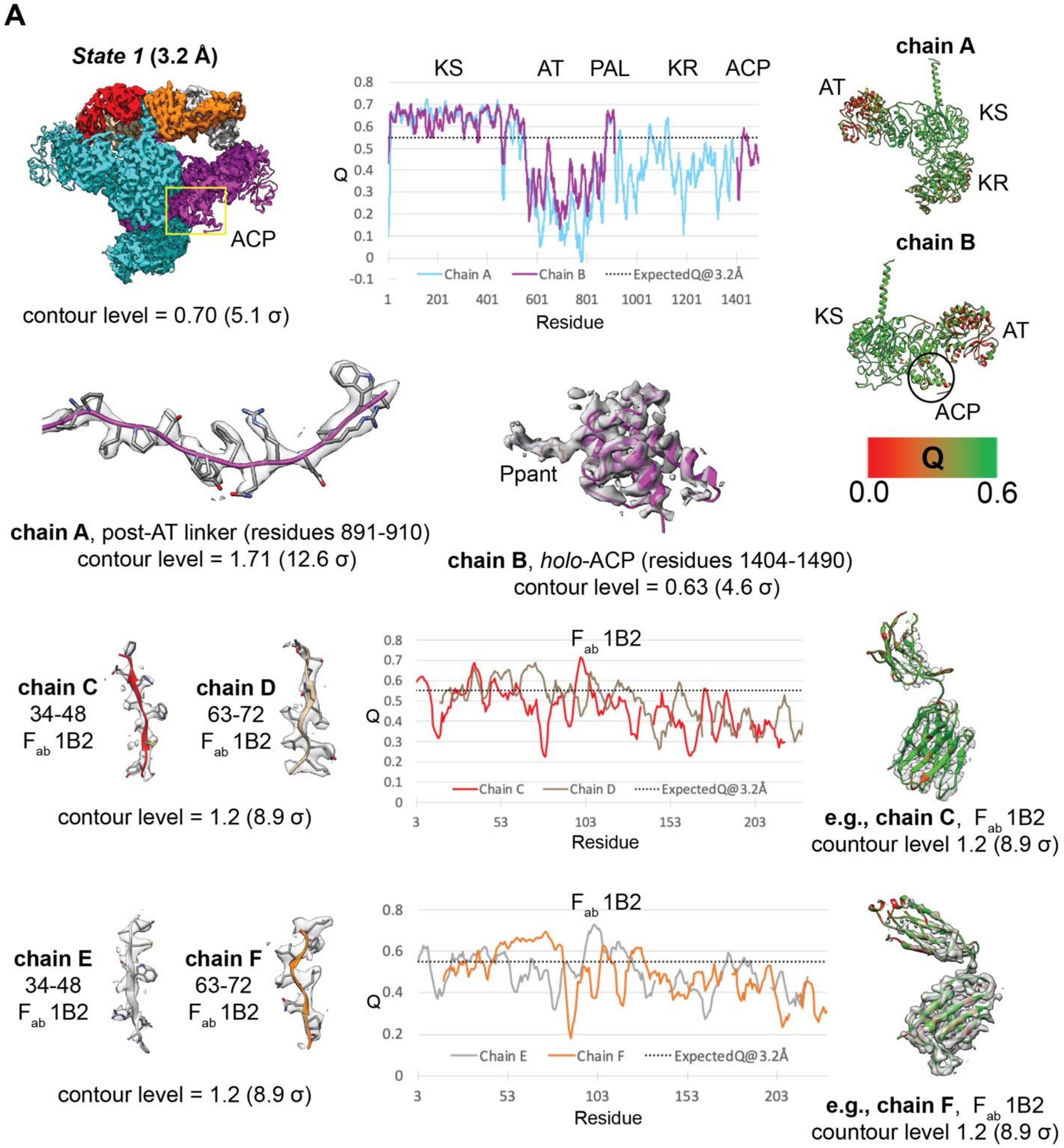


Fig. S11.

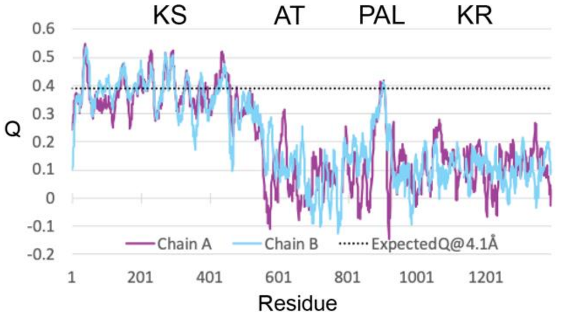
(**A**) Cryo-EM resolution maps of the final 3D reconstructions corresponding to the three M1-F_{ab} (pDC1) sub-classes (fig. S10). Note, the reduced local resolution at the KS:KR interfaces in the *State 2* map prevented any meaningful comparison with the higher resolution KS:KR interface in the *State 1* map. (**B**) Gold standard FSC plot for the 3D reconstructions from each sub-class particle set, calculated in cryoSPARC(*50*), with corresponding Euler angle distributions of the particle images to the right.

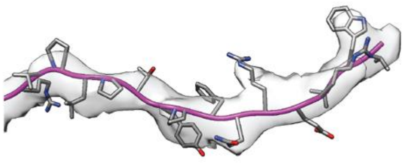

A
State 1 (3.2 Å)
ACP
contour level = 0.70 (5.1 σ)
KS
AT
PAL
KR
ACP
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
Q
1
201
401
601
801
1001
1201
1401
Residue
Chain A
Chain B
ExpectedQ@3.2Å
chain A
AT
KS
KR
chain B
KS
AT
ACP
Q
0.0
0.6
Ppant
chain A, post-AT linker (residues 891-910)
contour level = 1.71 (12.6 σ)
chain B, *holo*-ACP (residues 1404-1490)
contour level = 0.63 (4.6 σ)
chain C
34-48
F_{ab} 1B2
chain D
63-72
F_{ab} 1B2
contour level = 1.2 (8.9 σ)
F_{ab} 1B2
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Q
3
53
103
153
203
Residue
Chain C
Chain D
ExpectedQ@3.2Å
e.g., chain C, F_{ab} 1B2
countour level 1.2 (8.9 σ)
chain E
34-48
F_{ab} 1B2
chain F
63-72
F_{ab} 1B2
contour level = 1.2 (8.9 σ)
F_{ab} 1B2
0.8
0.6
0.4
0.2
0
Q
3
53
103
153
203
Residue
Chain E
Chain F
ExpectedQ@3.2Å
e.g., chain F, F_{ab} 1B2
countour level 1.2 (8.9 σ)

B

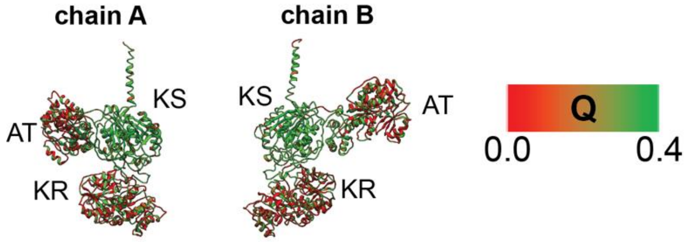
***State 2* (4.1 Å)**

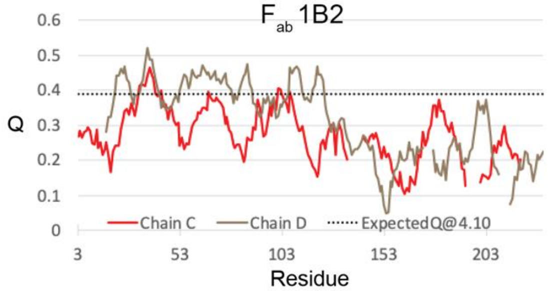
contour level = 0.40 (5.2 σ)



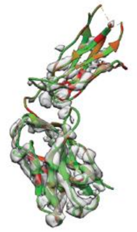


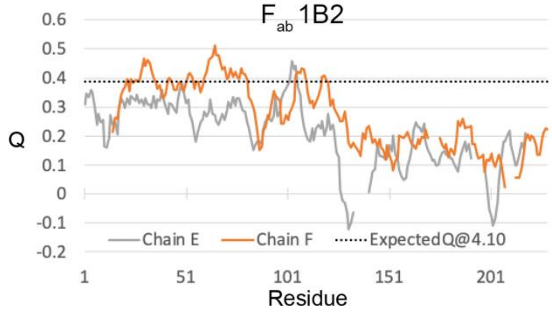
chain A, post-AT linker (residues 891-910)
contour level = 0.86 (11.3 σ)



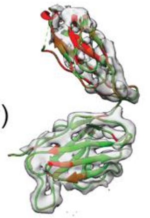


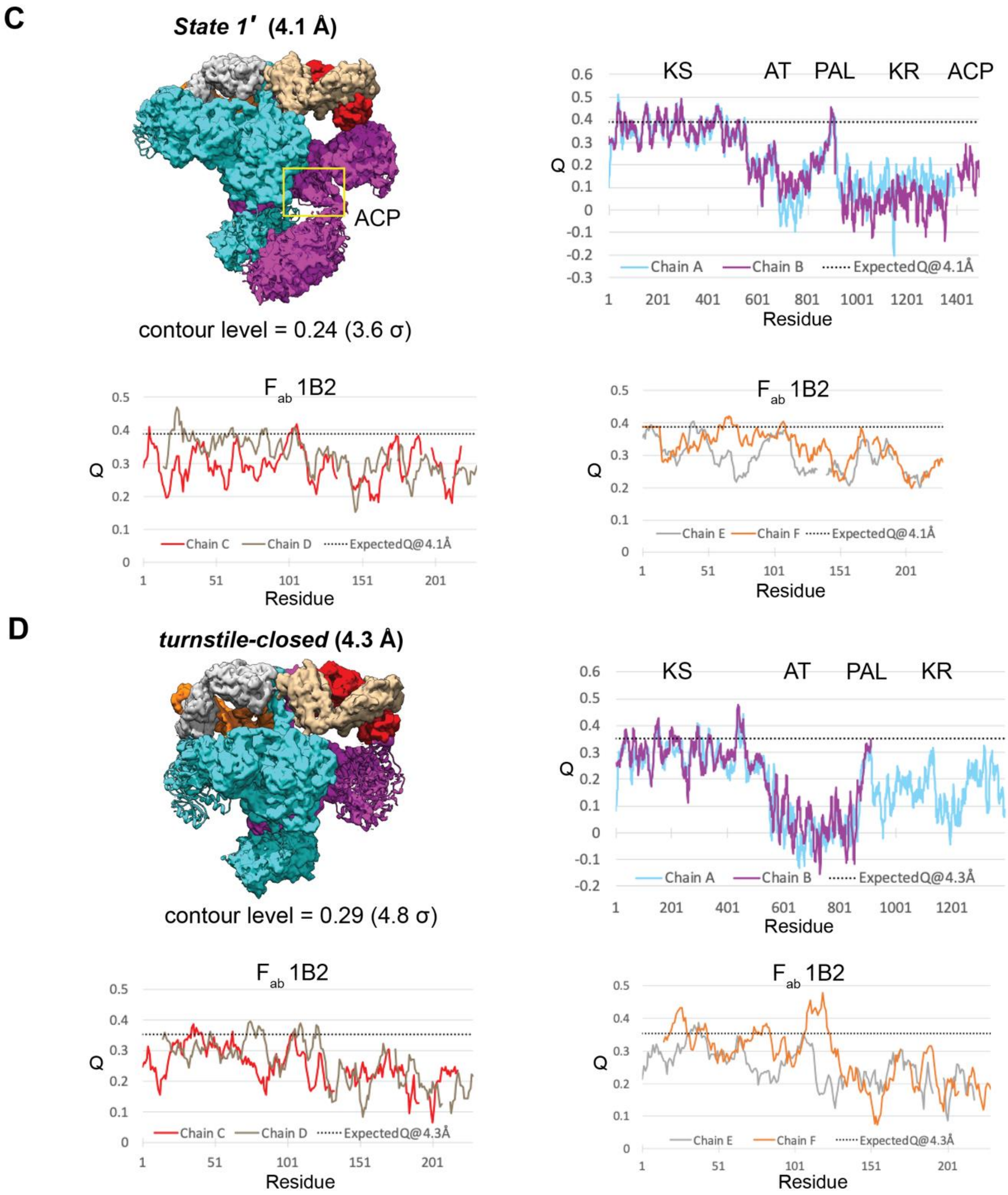
e.g., chain C, F_{ab} 1B2
countour level 1.00 (13.8 σ)





e.g., chain F, F_{ab} 1B2
countour level 0.85 (11.3 σ)




C
State 1′ (4.1 Å)
ACP
contour level = 0.24 (3.6 σ)
KS
AT
PAL
KR
ACP
Q
Residue
Chain A
Chain B
ExpectedQ@4.1Å
F$_{ab}$ 1B2
Chain C
Chain D
ExpectedQ@4.1Å
F$_{ab}$ 1B2
Chain E
Chain F
ExpectedQ@4.1Å
D
turnstile-closed (4.3 Å)
contour level = 0.29 (4.8 σ)
KS
AT
PAL
KR
Q
Residue
Chain A
Chain B
ExpectedQ@4.3Å
F$_{ab}$ 1B2
Chain C
Chain D
ExpectedQ@4.3Å
F$_{ab}$ 1B2
Chain E
Chain F
ExpectedQ@4.3Å

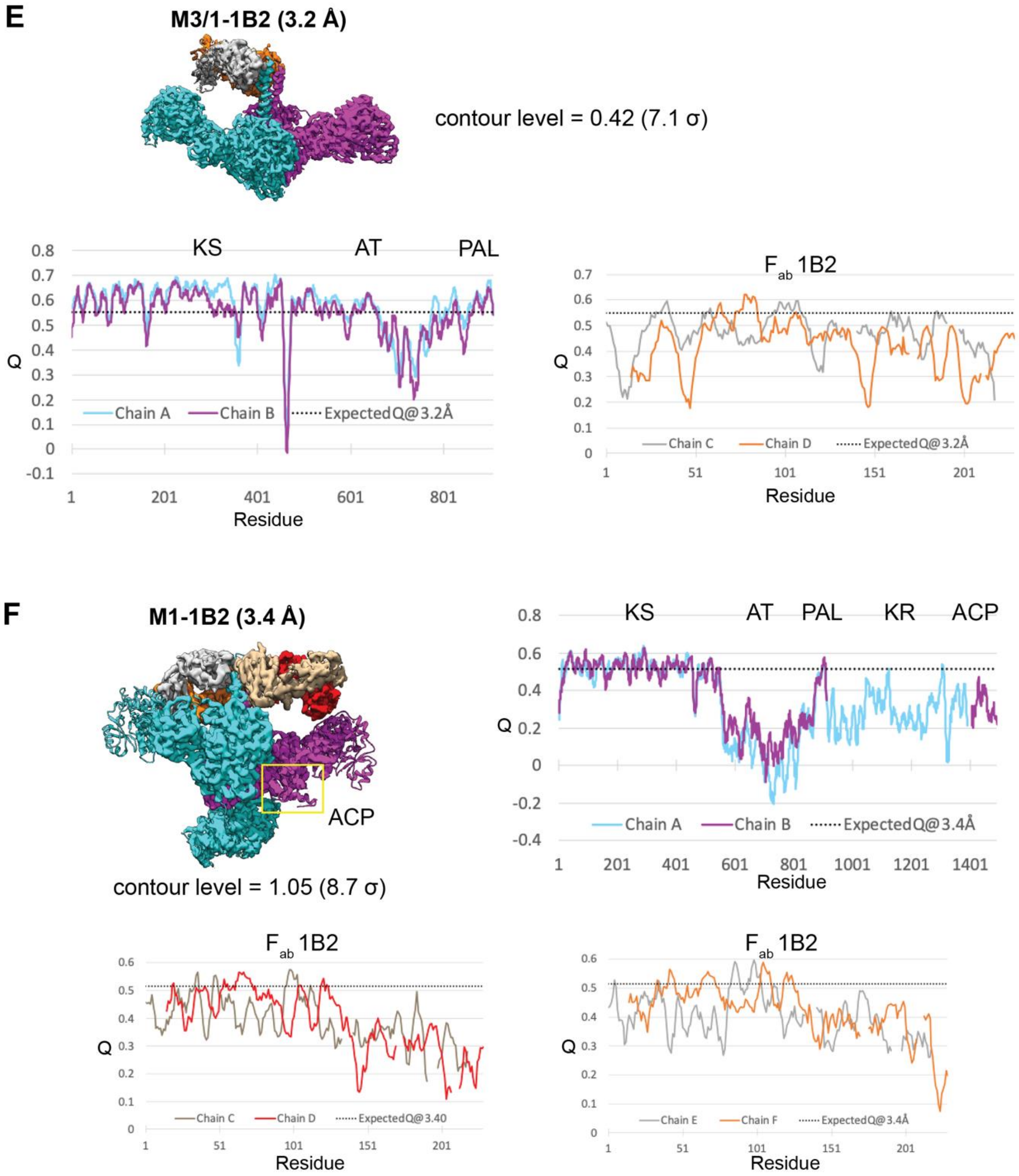


Fig. S12.
Visualization and Q-score analysis (*24*) for the (**A**) *State 1* (pDC1), (**B**) *State 2* (pDC1), (**C**) *State 1′* (pDC1), (**D**) *turnstile-closed* (pDC7), (**E**) M3/1-F_{ab} (pTED23), and (**F**) M1-F_{ab} (pDC7) models and cryo-EM maps (Ppant = 4′-phosphopantetheine; 1B2 = F_{ab}). The plots show Q-score per residue; dotted lines represent expected Q-score at the resolution of the map based on typical maps and models in the EMDB (PAL = post-AT linker connecting the AT to the KR). Regions near or above the line are resolved as expected at this resolution. Regions below the line are less resolved likely due to conformational variability.

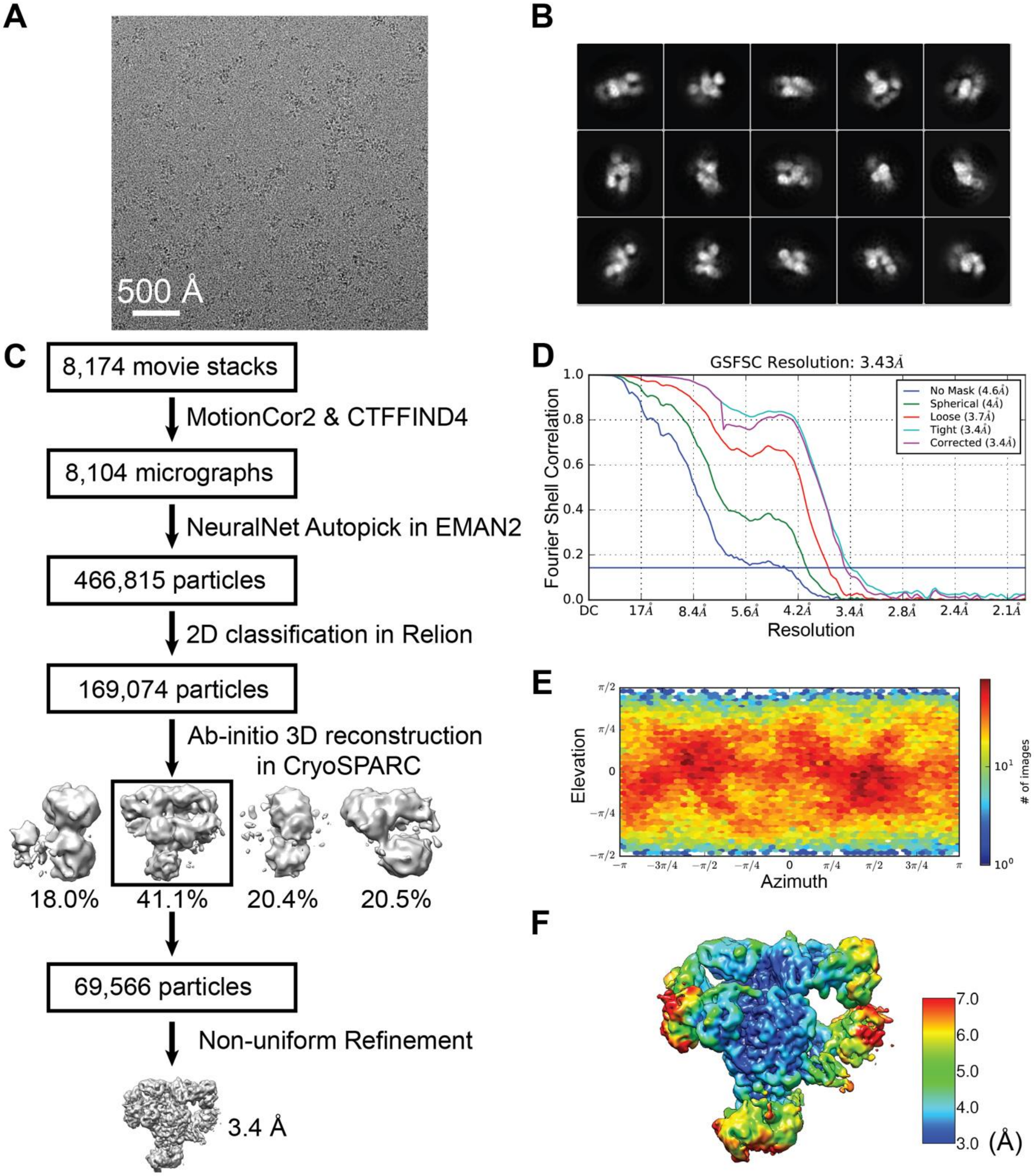


Fig. S13.
Single-particle cryo-EM analysis of M1-F_{ab} (pDC7). (**A**) Representative motion-corrected cryo-EM micrograph. (**B**) Reference-free 2D class averages. (**C**) Data processing workflow (PDB 7M7I/EMD-23714)(*46*, *47*, *49*, *50*, *57*). (**D**) Gold standard FSC plot for the 3D reconstruction, calculated in cryoSPARC(*50*). (**E**) Euler angle distribution of the particle images. (**F**) Resolution map of the final 3D reconstruction.

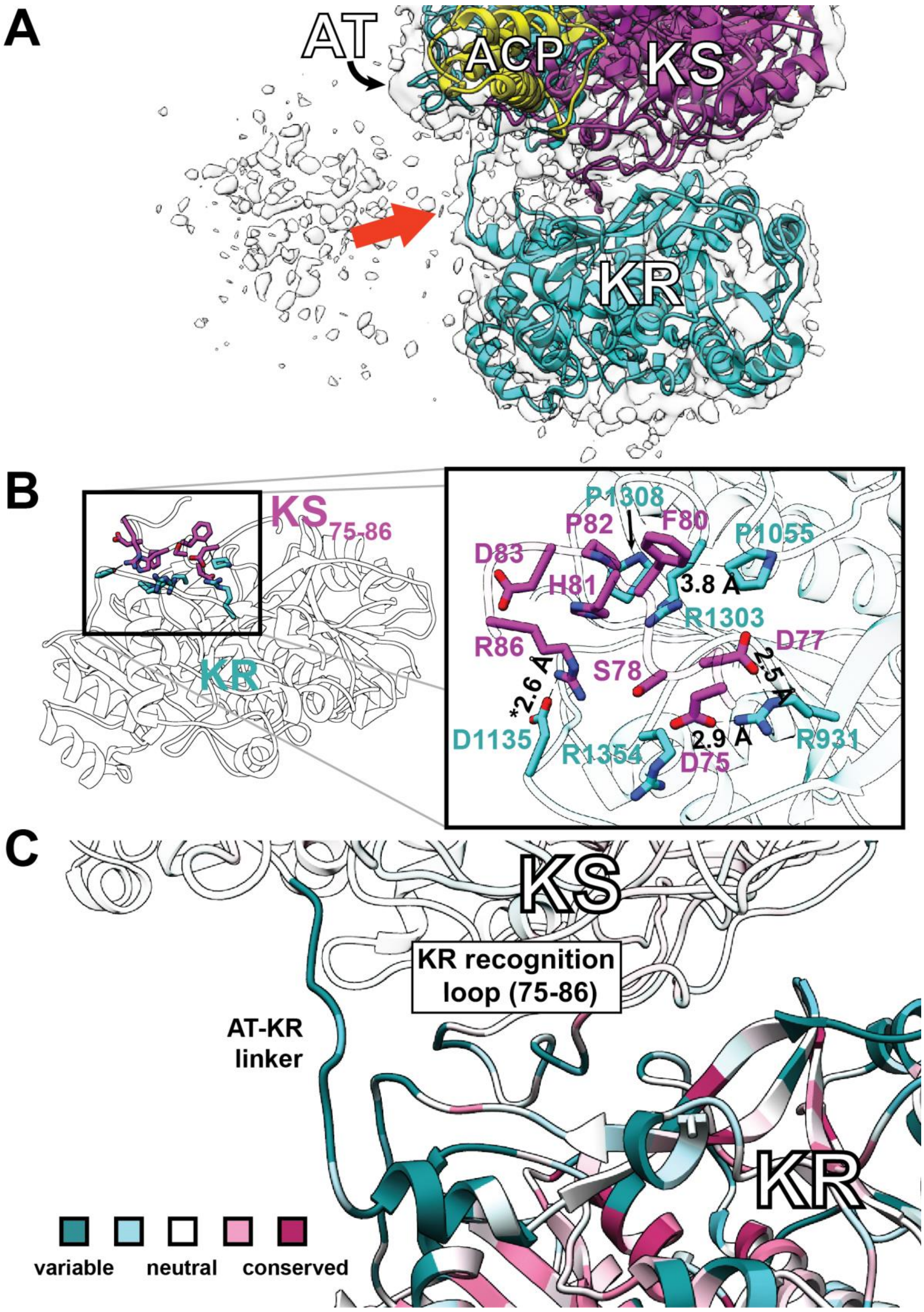
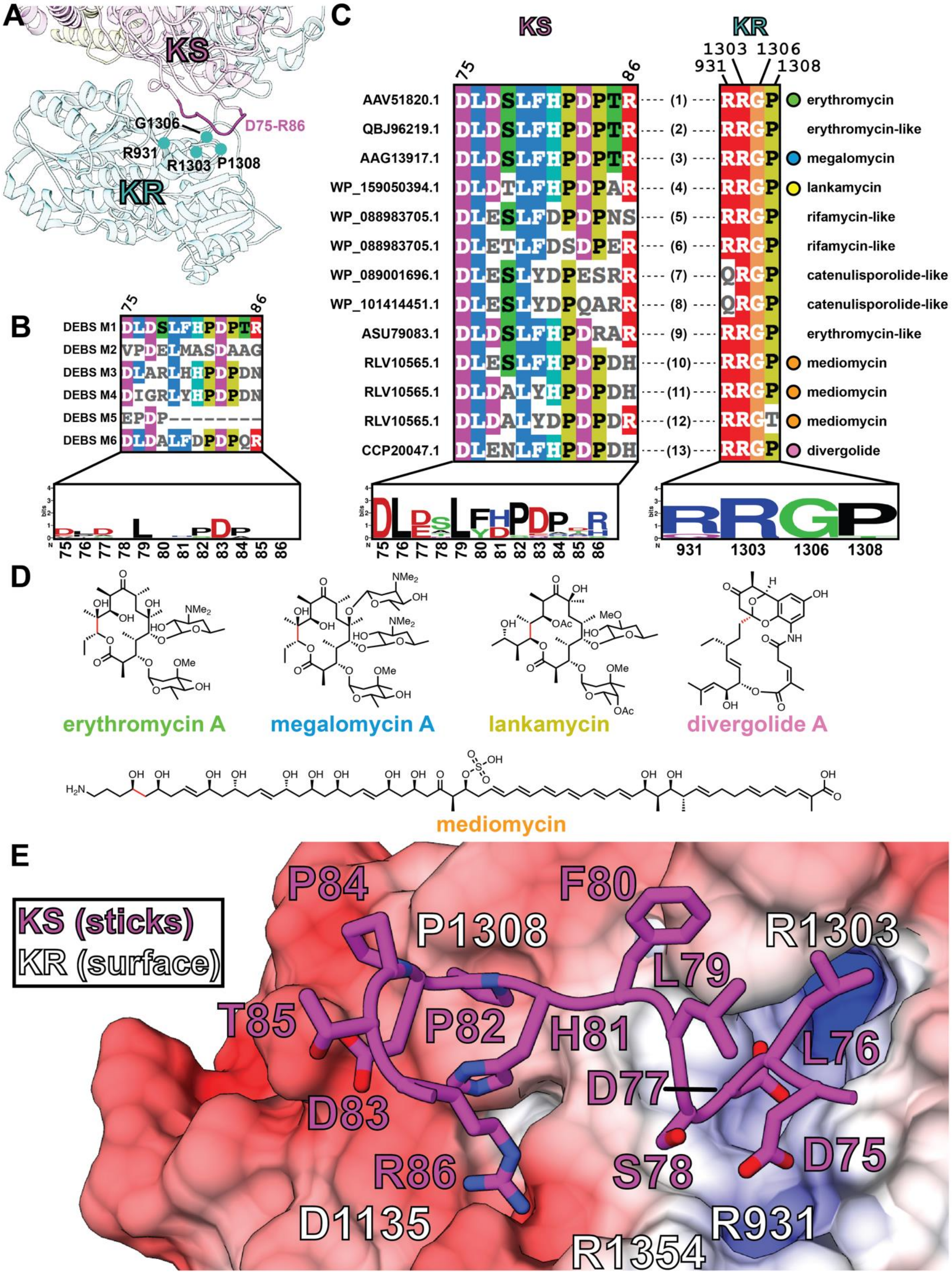

A
AT
ACP
KS
KR
B
KS75-86
KR
P1308
P82
F80
P1055
D83
H81
3.8 Å
R1303
D77
R86
S78
2.5 Å
*2.6 Å
D1135
R1354
2.9 Å
D75
R931
C
KS
KR recognition loop (75-86)
AT-KR linker
KR
variable
neutral
conserved

Fig. S14.

Establishing subunit connectivity of the KR and analysis of the KS:KR interface in the *State 1* structure of M1-F_{ab} (pDC1). (**A**) A continuous region of density bridging the AT and the KR of subunit A (cyan) from the M1-F_{ab} (pDC7) cryo-EM map (PDB 7M7I, see Supplementary Text and fig. S13) is highlighted by a red arrow and corresponds to the AT-KR linker (subunit B is in magenta and the ACP of subunit B is in yellow, for clarity). (**B**) A short unstructured loop from the KS of subunit B (residues D75 – R86) serves as the principal surface for KR (subunit A) recognition (PDB 7M7F). The KS:KR residue interactions are boxed and expanded (*the R86:D1135 interaction is not supported by the cryo-EM density but may reasonably occur based on proximity). Whereas the KS residues are clustered in a single 12-residue stretch, the KR residues are more dispersed across the KR sequence. (**C**) ConSurf analysis of M1 whereby 300 homologs were used in a multiple sequence alignment to generate conservation scores that are mapped onto the cryo-EM structure *State 1* (color-coded legend in the bottom left)(*59*). A significant proportion of residues within the AT-KR linker and KR recognition loop (KS residues D75 – R86) have variable sequences, in addition to KR residues proximal to this recognition loop in the KS domain.



A
KS
G1306
D75-R86
R931
R1303
P1308
KR
B
75
86
DEBS M1 DLDSLFHPDPTR
DEBS M2 VPDELMASDAAG
DEBS M3 DLARLHHPDPDN
DEBS M4 DIGRLYHPDPDN
DEBS M5 EPDP--------
DEBS M6 DLDALFDPDPQR
bits
75 76 77 78 79 80 81 82 83 84 85 86
C
KS
KR
1303 1306
931 1308
AAV51820.1 DLDSLFHPDPTR (1) RRGP erythromycin
QBJ96219.1 DLDSLFHPDPTR (2) RRGP erythromycin-like
AAG13917.1 DLDSLFHPDPTR (3) RRGP megalomycin
WP_159050394.1 DLDTLFHPDPAR (4) RRGP lankamycin
WP_088983705.1 DLESLFDPDPNS (5) RRGP rifamycin-like
WP_088983705.1 DLETLFDSDPER (6) RRGP rifamycin-like
WP_089001696.1 DLESLYDPESRR (7) QRGP catenulisporolide-like
WP_101414451.1 DLESLYDPQARR (8) QRGP catenulisporolide-like
ASU79083.1 DLDSLFHPDRAR (9) RRGP erythromycin-like
RLV10565.1 DLESLFHPDPDH (10) RRGP mediomycin
RLV10565.1 DLDALYHPDPDH (11) RRGP mediomycin
RLV10565.1 DLDALYDPDPDR (12) RRGT mediomycin
CCP20047.1 DLENLFHPDPDH (13) RRGP divergolide
931 1303 1306 1308
D
erythromycin A
megalomycin A
lankamycin
divergolide A
mediomycin
E
KS (sticks)
KR (surface)
P84
F80
P1308
R1303
L79
T85
P82
H81
L76
D77
D83
S78
D75
R86
D1135
R931
R1354

Fig. S15.

Sequence analysis of the KS:KR interface of DEBS M1-M6 and DEBS M1 homologs. (**A**) Macroscopic view of the KS:KR interface of the *State 1* structure featuring the KR recognition loop (i.e., KS residues D75–R86) and interacting KR residues R931, R1303, G1306, and P1308 used in the sequence analysis (PDB 7M7F). (**B**) Comparison of the KR recognition loop sequences of DEBS M1-M6 (residue numbering based on DEBS M1). The consensus sequence is shown below(*60*). (**C, left**) Comparison of the KR recognition loop sequences of 12 DEBS M1 homologs identified by NCBI BLAST (100 – 50% sequence identity)(*61*). The multiple sequence alignment (MSA) corresponds to the extracted KR recognition loops from within each sequence that are most homologous to that of DEBS KS1 (accession numbers shown for entries 1–13). The consensus sequence is shown below(*60*). (**C, right**) Comparison of the interacting KR residues (non-continuous in sequence) identified by MSA of the extracted KR domains that are part of the same module as the KS analyzed in the left MSA. For entries without characterized products, antiSMASH was used to predict features of their associated biosynthetic gene clusters (color-coded to the right)(*34*, *62*). (**D**) Structures of the associated natural products, or primary congeners, for entries 1–13. The C-C bonds highlighted in red are formed by each of the DEBS M1 homologs used in the sequence analysis. (**E**) The KR recognition loop (magenta) is overlaid onto the electrostatic potential surface of the KR, generated using PDB2PQR(*63*, *64*) and APBS(*65*), to visualize the electrostatic complementarity that supports the KS:KR interface – a feature that is partially conserved across the 13 entries (Fig. 3A,B and fig. S14B).

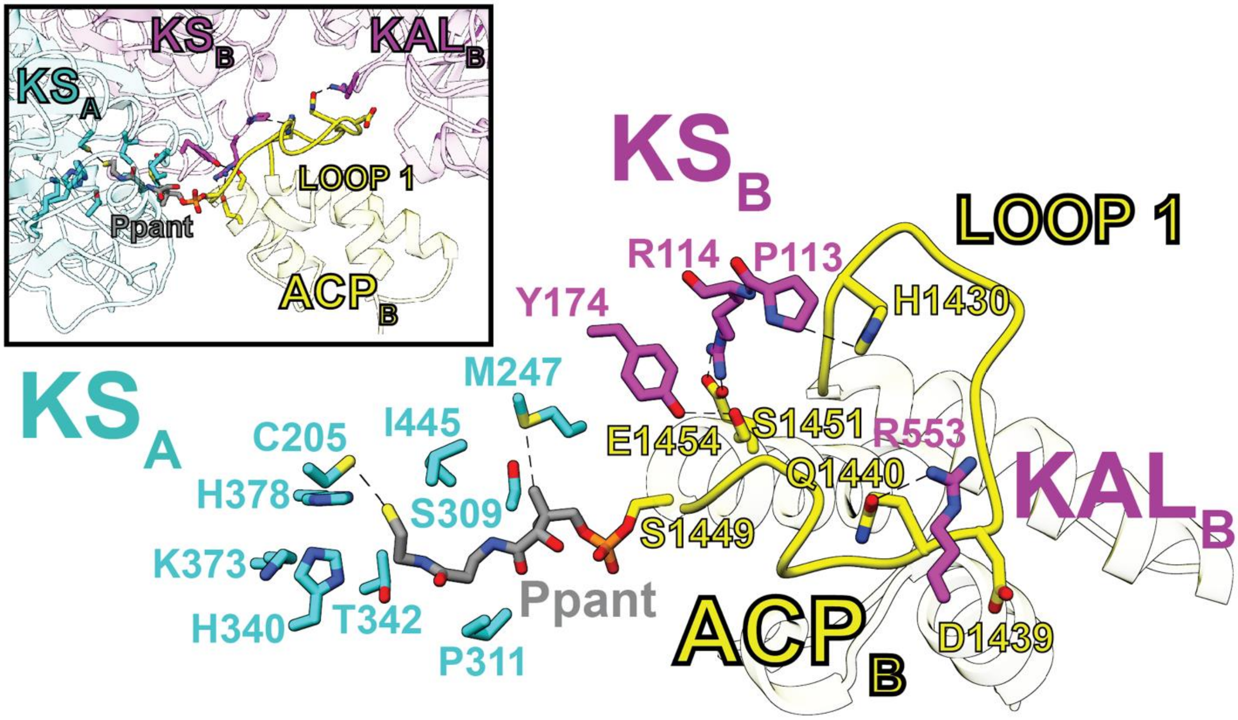


Fig. S16.

Focused view of the *holo*-ACP bound to its KS-AT cleft in the *State 1* structure (PDB 7M7F). This binding pose is proposed to be relevant to polyketide elongation given that the 4′-phosphopantetheine (Ppant) arm is bound in the KS active site and that loop 1 of the ACP donates most of the ACP surface involved in KS-AT cleft contacts – supported by previous mutational analysis(*29*, *30*). Helix 2 of the ACP contributes additional residues S1451 and E1454 that interact with the KS from its own subunit (B). Drawn interatomic distances are ≤ 3.9 Å. The inset features the same pose with ribbon diagrams to emphasize that loop 1 serves as the primary ACP epitope for KS-AT cleft recognition.

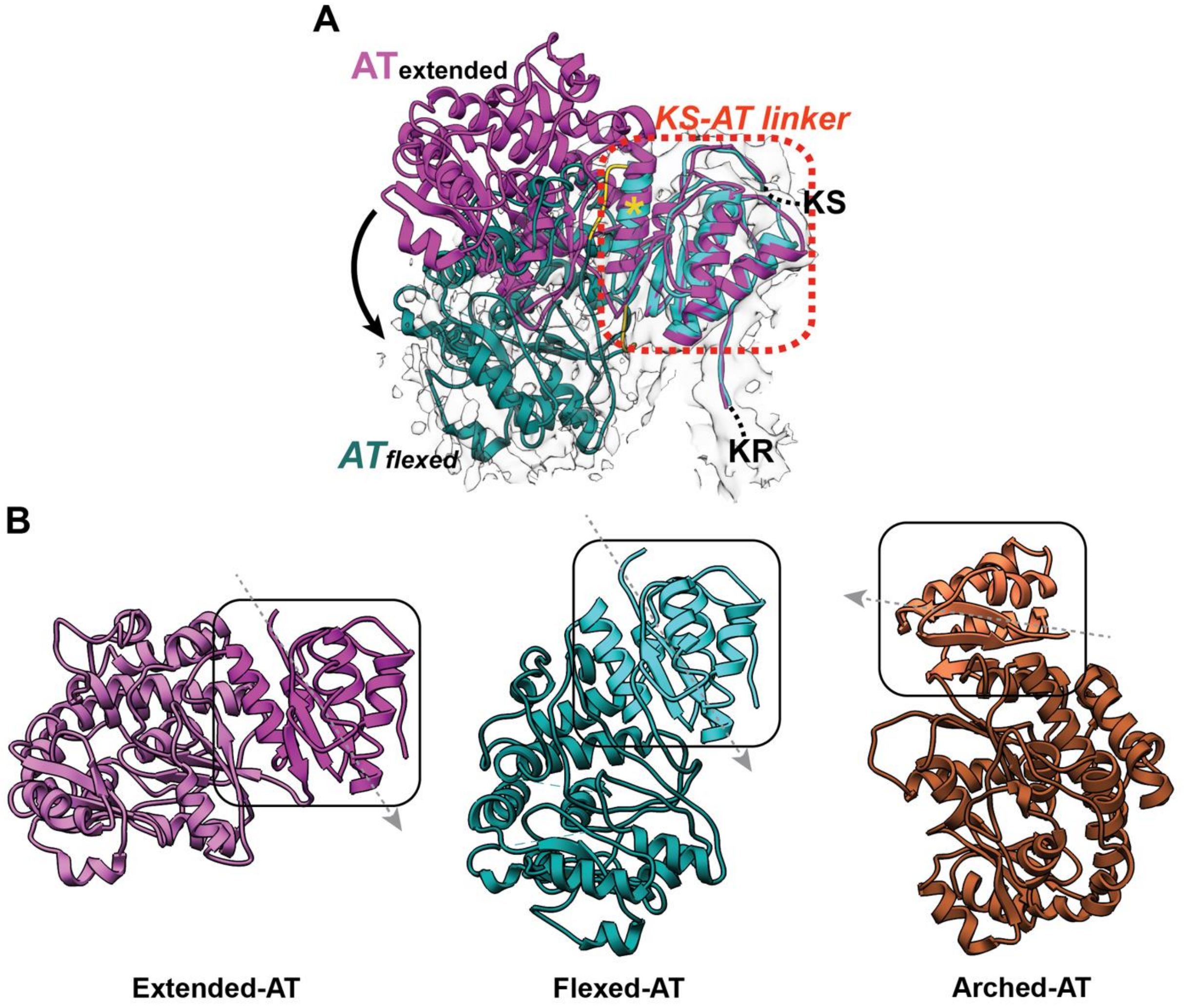


Fig. S17.

Conformational “flexing” of the AT domain in the *State 2* structure captured by molecular dynamics flexible fitting (MDFF)(*31*, *32*). (**A**) The pre-MDFF extended AT (magenta) transitions to the flexed state (dark cyan AT) during the course of the MDFF simulation (fitted into the *State 2* cryo-EM map). The superposed KS-AT linker subdomains before (magenta) and after (cyan) MDFF, boxed in red (dashed), highlights its conformational rigidity during AT flexing. Included in the rigid KS-AT linker is the helix of Leu865 – Ala878 which is C-terminal to the AT; marked with an asterisk. Residues that border the KS-AT linker and flexed AT (i.e., Thr551 – Gln555 and Arg860 – Ser864) are shown in gold (see accompanying Movie S1). (**B**) Comparison of the extended, flexed, and arched conformations of AT domains and their associated KS-AT linker, which are shaded differently and boxed for clarity. The models were first superposed via alignment of their KS domains. Dashed arrows are drawn nearly parallel with the first β-strand of the KS-AT linker to clearly visualize its relative orientations in these models. Whereas the linker deriving from extended and flexed ATs is similarly oriented, in the arched model it deviates by ~120°(*7*).

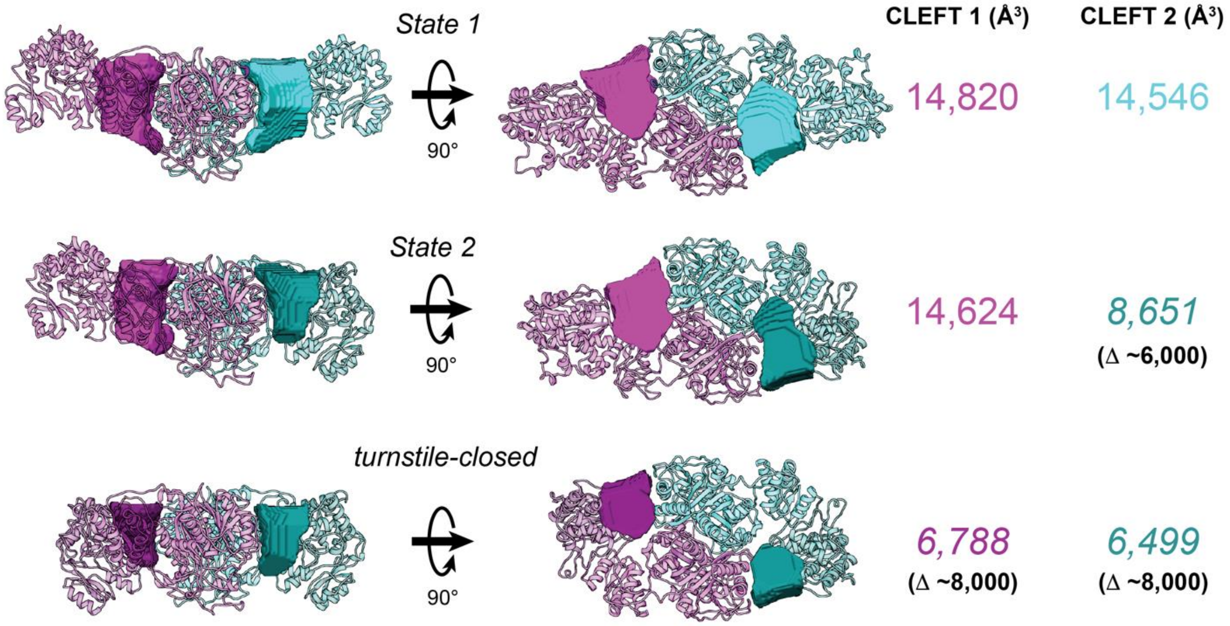


Fig. S18.

Comparison of the extended versus flexed KS-AT cleft volumes. To simplify the cleft volume calculation and visual representation, N-terminal coiled-coil docking, ACP, and KR domains were removed from the *State 1* (PDB 7M7F), *State 2* (PDB 7M7G), and *turnstile-closed* (PDB 7M7J) structures used in the analysis. KS-AT cleft volumes were determined using the Voss Volume Voxelator (3V) "channel finder" function(*66*). An outer probe radius of 30 Å and an inner probe radius of 6 Å were used for each of the determined clefts.

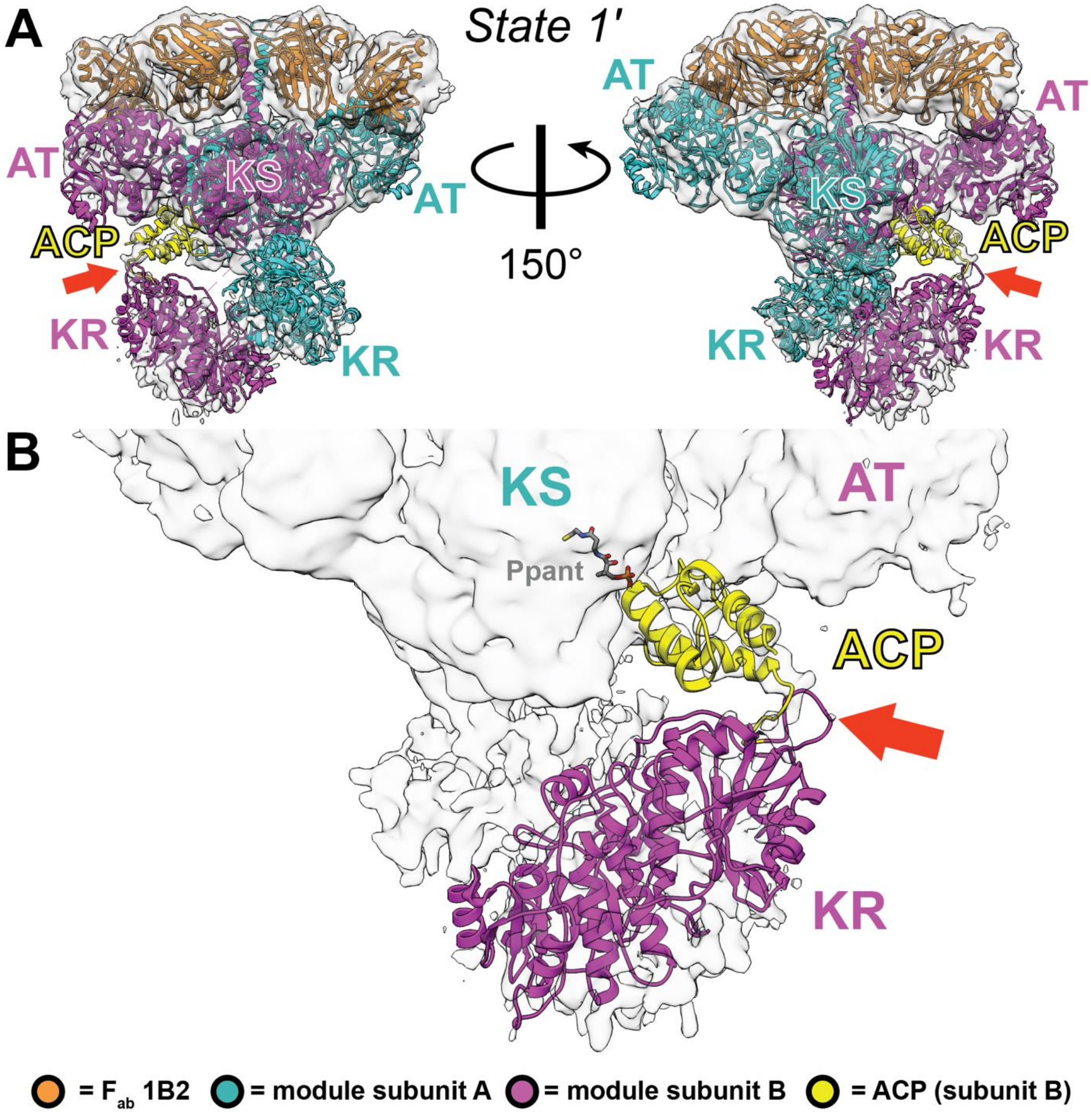


Fig. S19.

(**A**) The 4.1 Å cryo-EM *State 1′* structure of F_{ab}-bound M1 (pDC1; PDB 7M7H). Shown are two 150° related orientations and a red arrow highlighting the continuous map density between the KR and ACP of subunit B (magenta). (**B**) A closer view of the pose on the right in (A) with only the KR and ACP of subunit B shown for clarity, along with the cryo-EM map (Ppant = 4′-phosphopantetheine). The KR from subunit B in *State 1′* was fitted as a rigid body into the 4.1 Å cryo-EM map, followed by MDFF (*31*, *32*). Due to reduced local resolution in the regions of the KR and ACP, a simplified linkage was modeled in which all side chains (excepting Pro) were truncated up to the β-carbon in residues D1363–A1409.

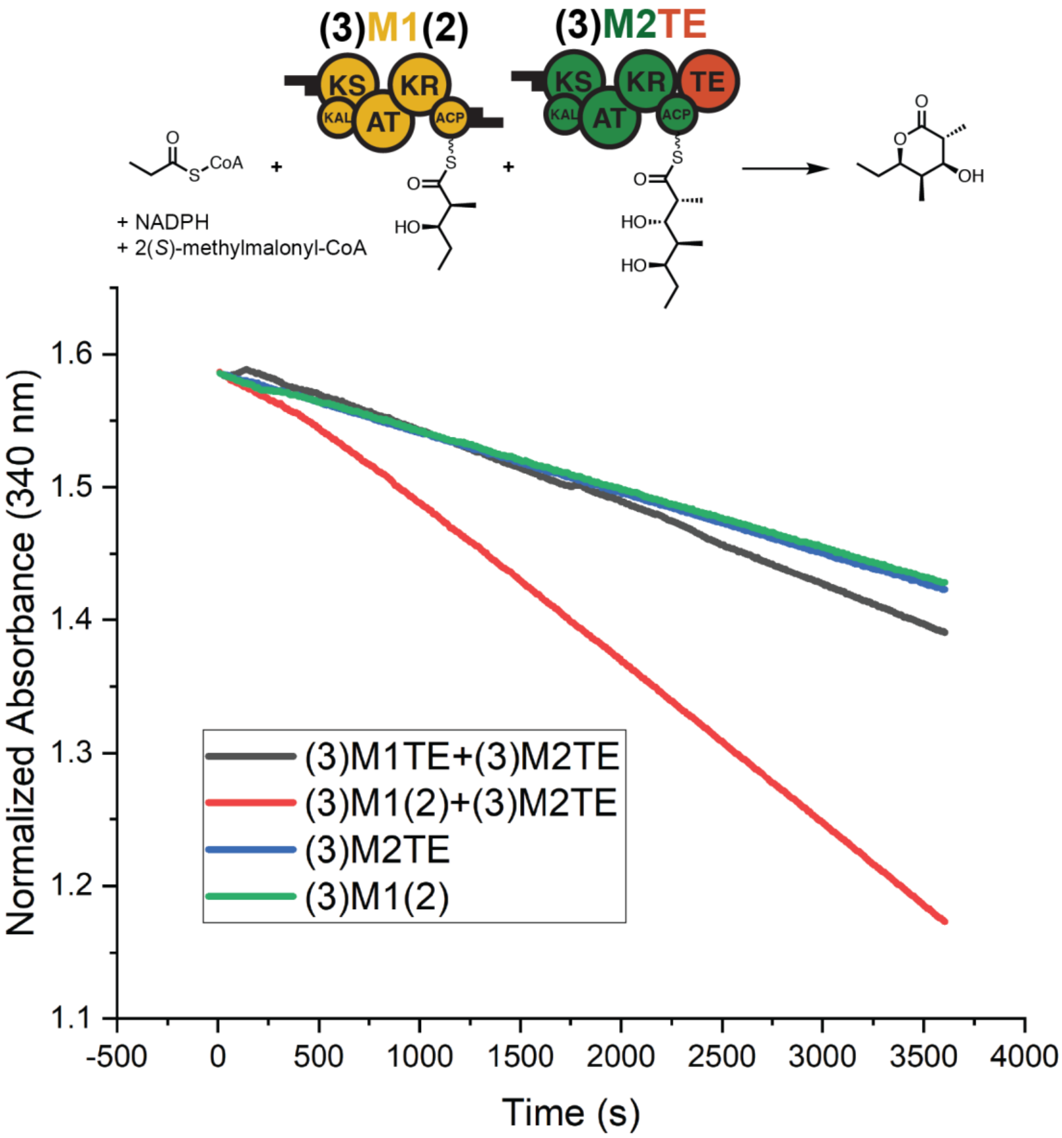


Fig. S20.
Catalytic activity of DEBS M1 in the context of a bimodule and a propionyl-CoA starter unit. NADPH depletion is monitored through absorbance at 340 nm as an indirect measure of triketide lactone formation(*14*). See the Methods for reaction details. Shape-complementary black tabs at the end of each module represent native DEBS docking domains that facilitate intermodular polyketide translocation(*40*), expressed as parenthetical numbers, and "KAL" denotes the KS-AT linker. (3)M1TE = "M1" derived from plasmid pDC1; (3)M1(2) = "M1" derived from plasmid pDC7.

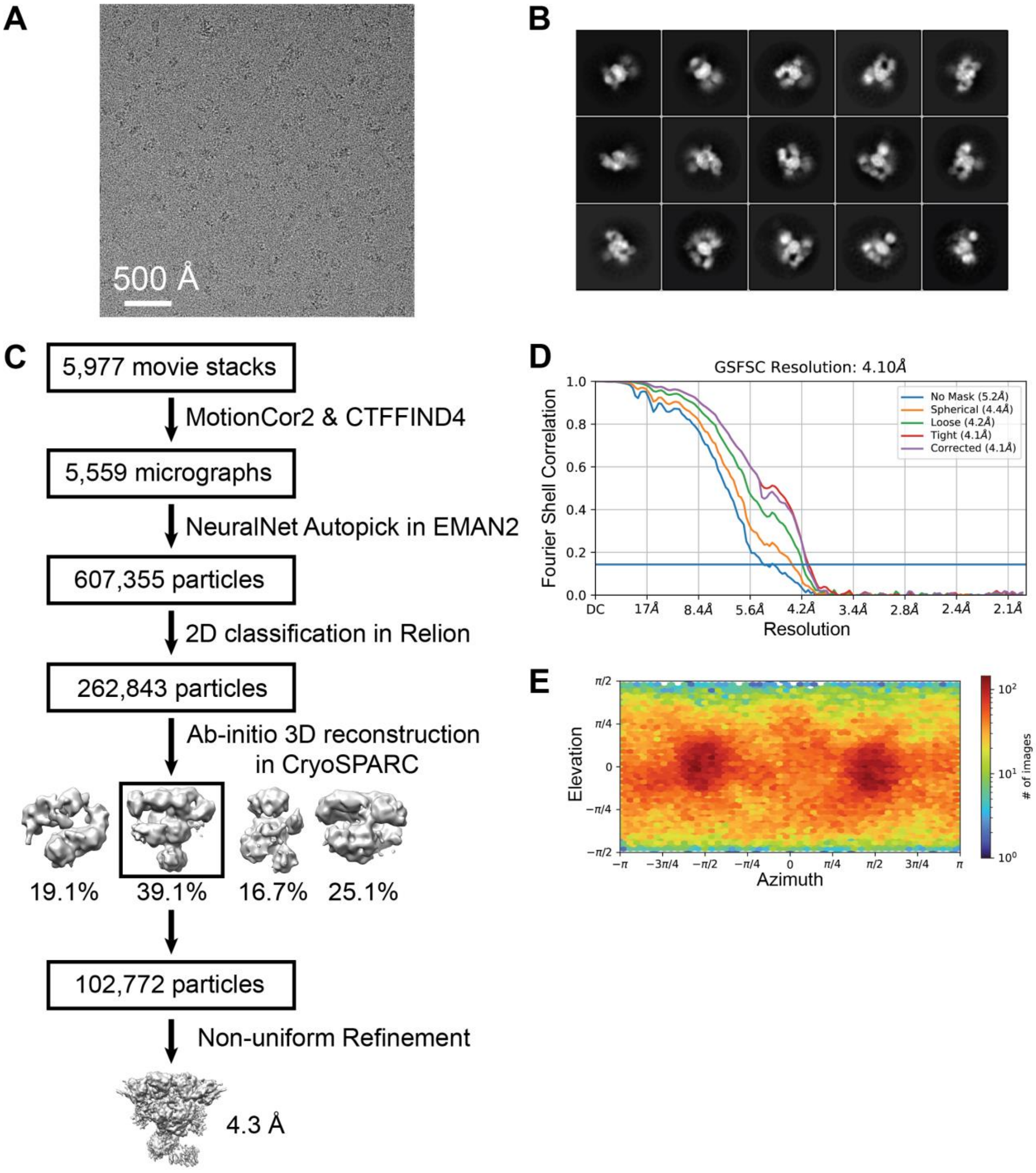


Fig. S21.

Single-particle cryo-EM analysis of F_{ab}-complexed diketide-M1 (pDC7) intermediate in the *turnstile-closed* state. (**A**) Representative motion-corrected cryo-EM micrograph. (**B**) Reference-free 2D class averages. (**C**) Data processing workflow (PDB 7M7J/EMD-23715)(*46*, *47*, *49*, *50*, *57*). (**D**) Gold standard FSC plot for the 3D reconstruction, calculated in cryoSPARC(*50*). (**E**) Euler angle distribution of the particle images.

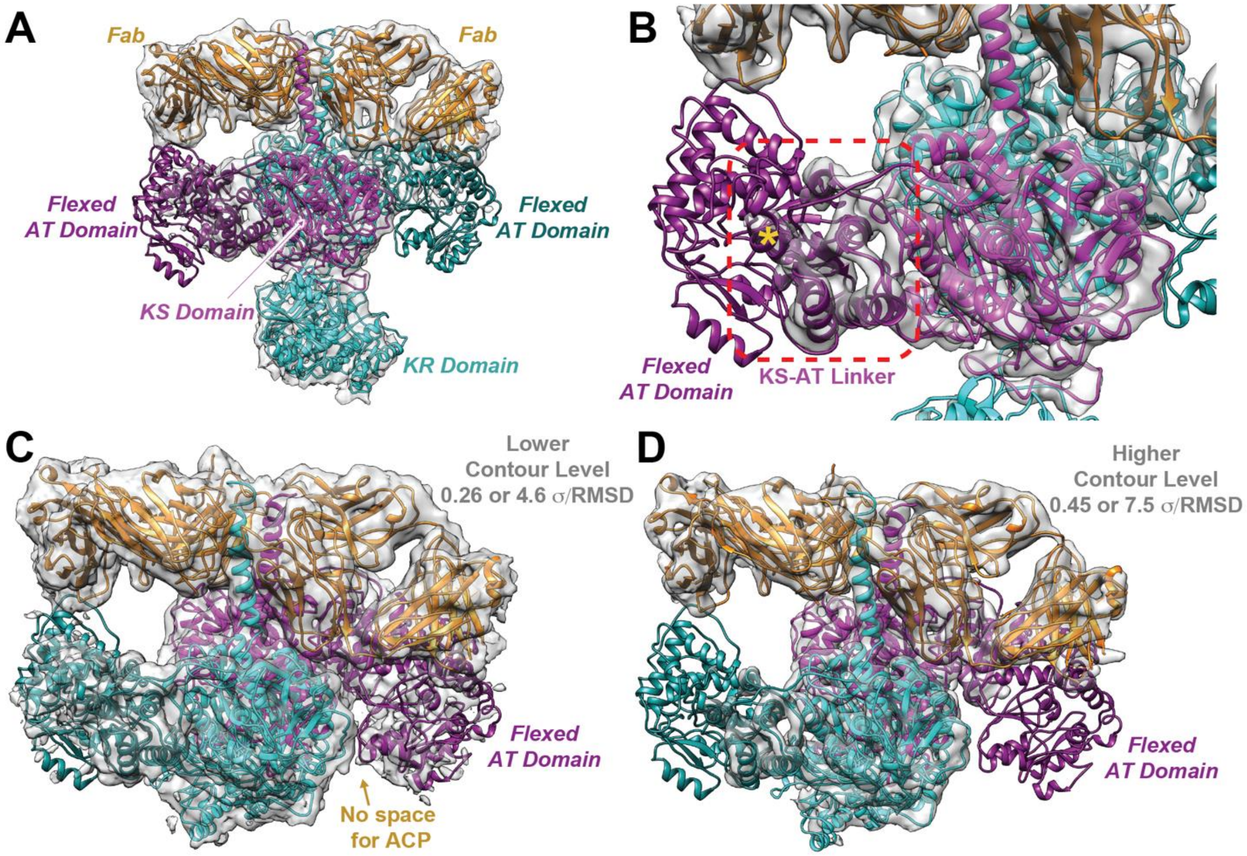


Fig. S22.

Summary and snapshots of Movie S2 featuring the 4.3 Å map of diketide-M1-F_{ab} (pDC7) at various threshold values. (**A**) Overview of the diketide-M1-F_{ab} in the *turnstile-closed* state (threshold = 0.45[7.5σ/RMSD]). (**B**) Close-up view of the KS-AT linker (boxed in dashed red) which could be unambiguously resolved, reinforcing its conformational rigidity during AT flexing. An α-helix (Leu865 – Ala878) is accented with a gold asterisk (as in fig. S17A) which is C-terminal to the AT domain but a constituent of the KS-AT linker subdomain. AT flexing is proposed to hinge on the two unstructured loops directly N/C-terminal to this helix (threshold = 0.45[7.5σ/RMSD]). (**C,D**) Another view of the *turnstile-closed* map contoured at 4.6 σ (C) or 7.5 σ (D). The KS-AT cleft constriction is better visualized at the lower contour level, owing to its poor local resolution (~7.5 Å; Fig. 4B). Only one of the two flexed ATs and KS-AT clefts is labeled, for clarity. Sigma values were calculated using the entire map volume and the equation: σ = (threshold – mean) / SD.

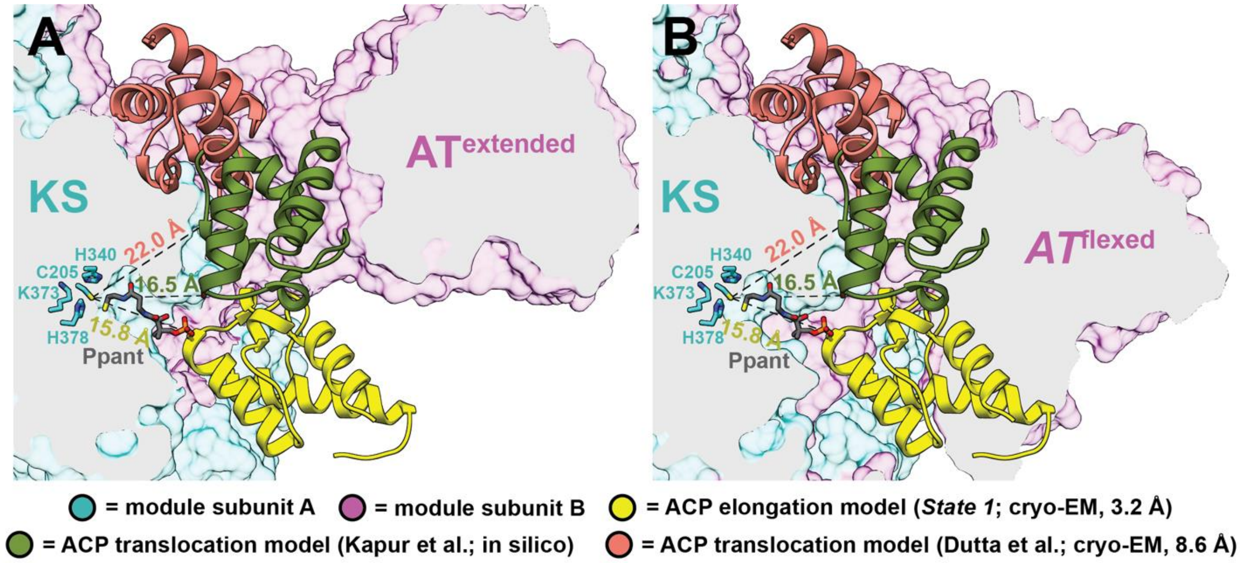


Fig. S23.

Comparison of the ACP binding modes for polyketide translocation and elongation and their compatibility with binding to an AT-flexed module (Ppant = 4′-phosphopantetheine). (**A**) ACPs corresponding to two previously reported translocation models(*6*, *29*) were compared to the experimentally observed ACP in the *State 1* structure (PDB 7M7F). The KS-AT cleft can accommodate all three ACPs, suggesting translocation and elongation are compatible with ACP binding to an extended AT cleft. Conversely, (**B**) superposing the same ACP coordinates onto a flexed AT cleft from the *turnstile-closed* structure results in incompatibility. In particular, ACP binding in one of the translocation models and the elongation model is precluded due to cleft constriction(*10*). While the translocation model proposed by Dutta et al. (EMD-5651) appears compatible with ACP binding to an AT-flexed cleft, the interatomic distance of 22.0 Å between the Ser-oxygen of the ACP and the Cys-sulfur of the KS is too long to accommodate an acyl-Ppant substrate in the KS active site. Furthermore, the direct path is blocked by KS residues surrounding the active site (made clear by the cross-sectional view in both panels). As a final note, all three ACPs could conceivably bind to a F_{ab}(1B2)-bound module, as no steric overlap was observed after superposition of these models (not shown).

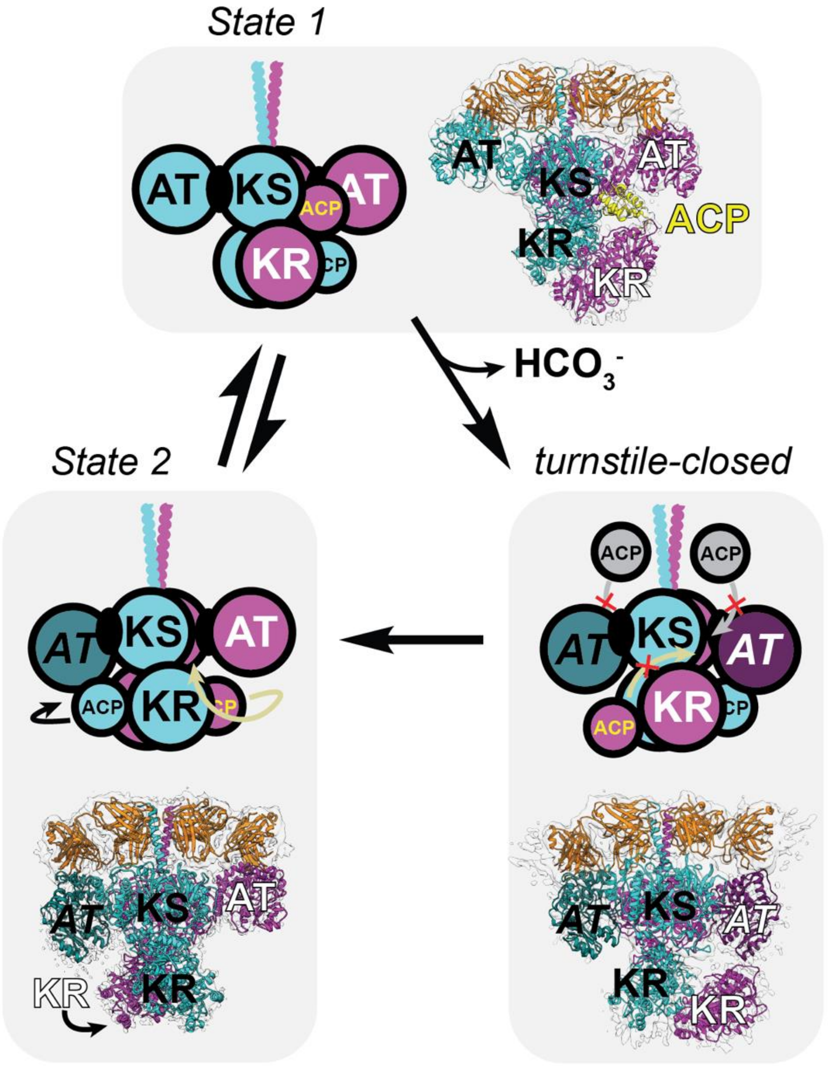


Fig. S24.
Dynamic model for an assembly-line PKS module. The observed structures (PDB 7M7H, PDB 7M7J, and PDB 7M7G; in clockwise order starting from the top) are drawn adjacent to the corresponding cartoon models from which the F_{ab}s have been removed for clarity. F_{ab} 1B2 = orange; module subunit A = cyan; module subunit B = magenta; hypothetical upstream ACP = gray; dark cyan/dark magenta = flexed ATs (subunit A/subunit B); black ovals = KS-AT linker subdomains. As in Figure 4A, the KR from subunit B has been modeled in the post-elongation state, although its true orientation was unresolved (PDB 7M7J).

Data Collection						
Sample	M3/1-F_{ab}	M1-F_{ab}			M1-F_{ab}	Diketide-M1-F_{ab}
Module expression plasmid	pTED23	pDC1			pDC7	pDC7
Microscope	Titan Krios G3i	Titan Krios G3i			Titan Krios G3i	Titan Krios G3i
Voltage (kV)	300	300			300	300
Camera	Falcon4	Falcon4			Falcon4	Falcon4
Pixel size (Å)	1.0	1.0			1.0	1.0
Total Dose (e-/Å^2)	50	50			50	50
Defocus range during data collection(µm)	-1.0 – -2.5	-1.0 – -2.5			-1.0 – -2.5	-1.0 – -2.5
Number of micrographs	3,974	10,234			8,104	5,977
Number of initial particles	381,647	1,406,538			466,815	607,355
Number of initial particles after 2D classification	205,348	619,899			169,074	262,843
Symmetry	C1	C1			C1	C1
Subclass	–	*State 1*	*State 2*	*State 1′*	–	–
Number of final particles	93,053	172,352	56,378	58,206	69,566	102,772
Resolution (0.143 FSC, Å)	3.2	3.2	4.1	4.1	3.4	4.3
EMDB ID	23710	23711	23712	23713	23714	23715
Atomic Model Refinement						
PDB ID	7M7E	7M7F	7M7G	7M7H	7M7I	7M7J
Clashscore (all atoms)	7.99	15.31	14.76	17.33	13.99	19.07
Poor rotamers (%)	9.50	15.86	18.86	5.15	2.38	5.04
Ramachandran outliers (%)	0.00	0.09	0.03	0.05	0.00	0.06
Ramachandran favored (%)	96.07	91.26	90.88	93.72	94.09	93.57
MolProbity score	2.45	3.11	3.17	2.70	2.40	2.73
Bond length (RMSD, Å)	0.006	0.012	0.012	0.005	0.004	0.005
Bond angles (RMSD, °)	1.088	1.721	1.673	0.920	0.761	0.926

Table S1.
Cryo-EM data collection, processing, and model validation statistics.

Name	Sequence (5′→3′)	PCR amplicon	PCR template
pDC1_M1_F	ATCCGCGAGCTGGAAGGCGAACCGGTCGCGGTCGTCGCG	i	pBL13
pDC1_M1_R	CGAGCTCACTAGTGAGGTGCGCCGCCAGCGTCCGCAC	i	pBL13
pDC1_TE_F	CTGGCGGCGCACCTCACTAGTGAGCTCGGCAGCGGGACTC	ii	pRSG34
pDC1_(3)_R	GCGACCGGTTCGCCTTCCAGCTCGCGGATGCGCTGCCG	ii	pRSG34
pDC7_(3)M1_F	ATGGCTAGCACTGACAGCGAGAAGGTGGC	i	pDC1
pDC7_(3)M1_R	CCCCGGACCTCGGTGCCGAGTTCGGCGGCGAGGTGCGCCGCCAGCGTCCG	i	pDC1
pDC7_(2)-(3)_F	CGAACTCGGCACCGAGGTCCGGGGGGAGG	ii	pBL36
pDC7_(2)-(3)_R	CACCTTCTCGCTGTCAGTGCTAGCCATATGT	ii	pBL36

Table S2.
Oligonucleotides used in this study. See the Methods for amplicon details and table S3 for template references.

Plasmid Name	Encoded Protein	Antibiotic Resistance	Reference
[†]pBL12	DEBS LDD(4)	Kan	(*14*)
pBL13	DEBS (5)M1(2)	Carb/Amp	(*14*)
pBL36	DEBS (3)M2(2)	Carb/Amp	(*14*)
pBL16	DEBS (3)M2TE	Carb/Amp	(*14*)
pRSG34	DEBS M3TE	Carb/Amp	(*12*)
pAYC02	DEBS KS3-AT3	Carb/Amp	(*13*)
PrpE-pET28	PrpE	Kan	(*10*)
*SCME-pET28	SCME	Kan	(*67*)
*MatB-pET28	MatB	Kan	(*67*, *68*)
n/a	F_{ab} 1B2	Carb/Amp	(*69*)
pTED23	DEBS M3/1[#]	Carb/Amp	(*21*)
pDC1	DEBS M1[#] (a.k.a. M1TE)	Carb/Amp	This study
pDC7	DEBS M1[‡]	Carb/Amp	This study

*A gift from Prof. Michelle Chang's lab (University of California, Berkeley)

[†] LDD(4) = LM(4)

[#]The encoded proteins contain a C-terminal TE domain and the N-terminal docking domain from DEBS M3 (see Protein Sequences above).

[‡]The encoded protein (lacking a TE domain) contains a C-terminal docking domain from DEBS M2 and the N-terminal docking domain from DEBS M3 (see Protein Sequences above).

Table S3.
Plasmids used in this study (Kan = kanamycin; Carb = carbenicillin; Amp = ampicillin).

References and Notes

1. D. J. Newman, G. M. Cragg, Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020). doi:10.1021/acs.jnatprod.9b01285 Medline

2. H. A. Kirst, The spinosyn family of insecticides: Realizing the potential of natural products research. *J. Antibiot. (Tokyo)* **63**, 101–111 (2010). doi:10.1038/ja.2010.5 Medline

3. T. Robbins, Y.-C. Liu, D. E. Cane, C. Khosla, Structure and mechanism of assembly line polyketide synthases. *Curr. Opin. Struct. Biol.* **41**, 10–18 (2016). doi:10.1016/j.sbi.2016.05.009 Medline

4. C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem.* **76**, 195–221 (2007). doi:10.1146/annurev.biochem.76.053105.093515 Medline

5. C. M. Kao, R. Pieper, D. E. Cane, C. Khosla, Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. *Biochemistry* **35**, 12363–12368 (1996). doi:10.1021/bi9616312 Medline

6. S. Dutta, J. R. Whicher, D. A. Hansen, W. A. Hale, J. A. Chemler, G. R. Congdon, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith, G. Skiniotis, Structure of a modular polyketide synthase. *Nature* **510**, 512–517 (2014). doi:10.1038/nature13423 Medline

7. J. R. Whicher, S. Dutta, D. A. Hansen, W. A. Hale, J. A. Chemler, A. M. Dosey, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith, G. Skiniotis, Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **510**, 560–564 (2014). doi:10.1038/nature13409 Medline

8. M. A. Skiba, A. P. Sikkema, N. A. Moss, A. N. Lowell, M. Su, R. M. Sturgis, L. Gerwick, W. H. Gerwick, D. H. Sherman, J. L. Smith, Biosynthesis of t-butyl in apratoxin A: Functional analysis and architecture of a PKS loading module. *ACS Chem. Biol.* **13**, 1640–1650 (2018). doi:10.1021/acschembio.8b00252 Medline

9. J. Wang, J. Liang, L. Chen, W. Zhang, L. Kong, C. Peng, C. Su, Y. Tang, Z. Deng, Z. Wang, Structural basis for the biosynthesis of lovastatin. *Nat. Commun.* **12**, 867 (2021). doi:10.1038/s41467-021-21174-8 Medline

10. B. Lowry, X. Li, T. Robbins, D. E. Cane, C. Khosla, A turnstile mechanism for the controlled growth of biosynthetic intermediates on assembly line polyketide synthases. *ACS Cent. Sci.* **2**, 14–20 (2016). doi:10.1021/acscentsci.5b00321 Medline

11. A. Chari, D. Haselbach, J.-M. Kirves, J. Ohmer, E. Paknia, N. Fischer, O. Ganichkin, V. Möller, J. J. Frye, G. Petzold, M. Jarvis, M. Tietzel, C. Grimm, J.-M. Peters, B. A. Schulman, K. Tittmann, J. Markl, U. Fischer, H. Stark, ProteoPlex: Stability optimization of macromolecular complexes by sparse-matrix screening of chemical space. *Nat. Methods* **12**, 859–865 (2015). doi:10.1038/nmeth.3493 Medline

12. R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284**, 482–485 (1999). doi:10.1126/science.284.5413.482 Medline

13. A. Y. Chen, N. A. Schnarr, C.-Y. Kim, D. E. Cane, C. Khosla, Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase. *J. Am. Chem. Soc.* **128**, 3067–3074 (2006). doi:10.1021/ja058093d Medline

14. B. Lowry, T. Robbins, C.-H. Weng, R. V. O'Brien, D. E. Cane, C. Khosla, In vitro reconstitution and analysis of the 6-deoxyerythronolide B synthase. *J. Am. Chem. Soc.* **135**, 16809–16812 (2013). doi:10.1021/ja409048k Medline

15. X. Li, N. Sevillano, F. La Greca, L. Deis, Y.-C. Liu, M. C. Deller, I. I. Mathews, T. Matsui, D. E. Cane, C. S. Craik, C. Khosla, Structure–function analysis of the extended conformation of a polyketide synthase module. *J. Am. Chem. Soc.* **140**, 6518–6521 (2018). doi:10.1021/jacs.8b02100 Medline

16. Y. Tang, A. Y. Chen, C.-Y. Kim, D. E. Cane, C. Khosla, Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem. Biol.* **14**, 931–943 (2007). doi:10.1016/j.chembiol.2007.07.012 Medline

17. A. L. Edwards, T. Matsui, T. M. Weiss, C. Khosla, Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J. Mol. Biol.* **426**, 2229–2245 (2014). doi:10.1016/j.jmb.2014.03.015 Medline

18. S. C. Tsai, L. J. Miercke, J. Krucinski, R. Gokhale, J. C. Chen, P. G. Foster, D. E. Cane, C. Khosla, R. M. Stroud, 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 (2001). doi:10.1073/pnas.011399198 Medline

19. L. Griffin, A. Lawson, Antibody fragments as tools in crystallography. *Clin. Exp. Immunol.* **165**, 285–291 (2011). doi:10.1111/j.1365-2249.2011.04427.x Medline

20. S. Wu, A. Avila-Sakar, J. Kim, D. S. Booth, C. H. Greenberg, A. Rossi, M. Liao, X. Li, A. Alian, S. L. Griner, N. Juge, Y. Yu, C. M. Mergel, J. Chaparro-Riggers, P. Strop, R. Tampé, R. H. Edwards, R. M. Stroud, C. S. Craik, Y. Cheng, Fabs enable single particle cryoEM studies of small proteins. *Structure* **20**, 582–592 (2012). doi:10.1016/j.str.2012.02.017 Medline

21. D. P. Cogan, X. Li, N. Sevillano, I. I. Mathews, T. Matsui, C. S. Craik, C. Khosla, Antibody probes of module 1 of the 6-deoxyerythronolide B synthase reveal an extended conformation during ketoreduction. *J. Am. Chem. Soc.* **142**, 14933–14939 (2020). doi:10.1021/jacs.0c05133 Medline

22. A. T. Keatinge-Clay, R. M. Stroud, The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure* **14**, 737–748 (2006). doi:10.1016/j.str.2006.01.009 Medline

23. S. R. Bagde, I. I. Mathews, J. C. Fromme, C.-Y. Kim, Modular polyketide synthase contains two reaction chambers that operate asynchronously. *Science* **374**, 723–729 (2021).

24. G. Pintilie, K. Zhang, Z. Su, S. Li, M. F. Schmid, W. Chiu, Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020). doi:10.1038/s41592-020-0731-1 Medline

25. H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in

macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016). doi:10.1093/nar/gkw408 Medline

26. M. Klaus, M. Grininger, Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **35**, 1070–1081 (2018). doi:10.1039/C8NP00030A Medline

27. S. Yuzawa, K. Deng, G. Wang, E. E. K. Baidoo, T. R. Northen, P. D. Adams, L. Katz, J. D. Keasling, Comprehensive in vitro analysis of acyltransferase domain exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synth. Biol.* **6**, 139–147 (2017). doi:10.1021/acssynbio.6b00176 Medline

28. A. Hagen, S. Poust, T. de Rond, J. L. Fortman, L. Katz, C. J. Petzold, J. D. Keasling, Engineering a polyketide synthase for in vitro production of adipic acid. *ACS Synth. Biol.* **5**, 21–27 (2016). doi:10.1021/acssynbio.5b00153 Medline

29. S. Kapur, A. Y. Chen, D. E. Cane, C. Khosla, Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22066–22071 (2010). doi:10.1073/pnas.1014081107 Medline

30. S. Kapur, B. Lowry, S. Yuzawa, S. Kenthirapalan, A. Y. Chen, D. E. Cane, C. Khosla, Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4110–4115 (2012). doi:10.1073/pnas.1118734109 Medline

31. L. G. Trabuco, E. Villa, K. Mitra, J. Frank, K. Schulten, Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673–683 (2008). doi:10.1016/j.str.2008.03.005 Medline

32. L. G. Trabuco, E. Villa, E. Schreiner, C. B. Harrison, K. Schulten, Molecular dynamics flexible fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* **49**, 174–180 (2009). doi:10.1016/j.ymeth.2009.04.005 Medline

33. S. A. Kautsar, K. Blin, S. Shaw, J. C. Navarro-Muñoz, B. R. Terlouw, J. J. J. van der Hooft, J. A. van Santen, V. Tracanna, H. G. Suarez Duran, V. Pascal Andreu, N. Selem-Mojica, M. Alanjary, S. L. Robinson, G. Lund, S. C. Epstein, A. C. Sisto, L. K. Charkoudian, J. Collemare, R. G. Linington, T. Weber, M. H. Medema, MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **48** (D1), D454–D458 (2020). Medline

34. K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47** (W1), W81–W87 (2019). doi:10.1093/nar/gkz310 Medline

35. P. Cimermancic, M. H. Medema, J. Claesen, K. Kurita, L. C. Wieland Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy, B. W. Birren, E. Takano, A. Sali, R. G. Linington, M. A. Fischbach, Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158**, 412–421 (2014). doi:10.1016/j.cell.2014.06.034 Medline

36. M. A. Skinnider, N. J. Merwin, C. W. Johnston, N. A. Magarvey, PRISM 3: Expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* **45** (W1), W49–W54 (2017). doi:10.1093/nar/gkx320 Medline

37. A. Nivina, K. P. Yuet, J. Hsu, C. Khosla, Evolution and diversity of assembly-line polyketide synthases. *Chem. Rev.* **119**, 12524–12547 (2019). doi:10.1021/acs.chemrev.9b00525 Medline

38. A. T. Keatinge-Clay, The uncommon enzymology of Cis-acyltransferase assembly lines. *Chem. Rev.* **117**, 5334–5366 (2017). doi:10.1021/acs.chemrev.6b00683 Medline

39. E. J. N. Helfrich, J. Piel, Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep.* **33**, 231–316 (2016). doi:10.1039/C5NP00125K Medline

40. R. W. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay, K. J. Weissman, The structure of docking domains in modular polyketide synthases. *Chem. Biol.* **10**, 723–731 (2003). doi:10.1016/S1074-5521(03)00156-X Medline

41. M. M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976). doi:10.1016/0003-2697(76)90527-3 Medline

42. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison 3rd, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009). doi:10.1038/nmeth.1318 Medline

43. D. G. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. Hutchison 3rd, H. O. Smith, Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008). doi:10.1126/science.1151721 Medline

44. T. Miyazawa, M. Hirsch, Z. Zhang, A. T. Keatinge-Clay, An in vitro platform for engineering and harnessing modular polyketide synthases. *Nat. Commun.* **11**, 80 (2020). doi:10.1038/s41467-019-13811-0 Medline

45. M. Vivoli, H. R. Novak, J. A. Littlechild, N. J. Harmer, Determination of protein-ligand interactions using differential scanning fluorimetry. *J. Vis. Exp.* **91**, 51809 (2014). doi:10.3791/51809 Medline

46. S. H. W. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012). doi:10.1016/j.jsb.2012.09.006 Medline

47. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline

48. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). doi:10.1016/j.jsb.2015.08.008 Medline

49. G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, S. J. Ludtke, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007). doi:10.1016/j.jsb.2006.05.009 Medline

50. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

doi:10.1038/nmeth.4169 Medline

51. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi:10.1002/jcc.20084 Medline
52. A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018). doi:10.1093/nar/gky427 Medline
53. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi:10.1107/S0907444910007493 Medline
54. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018). doi:10.1107/S2059798318006551 Medline
55. V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi:10.1107/S0907444909042073 Medline
56. H. Chen, L. Du, Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl. Microbiol. Biotechnol.* **100**, 541–557 (2016). doi:10.1007/s00253-015-7093-0 Medline
57. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018). doi:10.7554/eLife.42166 Medline
58. E. C. Meng, E. F. Pettersen, G. S. Couch, C. C. Huang, T. E. Ferrin, Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**, 339 (2006). doi:10.1186/1471-2105-7-339 Medline
59. H. Ashkenazy, E. Erez, E. Martz, T. Pupko, N. Ben-Tal, ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010). doi:10.1093/nar/gkq399 Medline
60. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004). doi:10.1101/gr.849004 Medline
61. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990). doi:10.1016/S0022-2836(05)80360-2 Medline
62. M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011). doi:10.1093/nar/gkr466 Medline
63. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, PDB2PQR: An automated

pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667 (2004). doi:10.1093/nar/gkh381 Medline

64. T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, N. A. Baker, PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522–W525 (2007). doi:10.1093/nar/gkm276 Medline
65. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037–10041 (2001). doi:10.1073/pnas.181342398 Medline
66. N. R. Voss, M. Gerstein, 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38**, W555–W562 (2010). doi:10.1093/nar/gkq395 Medline
67. A. J. Hughes, A. Keatinge-Clay, Enzymatic extender unit generation for in vitro polyketide synthase reactions: Structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* **18**, 165–176 (2011). doi:10.1016/j.chembiol.2010.12.014 Medline
68. M. C. Walker, B. W. Thuronyi, L. K. Charkoudian, B. Lowry, C. Khosla, M. C. Y. Chang, Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **341**, 1089–1094 (2013). doi:10.1126/science.1242345 Medline
69. J. Kim, R. M. Stroud, C. S. Craik, Rapid identification of recombinant Fabs that bind to membrane proteins. *Methods* **55**, 303–309 (2011). doi:10.1016/j.ymeth.2011.09.012 Medline