Coupled inter-subunit dynamics enable the fastest CO_2-fixation by reductive carboxylases

Hasan DeMirci[1,2,#], Yash Rao[1,4,#], Gabriele M. Stoffel[3,#], Bastian Vögeli[3,#], Kristina Schell[3,#], Alexander Batyuk[4], Cornelius Gati[1,5], Raymond G. Sierra[4], Mark S. Hunter[4], E. Han Dao[1,2], Halil I. Ciftci[2], Brandon Hayes[4], Fredric Poitevin[5], Kensuke Tono[6,7], David Adrian Saez[8], Esteban Vöhringer-Martinez[8], Samuel Deutsch[9], Yasuo Yoshikuni[9], Tobias J. Erb[3,10,*], Soichi Wakatsuki[1,6,*]

[1] Biosciences Division, SLAC National Accelerator Laboratory; [2] PULSE Institute, SLAC National Accelerator Laboratory; [3] Department of Biochemistry and Synthetic Metabolism, Max Planck Institute for terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany; [4] Linac Coherent Light Source, SLAC National Accelerator Laboratory; [5] Structural Biology Department, Stanford University; [6] RIKEN SPring-8 Center, Sayo, Hyogo, Japan; [7] Japan Synchrotron Radiation Research Institute, Sayo, Hyogo, Japan; [8]Departamento de Físico Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción, Chile; [9] U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Walnut Creek, CA, USA; [10] LOEWE Center for Synthetic Microbiology (SYNMIKRO); [#] these authors contributed equally; [*] corresponding authors

Abstract

Enoyl-CoA carboxylases/reductases (ECRs) are the most efficient CO_2-fixing enzymes described to date, outcompeting RubisCO, the key enzyme in photosynthesis in catalytic activity by more than an order of magnitude. However, the molecular mechanisms underlying ECR's extraordinary catalytic activity remain elusive. Here we used different crystallographic approaches, including ambient temperature X-ray Free Electron Laser (XFEL) experiments, to study the dynamic structural organization of the ECR from *Kitasatospora setae*. *K. setae* ECR is a homotetramer that differentiates into a dimer of dimers of open- and closed-form subunits in the catalytically active state, suggesting that the enzyme operates with "half-site reactivity" to achieve high catalytic rates. Using structure-based mutagenesis, we show that catalysis is synchronized in *K. setae* ECR across the pair of dimers by conformational coupling of catalytic domains and within individual dimers by shared substrate binding sites. Our results provide unprecedented insights into the dynamic organization and synchronized inter- and intra-subunit communications of nature's most efficient CO_2-fixing enzyme during catalysis.

INTRODUCTION

The capture and conversion of atmospheric CO_2 remains a challenging task for chemistry, resulting in an ever-increasing interest to understand and exploit CO_2 fixation mechanisms offered by biology[1]. The recently described family of enoyl-CoA carboxylases/reductases (ECRs) represent the most efficient CO_2-fixing enzymes found in nature to date[2,3]. ECRs catalyze the reductive carboxylation of a variety of enoyl-CoA thioester substrates at catalytic rates that are up to 20-fold higher than Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), an enzyme involved in the first carbon fixation step in the Calvin-Benson cycle of photosynthesis[1,4].

ECRs catalyze the reduction of α,β-unsaturated enoyl-CoAs using the reduced form of the cofactor nicotinamide adenine dinucleotide phosphate (NADPH). This generates a reactive enolate species, which acts as a nucleophile to attack a CO_2 molecule[2,3,5]. The structural details of the carboxylation reaction have remained elusive, due in part to the lack of high-resolution structures of ECRs containing catalytic intermediates and carboxylated products. Currently, there are five available ECR structures. However,

they all have different substrate specificities, ranging from short- (PDB: 3HZZ, 3KRT) to long-chain (4A0S[6]) and aromatic enoyl-CoA substrates (4Y0K[7]), and are from different biological backgrounds including primary (i.e. central carbon) metabolism (PDB: 4GI2) and secondary metabolism. Moreover, most of them were co-crystalized with NADPH or NADP+ only and do not contain CO_2, enoyl-CoA substrates or acyl-CoA products. This significantly limits our structural understanding of the enzyme's catalytic mechanism.

The aim of this study was to provide a detailed structural understanding of the carboxylation reaction of ECRs at the level of the oligomeric protein complex. To this end, we chose the ECR from *K. setae*, which shows high substrate specificity for crotonyl-CoA and superior catalytic efficiency (see Table 1). Using cryogenic X-ray crystallography at synchrotrons and room temperature serial femtosecond X-ray crystallography (SFX) at an XFEL, four high-resolution ECR structures were determined in different conformational states: the apo form and three holo forms, in binary complex with the reduced cofactor NADPH, in ternary complex with NADPH and butyryl-CoA, and in binary complex with the oxidized cofactor NADP+ (**Figure 1a**).

Here we show that the tetrameric complex assumes a dimer-of-dimers ("a pair of dimers") configuration during catalysis. The central oligomerization domains of ECR remain largely unchanged, while the peripheral catalytic domains move drastically to provide two sets of active site conformations, open- and closed-form, upon binding of the NADPH cofactor alone or in the presence of substrates. This coordinated motion is enabled by a tight coupling of catalytic domains across the pairs of dimers. Structure based mutagenesis of the interface of the catalytic domains supports this notion and provides compelling evidence that synchronization across the pair of dimers is a crucial factor in *K. setae* ECR to achieve the high catalytic efficiency. Further kinetic experiments demonstrate that subunit communication within the pair of dimers is important to synchronize open- and closed-states. Altogether, our data unveil a detailed picture of the dynamic structural organization and subunit synchronization of the ECR complexes, providing unprecedented insights into the functional organization of nature's most efficient CO_2-fixing enzyme during catalysis.

RESULTS

Apo ECR is a symmetric homotetramer, readily accessible for NADPH binding

We first determined the apo form of the ECR crystal structure from *K. setae* at 1.8 Å resolution by using synchrotron X-ray crystallography at cryogenic temperature (**Supplementary Table 1**). The asymmetric unit contains one homotetramer composed of four subunits arranged in a dimer of dimers geometry ("pair of dimers") similar to those of the previously reported binary (PDB: 4Y0K) and ternary (PDB: 4A0S) ECR structures. Overall, the tetramer shows a non-crystallographic, close to D2 (dihedral) symmetry (**Supplementary Figure S1, top right panel**) with four conformationally identical subunits (**Supplementary Figures S1&S2,** RMSD = 0.1 Å). The tetrameric structure of *K. setae* ECR is further supported by size-exclusion chromatography which showed that the apo enzyme eluted as a single peak at 205 kDa compared to the expected monomer molecular weight of 51.2 KDa corresponding to a functional complex of four subunits (**Supplementary Figure S3**).

Each ECR subunit consists of two domains – a larger catalytic domain formed by residues 1-212 and 364-445, and a smaller oligomerization domain formed by residues 212 to 363 (**Supplementary Figure S4**). The oligomerization domain comprises a Rossmann fold[8] with repeating αβ-motifs that forms a 6-stranded β-sheet (β12 to β17). The 6-stranded β-sheets of two neighboring subunits are combined into one 12-stranded β-sheet, forming the core of one dimer, A/C or B/D. Two of these 12-stranded β-sheets then form the core of the tetrameric complex (**Supplementary Figure S4**).

The catalytic domains of *K. setae* ECR are located at the periphery of the tetrameric complex. The active site of ECR is formed by helix 8 and surrounding loops at the interface with an adjacent subunit in the tetramer (**Supplementary Figure S4**). The active site cavities in the apo form are open and accessible for both the cofactors and substrates.

NADPH binding induces ECR into a dimer of dimers with distinct open and closed form subunits

To understand how cofactor binding affects the enzyme, we determined the crystal structure of the *K. setae* ECR-NADPH binary complex at 2.4 Å resolution by using serial femtosecond X-ray crystallography (SFX) at ambient temperature (**Figure 2, Supplementary Table 1**)[9–12]. The simple F_o-F_c difference electron density map allowed us to unambiguously place NADPH molecules in all four subunits. NADPH binds with its adenine moiety in the oligomerization domain and spans the catalytic domain, where its nicotinamide moiety is located (**Supplementary Figure S5**).

Notably, binding of NADPH breaks the dihedral D2 symmetry observed in the apo-form tetramer structure, while symmetry about the y-axis is retained, resulting in a non-crystallographic, almost cyclic C2 symmetry (**Supplementary Figure S1 bottom right**). In the NADPH-ECR binary complex, the four subunits of ECR differentiate into two forms (A & B and C & D), which are structurally distinct from each other (**Supplementary Figure S1**, RMSD = 0.5 Å between A & B and C & D, 1.8 Å between A & C, A & D, B & C and B & D respectively) (**Figure 2a&b**). The A & B subunits show cofactor-binding pockets that are open, referred hereafter as "open-form" state (**Figure 2b**). On the other hand, in the C & D subunits, the cofactor binding pocket is compressed inwards, which seals the NADPH cofactor within the catalytic domain, resulting in a "closed-form" state (**Figure 2b**).

The bulk of the NADPH cofactor is bound almost identically in the two closed-form subunits, C & D (**Figure 2d,e&f**). However, the nicotinamide moiety adopts two alternate conformations in the two open-form subunits A & B (**Figure 2e&f**), indicating a more flexible cofactor binding than in the closed-form subunits (**Figure 2d**). Possible conformations of the NADPH cofactor in the open and closed binding cavities and its flexibility were studied with molecular dynamics (MD) simulations in a dimer of subunits A and C. In the closed-form subunit C the cofactor kept its position in the binding site as observed in the X-ray crystal structure. When we placed the NADPH cofactor in the same position in the open subunit A and performed similar MD simulations, NADPH left this initial confirmation in all three 200 ns trajectories and adopted various alternate conformations in the open cavity, including the two that were observed in our high-resolution crystal structure (**Supplementary Figure S6**). These variable conformations in the open A subunit are allowed by a substrate binding pocket that is more than 5 Å wider than the closed C subunit (**Figure 2c**).

In summary, binding of the NADPH cofactors to the apo enzyme induces the four subunits of the enzyme to differentiate into open- and closed-form states in both dimers (**Figure 2b**) thus breaking the dihedral D2 symmetry to cyclic C2 symmetry. This coupled subunit rearrangement of *K. setae* ECR and the large active site differences within each pair of dimers suggest that catalysis is synchronized between the individual subunits of the complex, which will become clearer in the subsequent analysis sections below.

Ternary complex supports half-site reactivity in ECR catalysis

We next attempted to determine the structure of the *K. setae* ECR ternary complex crystallized in the presence of spent cofactor $NADP^+$ and the reaction product ethylmalonyl-CoA. The structure of the ternary complex, however, indicated that the carboxylate group was lost during the crystallization process, which resulted in butyryl-CoA, which is in line with the finding that ethylmalonyl-CoA is unstable and tends to decarboxylate at the active site of ECR into butyryl-CoA and CO_2 over time[2,13](**Supplementary Figure S7**). Numerous attempts of preserving ethylmalonyl-CoA in the crystal structure proved to be

extremely challenging and therefore we co-crystallized ECR with butyryl-CoA and NADPH and determined its structure at 1.7 Å resolution (**Figure 3**). This structure revealed that two butyryl-CoA molecules are bound at the active sites of the closed-form subunits B & D.

This ternary complex structure is overall very similar to the structure of ECR-NADPH binary complex. It also displays the non-crystallographic, pseudo C2 cyclic symmetry (**Figure 3a** and **Supplementary Figure S1 bottom right panel**) and comprises of open- and closed-form subunits that overlay very well with the open- and closed-form subunits of the ECR-NADPH binary complex (**Supplementary Figures S1&S2**, RMSD = 0.1, 2.1 Å respectively). The NADPH cofactor appears bound to all active sites, however, only the closed-form subunits B & D also contain the completely intact butyryl-CoA thioester (**Figure 3a&b**). This strongly suggests that the closed-form subunits represent the Michaelis complex in which substrate and cofactor are positioned for catalysis, while the open-form subunits represent catalytically incompetent complexes that are in place to perform the next round of catalysis.

ECR uses an elegant mechanism to align CoA-ester for catalysis in the closed-form subunit pairs. The active site of the closed-form subunits is sealed by the collective motion of loops 37-44, 88-94, 338-350, and helices 6, 7 and 21 of the catalytic domain (**Figure 3c**), which creates multiple interactions of the protein with the CoA-ester (**Figure 3d,e&f**). Notably, the CoA-ester extends from the catalytic closed-form domain into the neighboring open-form subunit within the same dimer pair, where Arg352, and Tyr353 interact with the phosphate backbone of CoA. The CoA moiety stretches further into the neighboring open-form subunit, where its adenosine tail interacts with a small binding pocket formed by three residues on the surface of the dimerization domain, Tyr328, Lys296, and Arg303 (**Figure 3g**).

When we inspected the adenine binding pocket of the closed-form subunits, we also observed electron density of adenine, indicating that the CoA-ester was bound (**Figure 3g right inset**). The electron density beyond the adenine ring, however, becomes disordered, suggesting that the part of the CoA molecule that extends into the active site of the neighboring open-form subunit remains flexible, which is corroborated by the higher anisotropy of the CoA binding site (**Figure 1c**). Quantum mechanical/molecular mechanics (QM/MM) simulations on a dimer of subunits A and C were performed to further evaluate the flexibility of the substrate in the open- and closed-form subunits. These simulations showed that the substrate residing in the closed subunit had significantly lower B-factors than the simulated substrate in the open subunit (**Supplementary Videos 1a-e, 2a-e**). In the open subunit, the acyl moiety was found to have a high degree of flexibility within the active site (**Supplementary Figure S8, Supplementary Videos 1&2**). Taken together, both the crystallographic analyses and QM/MM simulations are agreeing with the idea that the closed-, but not the open-form subunits represent the catalytically competent subunits. Furthermore, the organization of the ternary *K. setae* ECR complex into catalytically competent and incompetent subunits, suggests that the enzyme operates with half-site reactivity, in which active sites alternate during catalysis[14–16].

Product release returns ECR back into a symmetric homotetramer

Following catalysis, *K. setae* ECR has to release the product and oxidized cofactor. In order to understand the structural basis for this part of the catalytic cycle, the protein was co-crystallized with $NADP^+$ and its structure was solved at 1.8 Å resolution. The enzyme transitioned back to the D2 symmetry of the apo enzyme, with four conformationally identical subunits (**Supplementary Figures S1&S2**) all containing bound $NADP^+$ molecule (**Figure 4a**). Compared to the homotetrameric apo enzyme in which helix α6 and loop 88-93 of the catalytic domain stabilize the phosphate backbone of the $NADP^+$ molecule(**Figure 4b**), the corresponding helix and loop moved closer to each other (**Figure 4a**), similar to the ternary complex (**Figure 3d**), which leaves the homotetramer in an "all-closed" state. Notably, while the $NADP^+$ binding mode is comparable between the ternary complex and the complex with $NADP^+$ alone, the B-factors of

the latter are larger than those in the ternary complexes (**Figure 1b**), indicating that in the $NADP^+$ bound enzyme the atoms become in general more mobile, which may be advantageous for discharging the oxidized, spent cofactor. The configuration of the nicotinamide group of $NADP^+$ in the $NADP^+$ bound structure is similar to those of the NADPH in the cofactor-bound, closed form subunits (**Figure 4c**).

Swing motion of the peripheral catalytic domain during catalysis

Comparison of the high-resolution structures of apo, NADPH-bound, NADPH/butyryl-CoA-bound and $NADP^+$-bound *K. setae* tetrameric ECRs suggests that there are coordinated motions of the catalytic domains which are peripheral to the more rigid oligomerization domains (**Supplementary videos 3 & 4**). The apo-form and $NADP^+$-bound form show 4 equivalent subunits, while NADPH-bound and NADPH/butyryl-CoA-bound forms divide into two groups of open- (*i.e.*, catalytically incompetent) and closed-form (*i.e.*, catalytically competent) subunit pairs (**Figure 1a**). In order to understand how these global structural changes affect catalysis, a principal component analysis (PCA) was used to extract major structural differences among the four structures.

The PCA revealed 8 major contributions based on their singular values (**Supplementary Figure S9a**) and the movement of the catalytic domains described above was the strongest, followed by other less significant structural changes. The 8 PCA components were used to analyze the contributions of PC1 to PC8 to the structural changes between each of the 4 structures and the average tetramer structure. This analysis showed that the first three PCA components, PC1-3 (**Supplementary Videos 5a,b,&c**), can explain more than 50% of the structural changes (**Supplementary Videos 5d**). PC1 shows that the peripheral catalytic domains are coupled and swing up and down on either side of the central oligomerization domains; PC2 shows that each of the catalytic domains moves away from its partner catalytic domain, and finally the catalytic domains undergo subtle tilt motions in PC3. The deviations of the NADPH-bound and NADPH/butyryl-CoA-bound structures from the average structure is explained mainly by PC1, the $NADP^+$-bound form by PC2, and the apo form by PC3 (**Supplementary Figure S9b).**

Communication between pairs of dimers promotes catalysis

Given the coordinated motions of the peripheral catalytic domains during the catalysis, how is catalysis synchronized across the enzyme complex? One intriguing aspect of the ECR tetramer structures is that the catalytic domains share a common interface of 1636 $Å^2$ between the pairs of dimers (between the catalytic domains of A and D, and those of B and C, **Figure 5a&b**) suggesting that they move together as rigid bodies. A comparison of the overall domain movements between the apo-form and NADPH/product-bound ternary complex shows that the enzyme tetramer changes from the homotetrameric apo state to the open and closed-form subunit dimer pairs (**Figure 5a**). Upon binding of CoA and NADPH, neighboring catalytic domains rotate, which couples the widening of one active site with the compression of the other active site across the pair of dimers (**Figure 5a**). Thus, it seems to be a direct consequence of the rigid structure of the inter-catalytic domain interface that the enzyme will adopt two distinct conformational states in each dimer when it becomes catalytically active.

What are the molecular determinants that synchronize catalysis across the pair of open/closed form dimers? The inter-catalytic domain interface is mostly hydrophobic, but also features some electrostatic interactions (**Figure 5b**). Most notable are Asn218 of one subunit of one dimer that forms a hydrogen bond to Asn157 of the adjacent subunit of the other dimer, as well as Glu151 of one subunit of one dimer that forms hydrogen bonds to the main chain nitrogen of Asn133 (and/or Ala134) of the neighboring subunit of the other dimer (**Figure 5b**). Multiple sequence alignment showed that Glu151, Ans218 and Asn157 are highly conserved in ECRs from primary (i.e., central carbon) metabolism, which show faster CO_2-fixation kinetics (average k_{cat} 28 s^{-1}), but not in ECRs from secondary metabolism (**Figure 5c**), raising an interesting question about their roles in catalysis (average k_{cat} 1.2 s^{-1}).

Mutation of these residues, that are more than 20 Å away from the active site, dramatically affected the kinetic parameters of *K. setae* ECR (**Table 1**). In the E151D variant, the k_{cat} value was fivefold decreased, demonstrating that weakening the interaction of catalytic domains has profound effects on the catalytic rate of the enzyme. Mutations that targeted the asparagine interaction network showed also strong effects on the catalytic rate, but did additionally affect K_M of substrate binding. Most notable were variants N218E single and E151D/N157E/N218E triple variants that decreased the k_{cat} by more than 25- and 100-fold, respectively, highlighting that communication at the interface of the catalytic domains of the pair of dimers is an important determinant of the catalytic rate in *K. setae* ECR.

To exclude that the overall structure of the complex was not altered through these mutations, we used gel filtration, as well as native gel analysis to analyze the oligomerization state of the different enzyme variants (**Supplementary Figure S3a**). Gel filtration assays were performed under the same conditions as our kinetic measurements and showed that all mutant enzyme variants kept their tetrameric form. Only native gel analysis, which was performed under more disruptive conditions, showed slightly increase in the dimer and monomer fractions, indicating that interface interactions are weakened in these variants (**Supplementary S3b**). Overall, our mutational and kinetic data supports the hypothesis that synchronization of catalytic domains strongly contributes to catalytic rate and is conferred through hydrogen-bond network at the interface of the pair of dimers.

Shared substrate binding within dimers is important for catalysis

While our study on the interface between catalytic domains explained how communication is conferred between different dimers through the strong coupling of the two catalytic domains, each from two dimers (inter-dimer interaction, **Figure 5a**), it did not explain how catalysis is synchronized between the open- and closed-form subunits within the same dimer. We turned our attention back to the fact that CoA substrate binding is shared between the open- and closed-form subunits in each dimer through the adenine binding pocket (intra-dimer interaction, **Figure 5a**).

To understand the role of substrate adenine binding in catalysis, we characterized the kinetics of different single, double and triple mutant variants of the adenine binding site (**Figure 3g** and **Table 2**). Mutations in the adenine binding pocket, and in particular of Arg303, strongly increased the K_M of the CoA substrate as expected, but also decreased the apparent k_{cat} of the enzyme by a factor of two to three. Notably, a comparable decrease in k_{cat} was also observed in the wild-type enzyme when we used crotonyl-pantetheine, a truncated substrate that lacks the adenosine moiety and cannot bind to the adenine binding pocket. This indicated that shared cofactor binding between neighboring subunit is important for efficient catalysis, but did not provide a conclusive answer, how catalysis is synchronized across the subunits.

We noticed that the substrate adenine binding pocket is directly followed by a loop that carries a lysine residue (Lys332), which interacts with the active site of the neighboring subunit. Lys332 residue from the open-form subunit engages in a hydrogen bonding network with the nicotine amide group of the NADPH cofactor bound to the closed-form subunit through Gln165 and His365 of the neighboring subunit (**Figure 5d**). These interactions are not observed in the active site of the open-form subunit (**Figure 5d**), raising the question whether the hydrogen bonding network connected to the adenine binding pocket might be important for catalysis. In K332A and Q165A variants, k_{cat} was decreased two- to three-fold (**Table 2**). When we tested these variants with crotonyl-pantetheine, we saw much to our surprise that catalytic activity in the K332A variant was reduced by more than two orders of magnitude, leaving us with the suggestion that adenine binding together with the loop carrying Lys332 are important to synchronize catalysis between the two subunits within the dimer. Together with the inter-domain coupling, this intra-dimer synchronization drive fast CO_2-fixation by *K. setae* ECR.

CONCLUSIONS

Our structural studies of *K. setae* ECR revealed unprecedented details on the functional organization of nature's most efficient and fastest CO_2-fixing enzyme. During catalysis, the enzyme complex differentiates into distinct functional subunits. Binding of NADPH cofactor and substrates forces the homotetrameric apo enzyme into a dimer of dimers in which each dimer is constituted of an open- and a closed-form subunits. In the closed-form subunits the NADPH cofactor and CoA substrate are aligned with each other, suggesting that this is the catalytically competent state. The open-form subunits bind cofactor and the adenine rings of the substrates but the rest of the acyl-CoA substrate remains flexible and invisible in the active site. Thus, the open-subunit active sites seem to represent a catalytically incompetent state that is pre-organized for a next round of catalysis. Altogether, this structural reorganization of ECR strongly supports the idea that the enzyme operates with "half-site reactivity", according to which catalysis is synchronized across the enzyme tetramer and alters between the open- and closed-form subunits to increase the overall catalytic efficiency of the complex[14–17].

Interaction of the catalytic domains of neighboring subunits is crucial for efficient catalysis in *K. setae* ECR. Especially important is the interaction of catalytic domains between the pairs of dimers. As soon as this interaction is disturbed, the catalytic rate of the enzyme is severely diminished. This observation is consistent with theoretical and experimental data on half-site reactivity. Synchronization of the distant catalytic subunits can enhance the catalytic rate of enzymes several-fold[18,19]. Mutation of a single amino acid coupling the two catalytic sites of heptose isomerase GmhA reduced catalytic rate of GmhA to 6% of wild-type activity[20]. *Escherichia coli* thymidylate synthase is another example for an enzyme showing half-site reactivity[21]. Disturbing the interaction network in *E. coli* thymidylate synthase leads to a 400-fold decrease in k_{cat}[22,23], demonstrating that domain interactions are important factors in promoting enzyme catalysis[24].

Besides the inter-dimer domain interaction, our study on ECR also suggests joint substrate-binding between neighboring subunits as another potentially important mechanism of fast synchronized catalysis. The binding of the adenine end of the CoA ester into a pocket in the neighboring subunit seems to be connected back via a hydrogen-bonding network to the active site of the subunit where the CoA ester originated. This provides the missing link of how catalysis might be synchronized between the open- and closed-form subunits within one dimer. Taken together an attractive model of continuous turnover scheme emerges explaining the overall fast catalytic cycle of ECR; two consecutive reaction cycles alternate aided by the coupled inter-dimer catalytic domain motions (**Figure 5e**). In the first cycle (right half of **Figure 5e**), the open-form subunits A and B receive two sets of substrate and cofactor molecules, while the closed subunits C and D finish the previous reaction cycle, and release the products and $NADP^+$. As a result, the subunits A and B become closed and the subunits C and D switch to the open-subunit state. In the second cycle (left half of **Figure 5e**), subunits A and B perform the reaction and release the products and $NADP^+$ becoming open subunits, and the C and D subunits switch to closed state by acquiring a new set of substrate and NADPH.

While structural and biochemical data indicate that *K. setae* ECR achieves high catalytic rates by synchronizing active sites, this might not necessarily be true for other ECRs. A differentiation into dimers of dimers was not observed in NADPH-bound or ternary structures of other ECRs so far (e.g., PDB: 4Y0K and 4A0S respectively, which share substantial amino acid identity) (**Supplementary Figure S10**). Another reason might be that not all ECRs might perform synchronized catalysis. Note that ECRs fall into two different classes. Primary ECRs that operate in central carbon metabolism and secondary ECRs that serve in secondary metabolism, where they provide extender units for the synthesis of polyketides. Whereas primary ECRs are under strong evolutionary pressure and show on average k_{cat} values of 28 s^{-1} [25], secondary ECRs are not selected for high catalytic rates, which is also reflected by the fact that they show

an average k_{cat} value of 1.2 s^{-1} [25]. Thus, it might be tempting to speculate that secondary ECRs are not selected for high turnover rates during catalysis and thus might not display synchronized "half-site reactivity".

In summary, this work provides the first overall picture of the organization of the ECR homotetrameric complex. The observation of the differentiation of the apo tetramer into open and closed form subunits upon binding of NADPH seemed to have been made possible by room temperature data collection using the XFEL beam at SACLA, highlighting the power of ambient temperature crystallography to study larger scale motions in macromolecular crystals[26,27]. Further experiments using time-resolved X-ray crystallography at room temperature and mixing jets will be helpful to obtain a fully dynamic picture of the ECR complex during catalysis, which will be important to fill the gaps in the mechanistic understanding of nature's most efficient CO_2-fixing principle.

Data Availability

Coordinates of the four ECR structures have been deposited in the Protein Data Bank under accession codes, 6NA3 (apo), 6NA4 (Butryl-CoA/NADPH bound), 6NA5 ($NADP^+$ bound), and 6NA6 (NADPH-bound).

Acknowledgements

Authors acknowledge Takanori Nakane from University of Tokyo for his help with calibration of XFEL data from SACLA, RIKEN, Japan. We would like to thank Eriko Nango, Rie Tanaka and RIKEN SPring-8 Center for their help with data collection at SACLA, RIKEN, Japan. HD acknowledges support from NSF Science and Technology Center grant NSF-1231306 (Biology with X-ray Lasers, BioXFEL). The XFEL experiments were performed at BL3 of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2017A8055). The authors thank the beamline staff of Structural Molecular Biology Group, SSRL, SLAC and GM/CA CAT, Advance Photon Source, ANL for assistance on data collection. YR, RGS, MSH, and BH were supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. TJE and GS received support from the Max Planck Society, the European Research Council (ERC 637675 'SYBORG'), and the U.S. Department of Energy Joint Genome Institute a DOE Office of Science User Facility under Contract No. DE-AC02-05CH11231. DAS and EVM thank the Max-Planck Society for funding as a Max-Planck-Partner group. HD, and SW are supported by DOE Office of Science, Biological Environmental Research, and National Institute of Health, NIGMS. CG and SW were supported by National Science Foundation, Major Research Instrument grant. HD and SW's work was partially supported by Stanford PRECOURT Institute. Gregory M. Stewart of SLAC and Moe Wakatsuki for graphics work.

Table 1. Steady state analysis of *K. setae* ECR and variants targeting the catalytic domain interface between the pair of dimers. (Michaelis-Menten curves of *K. setae* ECR and its variants are provided in **Supplementary Figure S11**)

	Crotonyl-CoA			NADPH		CO_2	
Enzyme	K_M (µM)	K_i (µM)	k_{cat} (s^{-1})	K_M (µM)	k_{cat} (s^{-1})	K_M (µM)	k_{cat} (s^{-1})
Wild-type (WT)	21 ± 2	3650 ± 810	103 ± 3	37 ± 4	86 ± 2	90 ± 10	78 ± 2
E151D	28 ± 2	1958 ± 251	20 ± 1	72 ± 11	17 ± 1	80 ± 10	21 ± 1
N157E	515 ± 75	-	22 ± 1	105 ± 27	17.1 ± 0.4	40 ± 6	14 ± 0.2
N218E	272 ± 37	-	3.7 ± 0.2	66 ± 9	1.49 ± 0.06	630 ± 70	5.6 ± 0.2
E151DN157EN218E	245 ± 25	-	1.11 ± 0.04	26 ± 3	0.70 ± 0.02	440 ± 40	0.95 ± 0.03

Table 2. Apparent Michaelis-Menten parameters of *K. setae* ECR and variants targeting the adenine binding pocket and hydrogen binding network within the dimer as mean values ± standard error. (Michaelis-Menten curves of *K. setae* ECR and its variants are provided in **Supplementary Figure S11**)

	Crotonyl-CoA			Crotonyl-Pantetheine	
Enzyme	K_M (µM)	K_i (µM)	k_{cat} (s^{-1})	K_M (µM)	k_{cat} (s^{-1})
Wild-type (WT)	21 ± 2	3650 ± 810	103 ± 3	8658 ± 531	37 ± 1
K296A	107 ± 11	-	68 ± 2	-	-
Y328F	11 ± 2	4671 ± 1693	80 ± 3	-	-
R303K	702 ± 64	-	87 ± 3	-	-
R303A	516 ± 55	-	69 ± 3	2558 ± 769	16 ± 1
R303V	334 ± 33	-	31 ± 1	6930 ± 116	16 ± 1
K296A/Y328F	192 ± 34	-	39 ± 2	-	-
K296A/R303A/Y328F	2176 ± 280	-	29 ± 2	7772 ± 106	42 ± 3
K296A/R303K/Y328F	832 ± 138	-	53 ± 2	-	-
K332A	451± 130	3507 ± 1809	39 ± 7	2983 ± 509	0.33 ± 0.02
Q165A	27 ± 4	1873 ± 437	56 ± 3	5632 ± 521	8.5 ± 0.3

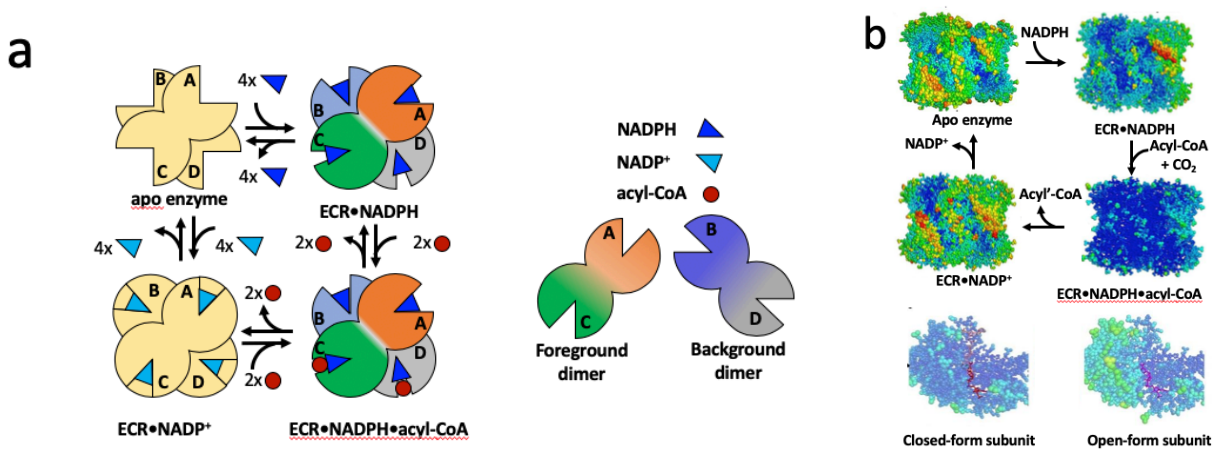


*Figure 1: Structural organization of the Kitasatospora setae ECR complex. **a**. Oligomeric organization of K. setae ECR, as inferred from the crystal structures solved in this study. The homotetrameric apo enzyme differentiates into a dimer of dimers of open (circles with 45° wedges) and closed (full circles) subunits upon binding of NADPH cofactor. The enzyme remains a dimer of dimers in the butyryl-CoA/NADPH-ternary complex and returns back into a homotetrameric state after product release. **b**. Anisotropic B-factors of the tetramer of the different ECR complexes solved in this study. **c.** Anisotropic B-factors of the active site of the open and closed form subunits of the ternary complex. Cofactors and acyl-CoA ester are shown as stick models in red and salmon, respectively.*

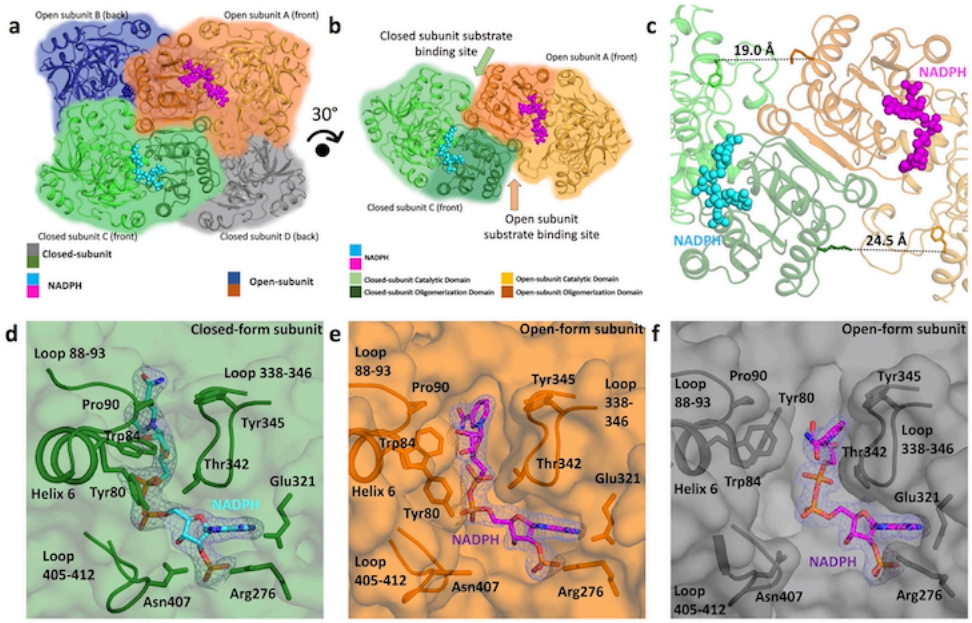


Figure 2: Binding of NADPH results in global and local conformational changes in K. setae ECR. ***a****. NADPH bound tetramer complex that is organized as dimer of dimers, a pair of closed (green) and open (orange) subunits and another pair containing closed (gray) and open (blue) subunits.* ***b****. The foreground dimer with open- (orange) and closed-form (green) subunits rotated by 30 degrees from the view in Figure 2a. Each monomer is composed of a catalytic and an oligomerization domain.* ***c****. Comparison of the putative substrate binding sites between the open and closed-form subunits.* ***d****. Surrounding residues and loops sealing NADPH in the closed-form subunit. In panels* ***d*** *to* ***f****, simple* $2F_O$-F_C *density contoured at 1.5 sigma level is shown for NADPH within 3 Å from the molecule.* ***e****. Surrounding residues and loops composing the looser binding of NADPH in the open-form subunit.* ***f****. Alternate binding of NADPH in the open-form subunit (gray subunit in* ***a****.), viewed in the same direction as in* ***e****.).*

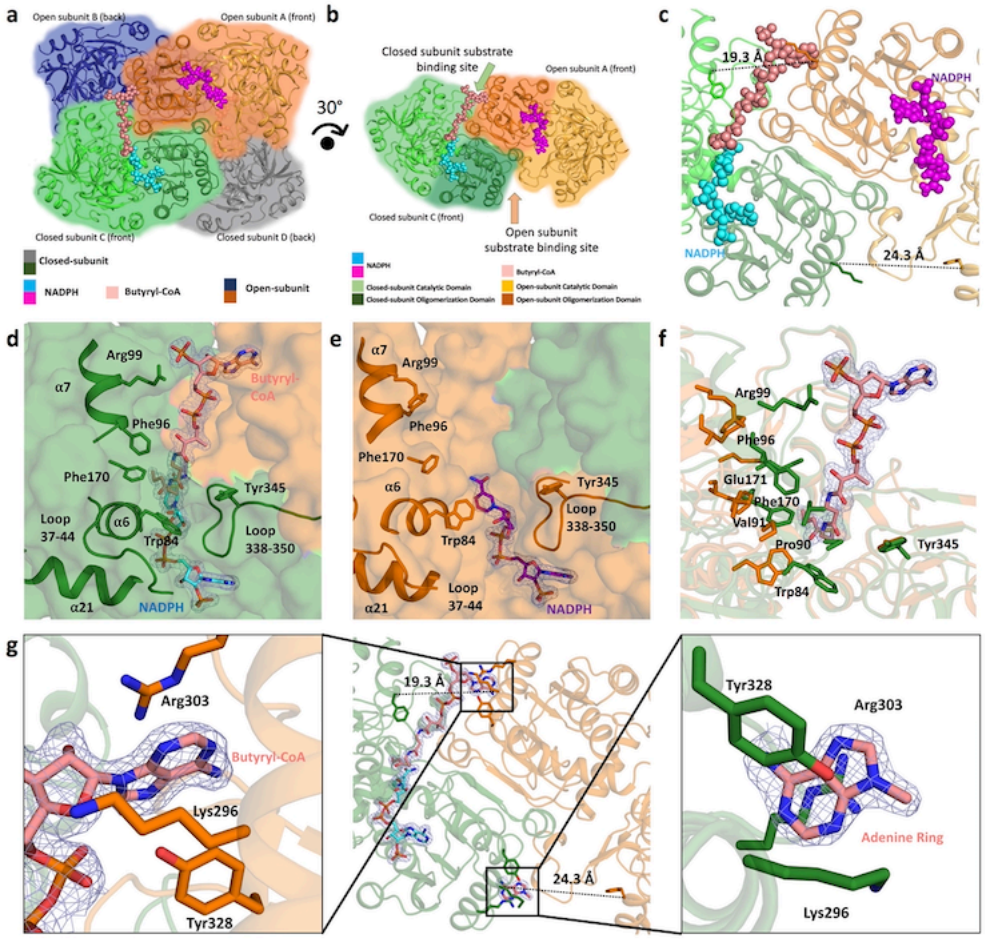


Figure 3: Structure of the ternary ECR complex. ***a****. ECR tetramer in complex with NADPH and butyryl-CoA organized as dimer of dimers, the foreground dimer with one closed subunit (green) with NADPH and butyryl-CoA and open (orange) subunit containing NADPH, and another pair in the background with one closed (gray) and open (blue) subunits.* ***b.*** *The foreground dimer with closed (green) and open (orange) subunits, rotated by 30 degrees from the view in Figure 3a. Butyryl-CoA and NADPH atoms are represented as spheres.* ***c****. Comparison of the product binding site between the open and closed-form subunits.* ***d****. Cartoon and stick representation of the closed-form subunit active site. In panels* ***d*** *to* ***g****, simple $2F_O$-F_C density contoured at 1.5 sigma level is shown for butyryl-CoA, or potion thereof, and NADPH within 3 Å from the molecules.* ***e****. Cartoon and stick representation of the open-form subunit active site.* ***f****. Superposition of the open-form subunit onto the closed-form subunit with stick representation of the residues surrounding butyryl-CoA.* ***g****. Comparison of the butyryl-CoA binding sites between open and closed-form subunits with electron density of the bound butyryl-CoA and NADPH at the active site of the closed subunit (green) and the adenine ring of butyryl-CoA at the active site of the open subunit (only the adenine ring electron density is visible). Left inset: the adenine binding pocket of the open-form subunit stabilizing the adenine ring of butyryl-CoA that stretches into the adjacent closed-form subunit. Right inset: the adenine binding pocket of the closed-form subunit holding the adenine ring of butyryl-CoA. Note that only the adenine ring of butyryl-CoA is visible, while the rest of the molecule is disordered. In both cases, three residues of the adjacent subunits, Lys296, Arg303, and Tyr328 together hold the adenine ring.*

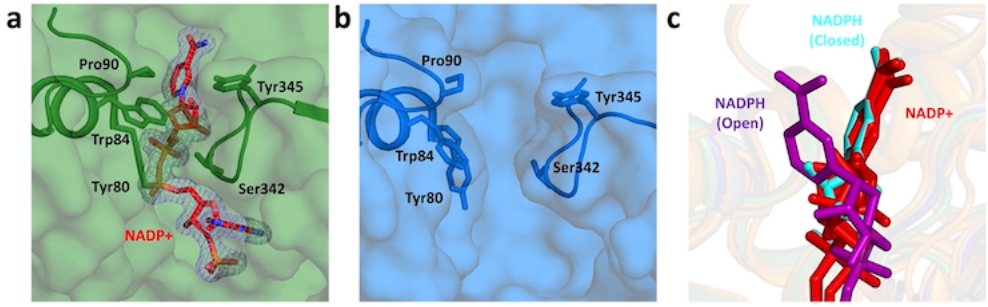
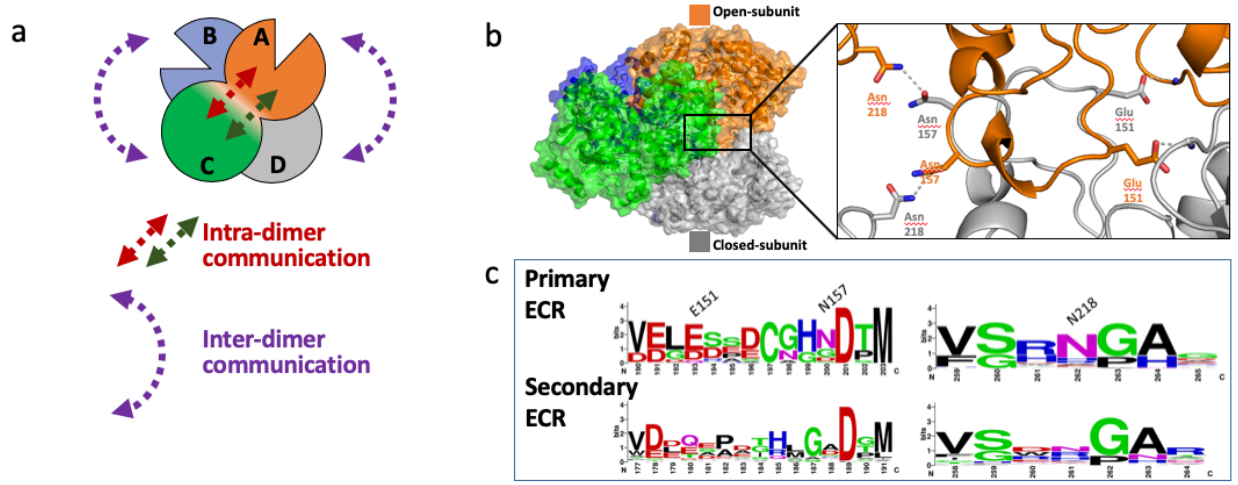
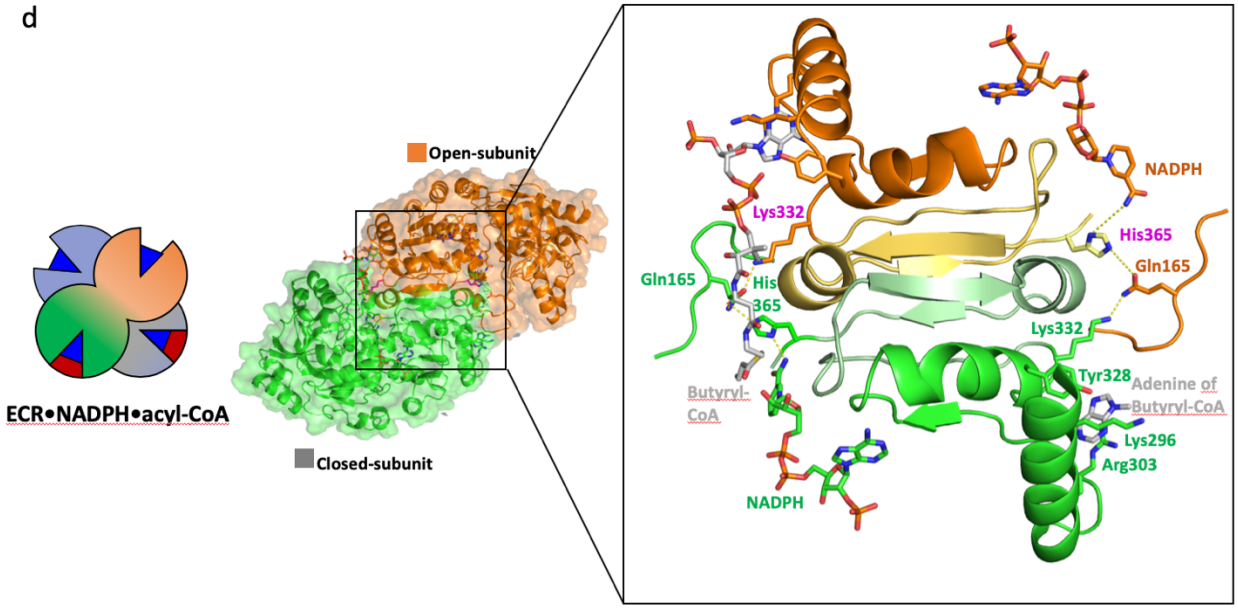


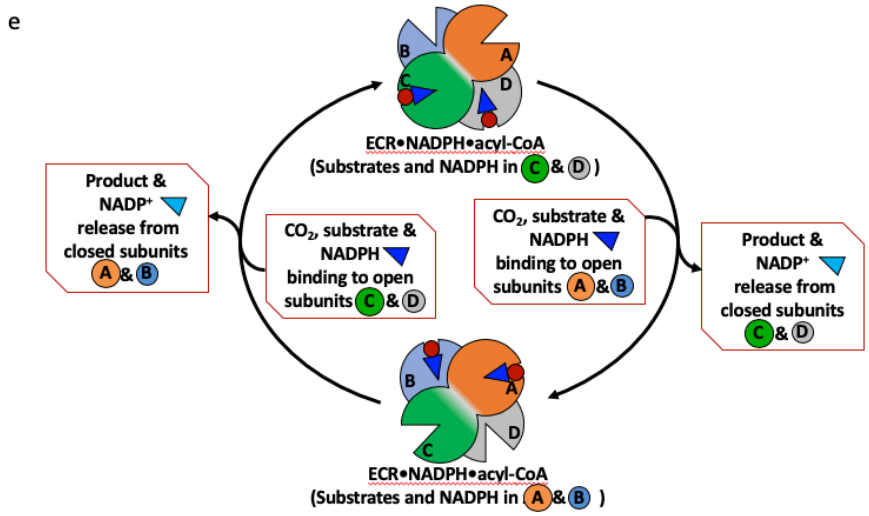
Figure 4: NADP⁺ binding at the active site of the NADP⁺ bound tetramer. a*. The simple $2F_O$-F_C electron density map of NADP⁺ bound in the K. setae ECR tetramer complex. Helix α6 and loop 88-93 (top left) close the binding site of NADP⁺.* ***b****. Cofactor binding site in the apo form subunit of the K. setae ECR tetramer. Helix α6 and loop 88-93 have moved away from the loop containing Ser342 and Tyr345, resulting in an open binding pocket for NADPH or NADP⁺. The two loops forming the cofactor binding pocket are wider apart by about 2 Å: the distances between the Cα atoms of Pro90 and Tyr345 are 13.1 Å in the open-form (blue) compared to 11.2 Å in the closed form (green).* ***c****. Superposition of the NADP⁺ molecules from all four subunits of the tetramer (red) with the open (purple) and closed-form (cyan) subunits of the NADPH/butyryl-CoA bound ternary complex.*



417



418



419

Figure 5. Inter- and intra-dimer communications drive fast CO_2 fixation by K. setae ECR. ***a.*** *Two distinct sets of communications: inter-dimer interactions between the catalytic domains from two dimers (purple arrows) and intra-dimer communication between the open and closed subunits within each dimer (brown arrows).* ***b****. Inter-dimer catalytic domain interface and positions of selected amino acids that were mutated in this study in order to affect the interface between the two catalytic domains (open-form subunit in orange and closed form subunit in gray). The right panel shows the mutual H-bonding interaction between Asn218 and Asn157 from open and closed form subunits and H-bonding between Glu151 and N-atom from protein backbone.* ***c****. Alignment of ECR protein sequences from the primary (upper row) and the secondary (lower row) metabolism represented as sequence logos. Numbering of residues, above first row, is according to their position in K. setae ECR.* ***d****. Communication between the closed (green) and open (orange) subunits across the two dimers of K. setae ECR. In the closed conformation the contacts between NADPH-His365-Glu165 and Lys332 of the adjacent open monomer allow for the correct intra-dimer communication. In the open conformation the communication network is compromised as indicated by the increased distances between the amino acid sidechains that cause the incorrect positioning of the nicotinamide ring of NADPH.* ***e****. Continuous turnover scheme in which the open-form subunits (A and B) become closed by releasing the product and $NADP^+$ while, concurrently, the closed subunits (C and D) bind the substrate and NADPH thus becoming closed subunits. These alternating cycles are aided by the swing motions of tightly coupled catalytic subunits like two blades swinging synchronously.*

References

1. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S. & Erb, T. J. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science (80-.).* **354**, 900–904 (2016).
2. Erb, T. J. *et al.* Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10631–6 (2007).
3. Erb, T. J., Brecht, V., Fuchs, G., Muller, M. & Alber, B. E. Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Natl. Acad. Sci.* **106**, 8871–8876 (2009).
4. Erb, T. J. & Zarzycki, J. Biochemical and synthetic biology approaches to improve photosynthetic CO 2 -fixation. *Curr. Opin. Chem. Biol.* **34**, 72–79 (2016).
5. Rosenthal, R. G. *et al.* Direct evidence for a covalent ene adduct intermediate in NAD(P)H-dependent enzymes. *Nat. Chem. Biol.* **10**, 50–55 (2014).
6. Quade, N., Huo, L., Rachid, S., Heinz, D. W. & Müller, R. Unusual carbon fixation gives rise to diverse polyketide extender units. *Nat. Chem. Biol.* **8**, 117–124 (2012).
7. Zhang, L. *et al.* Rational Control of Polyketide Extender Units by Structure-Based Engineering of a Crotonyl-CoA Carboxylase/Reductase in Antimycin Biosynthesis. *Angew. Chemie Int. Ed.* **54**, 13462–13465 (2015).
8. Rao, S. T. & Rossmann, M. G. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **76**, 241–56 (1973).
9. Tono, K. *et al.* Diverse application platform for hard X-ray diffraction in SACLA (DAPHNIS): application to serial protein crystallography using an X-ray free-electron laser. *J. Synchrotron Radiat.* **22**, 532–537 (2015).
10. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–7 (2011).
11. Schlichting, I. & Miao, J. Emerging opportunities in structural biology with X-ray free-electron lasers. *Curr. Opin. Struct. Biol.* **22**, 613–26 (2012).
12. Helliwell, J. R. How to solve protein structures with an X-ray laser. *Science* **339**, 146–7 (2013).
13. Vogeli, B. et al. Combining Promiscuous Acyl-CoA Oxidase and Enoyl-CoA Carboxylase/Reductases for Atypical Polyketide Extender Unit Biosynthesis. *Cell Chem Biol* **25**, 833–839 (2018).
14. Bernhard, S. A. & MacQuarrie, R. A. Half-site reactivity and the “induced-fit” hypothesis. *J. Mol. Biol.* **74**, 73–78 (1973).
15. Seydoux, F., Malhotra, O. P., Bernhard, S. A. & Stark, G. Half-Site Reactivit. *CRC Crit. Rev. Biochem.* **2**, 227–257 (1974).
16. Levitzki, A., Stallcup, W. B. & Koshland, D. E. Half-of-the-sites reactivity and conformational states of cytidine triphosphate synthetase. *Biochemistry* **10**, 3371–3378 (1971).
17. Fersht, A. R., Mulvey, R. S. & Koch, G. L. E. Ligand binding and enzymic catalysis coupled through subunits in tyrosyl-tRNA synthetase. *Biochemistry* **14**, 13–18 (1975).
18. Hill, T. L. & Levitzki, A. Subunit neighbor interactions in enzyme kinetics: half-of-the-sites

reactivity in a dimer. *Proc. Natl. Acad. Sci.* **77**, 5741–5745 (1980).
19. LEVITZKI, A. & KOSHLAND, D. E. The Role of Negative Cooperativity and Half-of-the-Sites Reactivity in Enzyme Regulation. in *Current Topics in Cellular Regulation* 1–40 (1976). doi:10.1016/B978-0-12-152810-2.50008-5
20. Vivoli, M., Pang, J. & Harmer, N. J. A half-site multimeric enzyme achieves its cooperativity without conformational changes. *Sci. Rep.* **7**, 16529 (2017).
21. Danenberg, K. D. & Danenberg, P. V. Evidence for a sequential interaction of the subunits of thymidylate synthetase. *J. Biol. Chem.* (1979).
22. Anderson, A. C., O'Neil, R. H., DeLano, W. L. & Stroud, R. M. The Structural Mechanism for Half-the-Sites Reactivity in an Enzyme, Thymidylate Synthase, Involves a Relay of Changes between Subunits †. *Biochemistry* **38**, 13829–13836 (1999).
23. Finer-Moore, J. S., Lee, T. T. & Stroud, R. M. A Single Mutation Traps a Half-Sites Reactive Enzyme in Midstream, Explaining Asymmetry in Hydride Transfer. *Biochemistry* **57**, 2786–2795 (2018).
24. Hammes, G. G., Benkovic, S. J. & Hammes-Schiffer, S. Flexibility, Diversity, and Cooperativity: Pillars of Enzyme Catalysis. *Biochemistry* **50**, 10422–10430 (2011).
25. Peter, D. M. Doctoral Thesis. ETH Zurich; Switzerland (2016).
26. Stagno, J. R. *et al.* Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* **541**, 242–246 (2017).
27. Sierra, R. G. *et al.* Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).

Online Methods

Amplification & cloning of K. setae ECR

The *K. setae* enoyl-CoA carboxylase reductase (ECR) coding sequence was codon optimized using the *E. coli* codon frequency table, and synthesis constraints were removed using the Build Optimization Software Tools (BOOST) developed by DOE-Joint Genome Institute (JGI), USA[1]. Overlapping synthetic DNA fragments were obtained from Thermo Fisher Scientific and cloned into the NdeI site of the pET16b vector (Novagen) by using the Gibson Assembly HiFi kit (SGI-DNA). The resulted colonies were sequence verified by the PacBio sequencing platform.

Site-directed Mutagenesis of K. setae ECR

Mutations are introduced by the similar methods as described in the previous section. Two fragments flanking the mutagenesis site were amplified and the Gibson assembly was performed as described above. Below is the FASTA sequence of the ECR protein and list of primers that we used to introduce catalytic site single mutations Y328F, R303K, and K296A, and the triple mutant K296A/R303K/Y328F.

tr|E4N096|E4N096 KITSK Putative crotonyl-CoA reductase OS= Kitasatospora setae
MQEILDAILSGDAASADYAALALPESYRAVTLHKGEERMFDGLASRDKDPRKSLHLDDVP
LPELGPGEALVAVMASSVNYNTVWSSIFEPVSTFGFLERYGRLSPLTARHDLPYHVLGSD
LAGVVLRTGAGVNAWKPGDEVVAHCLSVELESPDGHNDTMMDPEQRIWGFETNFGGLAQL
ALVKTNQLLPKPKHLTWEEAASPGLVNSTAYRQLVSRNGAGLKQGDNVLIWGASGGLGSY
ATQYALAGGATPICVVSSPRKADICRAMGAEAIIDRSAEGYRFWKDEHHQDPREWKRLGG
KIREFTGGEDVDIVFEHPGRETFGASVYVTRKGGTIVTCASTSGYMHQYDNRYLWMSLKR
IVGSHFANYREAFEANRLVAKGKIHPTLSKVYALEETGQAALDVHHNKHQGKVGVLCLAP
REGLGVTDPELRSKHLTKINAFRNV

Single mutations were introduced with the following *_F&*_R primer pairs:
E4N096_DMP_064_Y328F_F: catccgtgttcgtgacccgcaaaggtggcactatcg
E4N096_DMP_064_Y328F_R: gcgggtcacgaacacggatgcaccgaaggtttcgcg
E4N096_DMP_064_R303K_F: ggtggcaaaatcaaggaattcaccggtggggaagacgtgg
E4N096_DMP_064_R303K_R: aattccttgattttgccacccagacgtttccactcacg
E4N096_DMP_064_K296A_F: agtgggcccgtctgggtggcaaaatccgtgaattcaccg
E4N096_DMP_064_K296A_R: ccagacgggcccactcacgcgggtcttggtggtgttcg

The double and triple mutants were introduced in the following order: By using Y328F plasmid we introduced R303K mutation to generate double mutant Y328F/R303K
For the triple mutant we used the following special primers pair
E4N096_DMP_064_triple_F: agtgggcccgtctgggtggcaaaatcaaggaattcaccg
E4N096_DMP_064_triple_R: ccagacgggcccactcacgcgggtcttggtggtgttcg

Mutagenesis of K. setae ECR subunit interface residues

Two fragments franking the mutagenesis site were amplified and the Gibson assembly was performed as with the list of primers that we used to introduce various combinations of subunit interface single mutations N157E, N218E, E151D, E151R, E151K, E151L, E151I

Seven single mutants were introduced with the following *_F&*_R primer pairs:

ECR_N157E_F: ccggacggtcacgaagacactatgatggacccagagcagc
ECR_N157E_R: catcatagtgtcttcgtgaccgtccggagattccagttcaacagacagg
ECR_N218E_F: gctggtgtctcgtgaaggcgccggcctgaaacagggtgacaacg
ECR_N218E_R: caggccggcgccttcacgagacaccagctgacgataagcggtagagttaacg
ECR_E151D_F: gtctgttgaactggattctccggacggtcacaacgacactatgatgg
ECR_E151D_R: gaccgtccggagaatccagttcaacagacaggcagtgagcaaccacctcgtcacc
ECR_E151R_F: gtctgttgaactgaggtctccggacggtcacaacgacactatgatgg
ECR_E151R_R: gaccgtccggagacctcagttcaacagacaggcagtgagcaaccacctcgtcacc
ECR_E151K_F: gtctgttgaactgaagtctccggacggtcacaacgacactatgatgg
ECR_E151K_R: gaccgtccggagacttcagttcaacagacaggcagtgagcaaccacctcgtcacc
ECR_E151L_F: gtctgttgaactgctgtctccggacggtcacaacgacactatgatgg
ECR_E151L_R: gaccgtccggagacagcagttcaacagacaggcagtgagcaaccacctcgtcacc
ECR_E151I_F: gtctgttgaactgatctctccggacggtcacaacgacactatgatgg
ECR_E151I_R: gaccgtccggagagatcagttcaacagacaggcagtgagcaaccacctcgtcacc

Five triple mutants were obtained with the following *_F&*_R primer pairs respectively (third mutation varies):

1) Asn157Glu, Asn218Glu, Glu151Asp
   ECR_N157E_N218E_E151D_F:
   ctgtctgttgaactggattctccggacggtcacgaagacactatgatggacccagagcagcgcatctgg
   ECR_N157E_N218E_E151D_R:
   gtccatcatagtgtcttcgtgaccgtccggagaatccagttcaacagacaggcagtgagcaaccacctcg
2) Asn157Glu, Asn218Glu, Glu151Arg
   ECR_N157E_N218E_E151R_F:
   ctgtctgttgaactgaggtctccggacggtcacgaagacactatgatggacccagagcagcgcatctgg
   ECR_N157E_N218E_E151R_R:
   gtccatcatagtgtcttcgtgaccgtccggagacctcagttcaacagacaggcagtgagcaaccacctcg
3) Asn157Glu, Asn218Glu, Glu151Lys
   ECR_N157E_N218E_E151K_F:
   ctgtctgttgaactgaagtctccggacggtcacgaagacactatgatggacccagagcagcgcatctgg
   ECR_N157E_N218E_E151K_R:
   gtccatcatagtgtcttcgtgaccgtccggagacttcagttcaacagacaggcagtgagcaaccacctcg
4) Asn157Glu, Asn218Glu, Glu151Leu
   ECR_N157E_N218E_E151L_F:
   ctgtctgttgaactgctgtctccggacggtcacgaagacactatgatggacccagagcagcgcatctgg
   ECR_N157E_N218E_E151L_R:
   gtccatcatagtgtcttcgtgaccgtccggagacagcagttcaacagacaggcagtgagcaaccacctcg
5) Asn157Glu, Asn218Glu, Glu151Ile
   ECR_N157E_N218E_E151I_F:
   ctgtctgttgaactgatctctccggacggtcacgaagacactatgatggacccagagcagcgcatctgg
   ECR_N157E_N218E_E151I_R:
   gtccatcatagtgtcttcgtgaccgtccggagagatcagttcaacagacaggcagtgagcaaccacctcg

Mutagenesis of Adenine binding residues and intra-dimer communication residues

Variants of the *K. Setae* Ecr were generated with the QuikChange® Site-Directed Mutagenesis Kit (Stratagene, La Jolla, USA) using 60 ng of template plasmid and the following forward and reverse primer pairs:

Mutation	**Forward Primer**	**Reverse Primer**
R303A	CTGGGTGGCAAAATCGCTGAATTCACCGGTG	CTGGGTGGCAAAATCGCTGAATTCACCGGTG
R303V	CTGGGTGGCAAAATCGTGGAATTCACCGGTGGG	CTGGGTGGCAAAATCGTGGAATTCACCGGTGGG
K332A	GTACGTGACCCGCGCAGGTGGCACTATC	GTACGTGACCCGCGCAGGTGGCACTATC
Q165A	CACTATGATGGACCCAGAGGCACGCATCTGGGGCTTCGAAAC	CACTATGATGGACCCAGAGGCACGCATCTGGGGCTTCGAAAC

Cell lysis, protein purification, and characterization

The cells were harvested by centrifugation (3000 rpm, 30 min) and the cell pellet was pooled. The pellet was resuspended in a lysis buffer containing 50 mm Tris-HCl pH 8.5, 1 M NaCl, 5% glycerol supplemented with 100µl Triton x100 per 100ml of final buffer volume (Sigma-Aldrich). The suspension was sonicated at 50% amplitude for 30 seconds three times. Immediately after the lysis, the suspension was ultra-centrifuged at 33,000 rpm for 40 minutes at 4°C.

The soluble fraction was pooled and was applied to a 10 ml Ni-NTA column and purified using an AKTA prime FPLC setup. The column was washed with 2 column volumes of HisA loading buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl, 10 mM imidazole) for equilibration. Preliminary attempts of His-tag purification were unsuccessful since the protein would precipitate out of solution during application to the column. This was remedied by adding 1 M L-proline (Sigma-Aldrich) to the lysis and HisB elution buffers to ensure the protein remains soluble. The soluble portion was then applied to Ni-NTA column, and then eluted using HisB elution buffer containing 50 mM Tris-HCl pH 8.5, 300 mM NaCl, 500 mM imidazole. The eluted fractions were collected on a fraction collector, and their purities were analyzed by SDS-PAGE, and pure fractions were pooled and concentrated to 10 mg/ml using MilliPore Amicon Ultra 30KDa molecular-weight cutoff concentrators.

Determination of the oligomeric state of KsCcr

Oligomeric state of *K. setae* ECR was determined by analytical size-exclusion chromatography. 260 µl containing 500 µg of purified protein were injected into a pre-equilibrated S200 INCREASE 10-300GL (GE Healthcare) column. Runs were performed using a 100 mM KH_2PO_4 pH=8.0 buffer at a flow of 0.75 ml/min. Protein size was determined by comparing the obtained retention volumes (RV) with a Gel filtration standard protein mixture (BioRad).

Protein	**RV (mL)**	**MW (kDa)**
K. setae ECR WT	12.98	205.1
K. setae ECR E151D	12.97	202.8
K. setae ECR N157E	13.00	199.1
K. setae ECR N218E	13.01	197.7
K. setae ECR E151DN157E N218E	12.92	207.5
Gel filtration standard		
Thyroglobulin	9.45	670
Gamma-globulin	12.65	158
Ovalbumin	15.52	44
Myoglobin	17.30	17

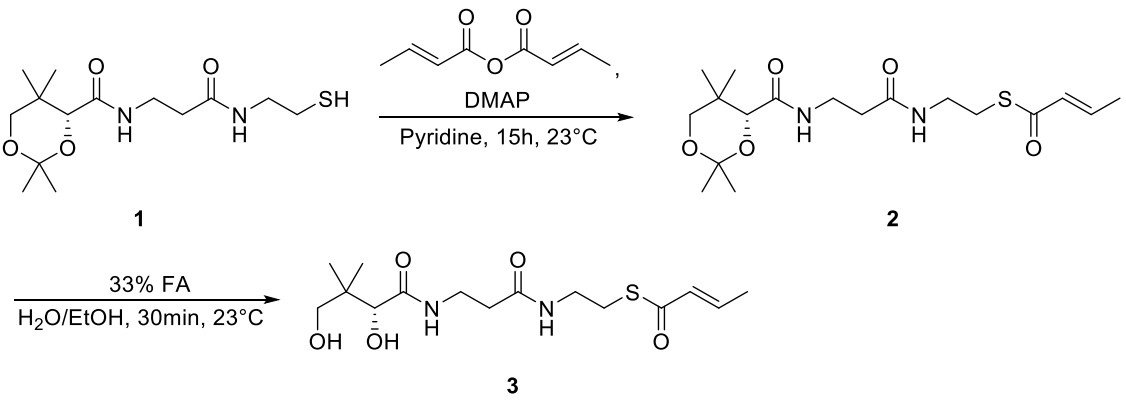
Spectrophotometric Enzyme assays

Assays were performed on a Cary-60 UV/Vis spectrophotometer (Agilent) at 30°C using quartz cuvettes (1 or 10 mm path length; Hellma). Reactions contained 20 µg/ml carbonic anhydrase and were performed in 100 mM K_2HPO_4 pH = 8.0. Kinetic parameters for one substrate were determined by varying its concentration while the others were kept constant at 10 times their K_M value. Reaction procedure was monitored by following the oxidation of NADPH at 365 nm ($\varepsilon_{NADPH,365nm}$ = 3.33 M^{-1} cm^{-1}). Each concentration was measured in triplicates and the obtained curves were fit using GraphPad Prism 8. Hyperbolic curves were fit to the Michaelis-Menten equation to obtain apparent k_{cat} and K_M values. For mutants revealing substrate inhibition, the data was fit to $v_0 = (V_{Max}\,[S])/(K_M + [S]\,((1+[S])/K_i)))$.

Chemical Synthesis of CoA-esters

Crotonic Anhydride, Carbonic anhydrase from bovine erythrocytes, 1,1-Carbonyldiimidazole (CDI) and 4-dimethylaminopyridine (DMAP) were purchased from Sigma Aldrich AG, Coenzyme A trilithium from Roche Diagnostics, NADPH Na_4 (98%) and pyridine from Carl Roth GmbH. Solvents and salts were all analytical grade or better. Crotonyl-CoA was synthesized as previously reported[2]. Briefly 200 mg of CoA trilithium salt were dissolved in 4 ml of 0.4 M $KHCO_3$ and stirred on ice for 45 min. After addition of 64 µl of crotonic anhydride the reaction procedure was tested by mixing 5 µl of reaction mixture with 20 µl of an aqueous solution of DTNB (5,5'-dithio-bis-[2-nitrobenzoic acid]). Crotonyl-CoA was purified by Preparative RFLC/MS over a Gemini 10 µm NX-C18 110 Å, 100 x 21.2 mm, AXIA packed column (Phenomenex) using a methanol gradient from 5% to 35% over 15 min with 25 mM ammonium formate pH = 8.1 (Buffer 8.1) as the aqueous phase. Fractions containing the product were pooled, lyophilized and stored at -20°C.

Synthesis of crotonyl-pantetheine (**3**) was performed according to scheme 1 as previously reported[2].



Scheme 1: Reaction conditions for the synthesis of crotonyl-pantetheine

Pantetheine **1** (0.50 g, 1.57 mmol), DMAP (0.02 g, 0.19 mmol) and crotonic anhydride (0.50 ml, 3.37 mmol) in pyridine (12.5 ml) were stirred for 15 h at 23° C then 1 h at 50 °C. Pyridine was removed under reduced pressure and the product dissolved in saturated aqueous $NaHCO_3$ (1 ml) and water (1 ml). The aqueous phase was extracted with CH_2Cl_2 (3x 5 ml), dried over $MgSO_4$, filtered and the solvent removed under reduced pressure. The obtained product was purified over FC (SiO_2; EtOAc/hexane, 2:3 → EtOAc) to afford **2**. **2** then stirred in Water/EtOH/FA, 1:1:1 for 30 min at 23 °C. After completion the solution was lyophilized, the solid dissolved in 0.5 % aqueous TFA and then purified with HPLC (C18; 25mM Ammonium formate pH = 8.1/MeOH, 5% → 95%) and lyophilized to yield **3** as a transparent thick oil. For use in assays the compound was resuspended in water and stored at -20°C if not used.

Analysis of Ethylmalonyl-CoA stability

Reactions to measure the stability of ethylmalonyl-CoA under crystallization conditions were performed in 200 mM TrisHCl pH = 7.5, 20% Polyacrylic acid sodium salt 5100 at 19°C. Reactions contained 600 µM Ethylmalonyl-CoA, $NADP^+$ µM and 1 µM KsECR WT. Samples were quenched at different time points using 50% formic acid and spinned at 17'000 g for 10 min to precipitate the protein. The reaction was diluted 10 times into 5% methanol/Buffer 8.1 and analyzed by UHPLC over a Sonoma C18(2), 3 µm 100 Å, 100 x 2.1 mm using a 5 to 45% methanol gradient over 14.5 min.

Crystallization of K.setae ECR complexes

72-well sitting-drop crystallization trays (Terasaki) were set up and screened against a library of various crystallization conditions (Molecular Dimensions, Hampton). Each crystallization well contained 0.77 µl of 10 mg/ml *K. setae* ECR protein kept in 500 mM Imidazole, 300mM NaCl, 1M proline and TRIS-HCl pH 8.5 mixed with 0.77 µL of the various crystallization buffers. Each well was sealed with 16.6 µL of 100% paraffin oil (Hampton Research) to slow the crystallization process. Crystals of apo ECR protein were observed in various morphologies after 24 hours of incubation. The initial crystallization conditions were from various MIDAS, Crystal Screen, and PGA-LM screening conditions (Molecular Dimensions, Hampton Research). The apo ECR was crystallied from a solution containing 100 mM TRIS pH 8.0 and 20% w/v poly (acrylic acid sodium salt) 5100 and resulted in 30-micron plate-like crystals. It is important

to note that all structures were solved using this condition as basis, with addition reagents as needed. The binary and ternary ECR complexes were co-crystallized with final concentration of 5 mM of each respective ligand and cofactor with a protein concentration of 10 mg/ml. Alternative crystallization conditions were used either to obtain larger crystals for higher resolution synchrotron structures or higher microcrystal density for SFX experiments. For the crystallization of *K. setae* ECR-butyrylCoA-NADPH ternary complex, the crystallization condition contained 17% w/v PEG 10000, 100 mM BisTris-HCl pH 7.5, 100 mM ammonium acetate and resulted in 50-micron plate-like crystals. For *K. setae* ECR-NADP+ binary complex, the crystallization solution contained 0.2 M ammonium formate, 10% (w/v) polyvinylpyrrolidone, 20% (w/v) PEG 4000 and resulted in 50-micron plate-like crystals. No further seeding was required for any of the synchrotron structures, and crystals were harvested after 30 minutes incubation with 30% (v/v) glycerol as a cryoprotectant. For the SFX experiments, neither of the synchrotron crystallization conditions of crystals would be sufficient due to size limitations and an optimal crystal density of 10^9 to 10^{11} crystals/ml could not be obtained. To test various crystal conditions, a batch method was employed with equal parts of protein and crystal condition to see if increased crystal densities could be achieved in 15 mL Corning conical falcon tubes. Initial tests were total volume of 1 ml (0.5 ml 10 mg/ml protein and 0.5 ml crystal condition) and incubated for 48 hours. From the tubes with crystals present, the 1 ml crystal slurry was used to seed a 10 ml total crystallization solution. The best crystals were obtained in final 10 ml sample solution consisted of the 1 ml seed crystal slurry solution, 4.5 ml of 10 mg/ml protein solution and 4.5 ml of crystallization buffer containing 0.03M Magnesium chloride hexahydrate, 0.03M Calcium chloride dihydrate, 0.05M imidazole, 0.05M MES-KOH pH 6.5, 15% v/v glycerol and 15% v/v PEG. Prior to sample injection, the crystals were filtered using a 20-micron nylon mesh filter to separate the contaminant of large crystals from the smaller ones (Millipore).

Data collection, processing and structure determination

For the apo, ternary, and NADP+ binary complex structures, the crystals were flash cooled in liquid nitrogen. The apo (1.8 Å) and NADP+ complex (1.75 Å) diffraction datasets were collected at 100 K on Beamline 23ID-B, the Advanced Photon Source, Argonne National Laboratory (Argonne, Illinois, USA), equipped with an Eiger 16M detector. The butyryl-CoA ternary complex (1.7 Å) diffraction dataset was collected at 100 K on Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory (Menlo Park, California, USA), equipped with a Dectris Pilatus 6M detector. The *K. setae* apo crystals belonged to the space group $P2_12_12_1$ with unit cell dimensions a= 78.1 Å, b= 153.0 Å, c= 202.7 Å and α=β=γ=90°. The *K. setae* ternary complex crystals belonged to the space group $P2_1$ with unit cell dimensions of a = 109.3 Å, b = 78.8 Å, c = 138.8 Å and α=90°, β=108.1°, γ=90°. The *K. setae* NADP+ complex crystals belonged to the space group $P2_12_12_1$ with unit cell dimensions of a = 77.0 Å, b = 146.7 Å, c = 200.2 Å and α=β= γ=90°. The *K. setae* ECR-NADPH binary complex was determined using serial femtosecond X-ray crystallography (SFX) at an X-ray Free Electron Laser (XFEL) and was carried out on May 2017 at SACLA beamline 3 (Hyogo, Japan) (Proposal number 2017A8055)[3]. The SALCA beam had a pulse duration of 10 fs. The photon energy was 10kEV. The in air concentric Electrokinetic Microfluidic Sample Holder (coMESH) injector[4] installed at DAPHNIS[5] chamber was used to introduce samples suspended in mother liquor to the 10 fs-long X-ray pulses. X-ray diffraction data was recorded by using the multiport CCD (MPCCD)[6] detector. Data analysis was performed on the SACLA High Performance Computing Cluster consisting of several steps of parameter optimization. Diffraction images were collected with consistent experimental parameters (attenuation, transmission, detector distance etc.) during one 12-hour shift. Crystal hits were identified with the program Cheetah[7]. The raw data were processed with CrystFEL's indexamajig against given cell parameters of the *K. setae*

NADPH (XFEL) complex microcrystals belonging to the space group $P2_1$ with unit cell dimensions of a = 109.8 Å, b = 78.1 Å, c = 138.9 Å and α=90°, β=107.8°, γ=90°.

The data processing for synchrotron structures were carried out using *autoXDS* and scaling was done with XSCALE[8,9]. A set of 5% of randomly chosen reflections were set aside for the calculation of the free R factor (R_{free}). The apo structure was solved using by *PHENIX*[10,11] and *PHASER*[12,13] molecular replacement program. Initial search model for molecular replacement is generated by using *SWISS-MODEL*[14] server against an unpublished CCR structure of a putative crotonyl-CoA carboxylase/reductase (PDB code 4GI2, deposited by S. Weidenweber, T.J. Erb, U. Ermler). The *K. setae* apo structure served as the model for solving the binary and ternary-complex synchrotron structures and also SFX structure. This resulted in four monomers in the asymmetric unit. The refinement was carried out using *PHENIX* refinement, utilizing automatically generated TLS groups based on the structure and ordered solvent to place the water molecules[15,16]. Following the first round of refinement, the structure was manually adjusted to the electron density and waters were added using *COOT* at one sigma cutoff [17,18]. The *K. setae*/NADP+ complex also shares the same space group as the apo form, and was solved directly using Phenix molecular replacement[10,11]. The NADP+ structure and restraint files were taken from previously solved CCR/NADP+ complexes (PDB: 4Y0K).

MD and QM/MM studies of the open and closed subunits of ECR

The flexibility of NADPH in the open and closed cavity were studied with a binary complex using the CHARMM22 force field[19,20] and parameters from Pavelites *et al.* for NADPH[21] in 200 ns explicit solvent simulations (TIP3P) at 298 K and 1 atm with the AMBER16 software package[22] (dt = 2 fs, tau = 1 ps, PME cutoff = 8.0 Å , SHAKE). Additionally, in three independent 200 ns simulations NADPH was positioned in the open cavity in the conformation observed in the closed cavity to test if these conformations are also visited in the open form.
To study crotonyl-CoA binding in the ternary complex we first extracted a dimer with one open- and another closed-subunits (subunit A & C) from the butyryl-CoA/NADPH ECR ternary X-Ray structure. We added the unresolved butyryl-CoA molecule to the open subunit aligning the protein chains from the closed subunit on the open one and shifting the butyryl-CoA coordinates from the closed subunit. The butyryl-CoA molecules were then modified to obtain the substrate crotonyl-CoA deleting the two hydrogen atoms. The resulting dimer consisted of one closed and one open subunit each with NADPH and one crotonyl-CoA molecule. The system was solvated and equilibrated (500ps NVT, 5ns NPT, 100ns NVT) as described above and substrate and NADPH were restrained to their initial configuration to relax the protein and the solvent. From these equilibrated configurations five structures were randomly extracted to study the behavior of the cofactor and substrate molecules. Five trajectories of 2 ns each for the closed and open cavity were performed with the QM/MM method using the DFTB3 Hamiltonian[23] and the 3ob parameter[24,25] set to describe NADPH and the crotonyl fragment. An electrostatic embedding using the link atom method at 298K and 1 atm was used together with a time step of 1 fs in AMBER16 software package. Parameters for the CoA fragment of the substrate were taken from Aleksandrov *et al*[26].

References

1. Oberortner, E., Cheng, J.-F., Hillson, N. J. & Deutsch, S. Streamlining the Design-to-Build Transition with Build-Optimization Software Tools. *ACS Synth. Biol.* **6**, 485–496 (2017).
2. Agarwal, V. *et al.* Chemoenzymatic Synthesis of Acyl Coenzyme A Substrates Enables in Situ

Labeling of Small Molecules and Proteins. *Org. Lett.* (2015). doi:10.1021/acs.orglett.5b02113

3. Ishikawa, T. *et al.* A compact X-ray free-electron laser emitting in the sub-ångström region. *Nat. Photonics* **6**, 540–544 (2012).
4. Sierra, R. G. *et al.* Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).
5. Tono, K. *et al.* Diverse application platform for hard X-ray diffraction in SACLA (DAPHNIS): application to serial protein crystallography using an X-ray free-electron laser. *J. Synchrotron Radiat.* **22**, 532–537 (2015).
6. Kameshima, T. *et al.* Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments. *Rev. Sci. Instrum.* (2014). doi:10.1063/1.4867668
7. Barty, A. *et al.* Cheetah : software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47**, 1118–1131 (2014).
8. Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
9. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 133–144 (2010).
10. Adams, P. D. *et al.* PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
11. Adams, P. D. *et al.* PHENIX : building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
12. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* (2007). doi:10.1107/S0021889807021206
13. McCoy, A. J. Acknowledging Errors: Advanced Molecular Replacement with Phaser. in *Methods in Molecular Biology* 421–453 (2017). doi:10.1007/978-1-4939-7000-1_18
14. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gky427
15. Winn, M. D., Isupov, M. N. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2001). doi:10.1107/S0907444900014736
16. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS Refinement in REFMAC at Moderate Resolutions. *Methods Enzymol.* (2003). doi:10.1016/S0076-6879(03)74014-2
17. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot . *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/s0907444910007493
18. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2004). doi:10.1107/S0907444904019158
19. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J. Comput. Chem.* (2004). doi:10.1002/jcc.20065
20. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* (1998). doi:10.1021/jp973084f
21. Pavelites, J. J., Gao, J., Bash, P. A. & Mackerell, A. D. A molecular mechanics force field for NAD+, NADH, and the pyrophosphate groups of nucleotides. *J. Comput. Chem.* (1997). doi:10.1002/(SICI)1096-987X(19970130)18:2<221::AID-JCC7>3.0.CO;2-X
22. Case, D. A. *et al.* Amber 2017. *University of California, San Francisco* (2017). doi:citeulike-article-id:2734527

23. Gaus, M., Cui, Q. & Elstner, M. DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.* (2011). doi:10.1021/ct100684s
24. Gaus, M., Goez, A. & Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* (2013). doi:10.1021/ct300849w
25. Gaus, M., Lu, X., Elstner, M. & Cui, Q. Parameterization of DFTB3/3OB for sulfur and phosphorus for chemical and biological applications. *J. Chem. Theory Comput.* (2014). doi:10.1021/ct401002w
26. Aleksandrov, A. & Field, M. Efficient solvent boundary potential for hybrid potential simulations. *Phys. Chem. Chem. Phys.* (2011). doi:10.1039/c0cp02828b

Supplementary information

Supplementary Table 1. Data collection and refinement statistics

	K. setae ECR Apo form	*K. setae* ECR NADP+ Butyryl-CoA	*K. setae* ECR $NADP^+$	*K. setae* ECR NADPH (XFEL)
Data collection				
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_1$
Cell dimensions				
a, *b*, *c* (Å)	78.1, 153.0, 202.7	109.3, 78.8, 138.8	77.0, 146.7, 200.2	109.8, 78.1, 138.9
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 108.1, 90.0	90.0, 90.0, 90.0	90.0, 107.8, 90.0
Resolution (Å)	16.9 – 1.80 (1.86 – 1.80)	38.4 – 1.78 (1.84 – 1.78)	47.7 – 1.75 (1.81 – 1.75)	31.8 – 2.34 (2.42 – 2.34)
R_{sym} or R_{merge}	0.12 (2.38)	0.08 (0.78)	0.17 (3.26)	-
R_{split}	-	-	-	0.31 (1.36)
$I / \sigma I$	12.3 (1.56)	16.6 (2.07)	11.9 (0.89)	6.7 (0.74)
Completeness (%)	100.0 (100.0)	98.8 (98.5)	96.5 (88.5)	100.0(100.0)
Redundancy	6.14	6.67	6.21	129.2
Refinement				
Resolution (Å)	16.93 – 1.80 (2.42 – 2.34)	38.34 – 1.78 (1.81 – 1.78)	47.69 – 1.75 (1.82 – 1.75)	31.75 – 2.34 (2.42 – 2.34)
No. reflections	223907 (18723)	227701 (16317)	220094 (24907)	93978 (9367)
R_{work} / R_{free}	0.20/0.24	0.17/0.20	0.19/0.21	0.28/0.31
No. atoms				
Protein	13791	14067	13796	13796
Ligand/ion	45	349	192	192
Water	1369	1885	1749	587
B-factors				
Protein	48.6	25.6	33.0	35.9
Ligand/ion	54.2	33.2	39.4	41.6
Water	48.2	35.3	37.3	37.1
R.m.s. deviations				
Bond lengths (Å)	0.010	0.006	0.007	0.015
Bond angles (°)	1.17	0.99	1.15	1.53

*Single crystal used for all datasets except the XFEL NADPH dataset.

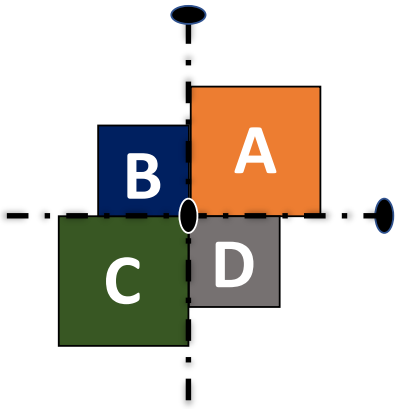
**Values in parentheses are for highest-resolution shell.

Apo

Chain	RMSD (Å)
A/B	0.11
A/C	0.08
A/D	0.11
B/C	0.11
B/D	0.08
C/D	0.08

NADPH complex

Chain	RMSD (Å)
A/B	0.50
A/C	1.84
A/D	1.78
B/C	1.74
B/D	1.70
C/D	0.52

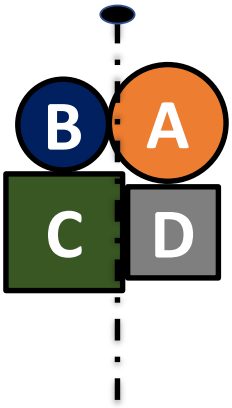


ECR + NADPH + Butyryl-CoA

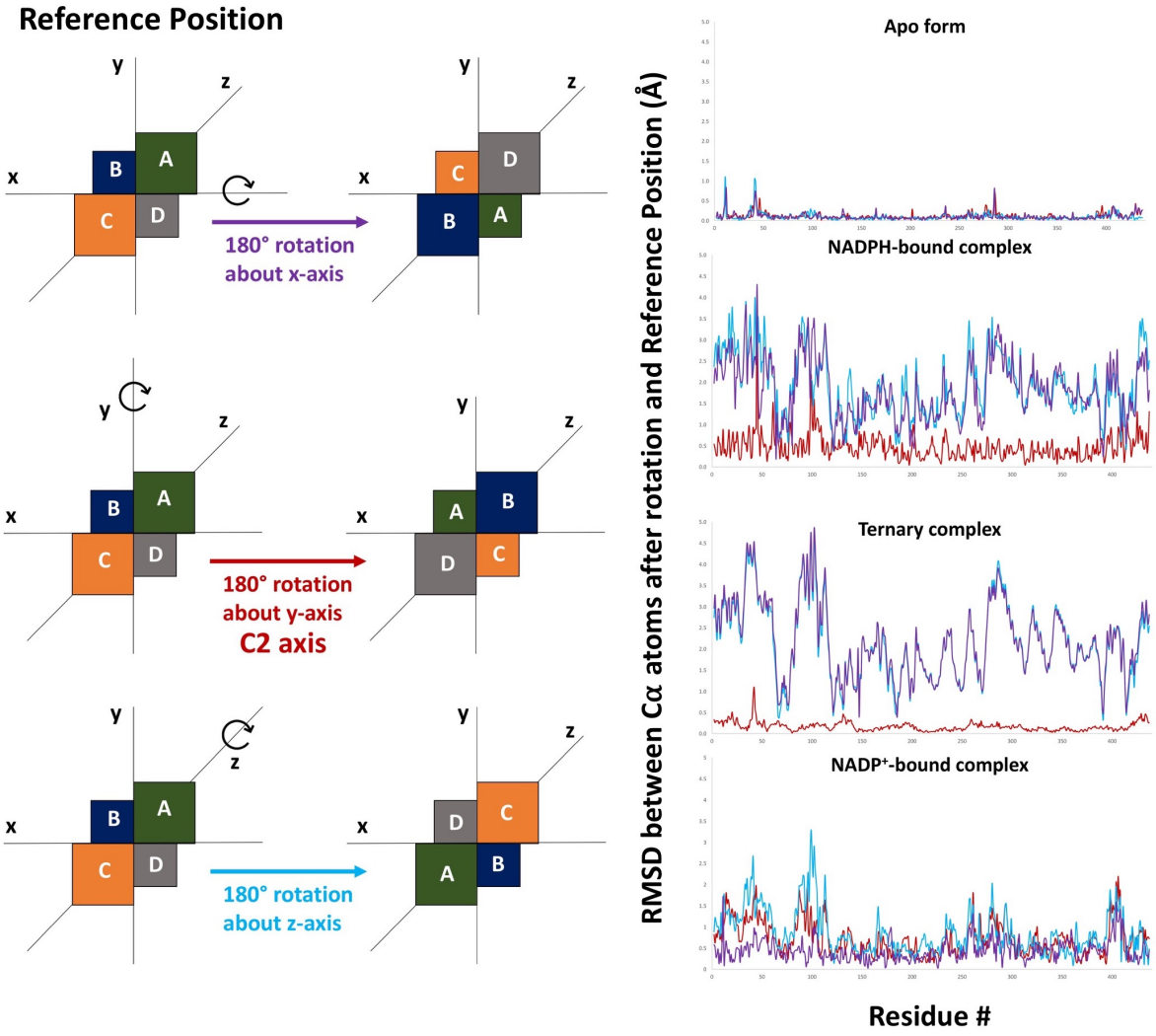
Chain	RMSD (Å)
A/B	0.10
A/C	2.04
A/D	2.06
B/C	2.07
B/D	2.04
C/D	0.08

NADP+ complex

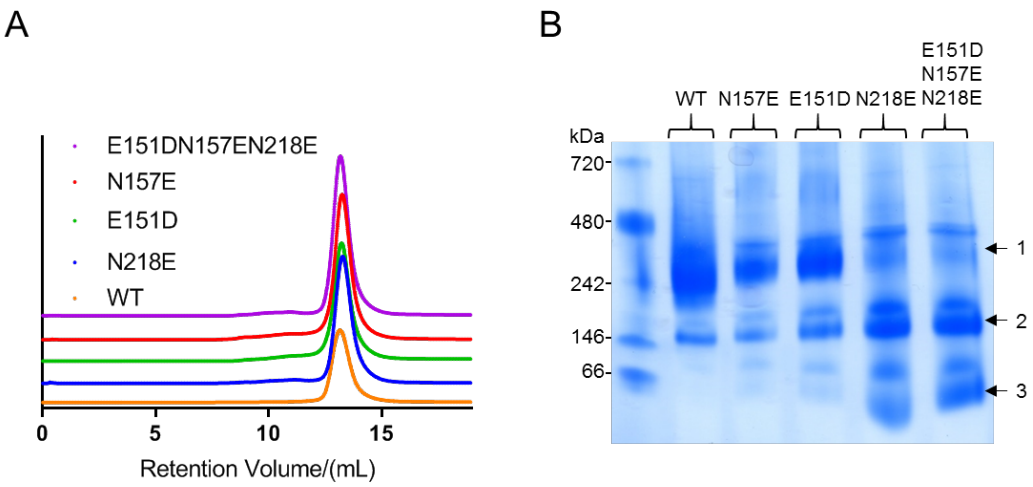
Chain	RMSD (Å)
A/B	0.18
A/C	0.70
A/D	0.25
B/C	0.85
B/D	0.35
C/D	0.49



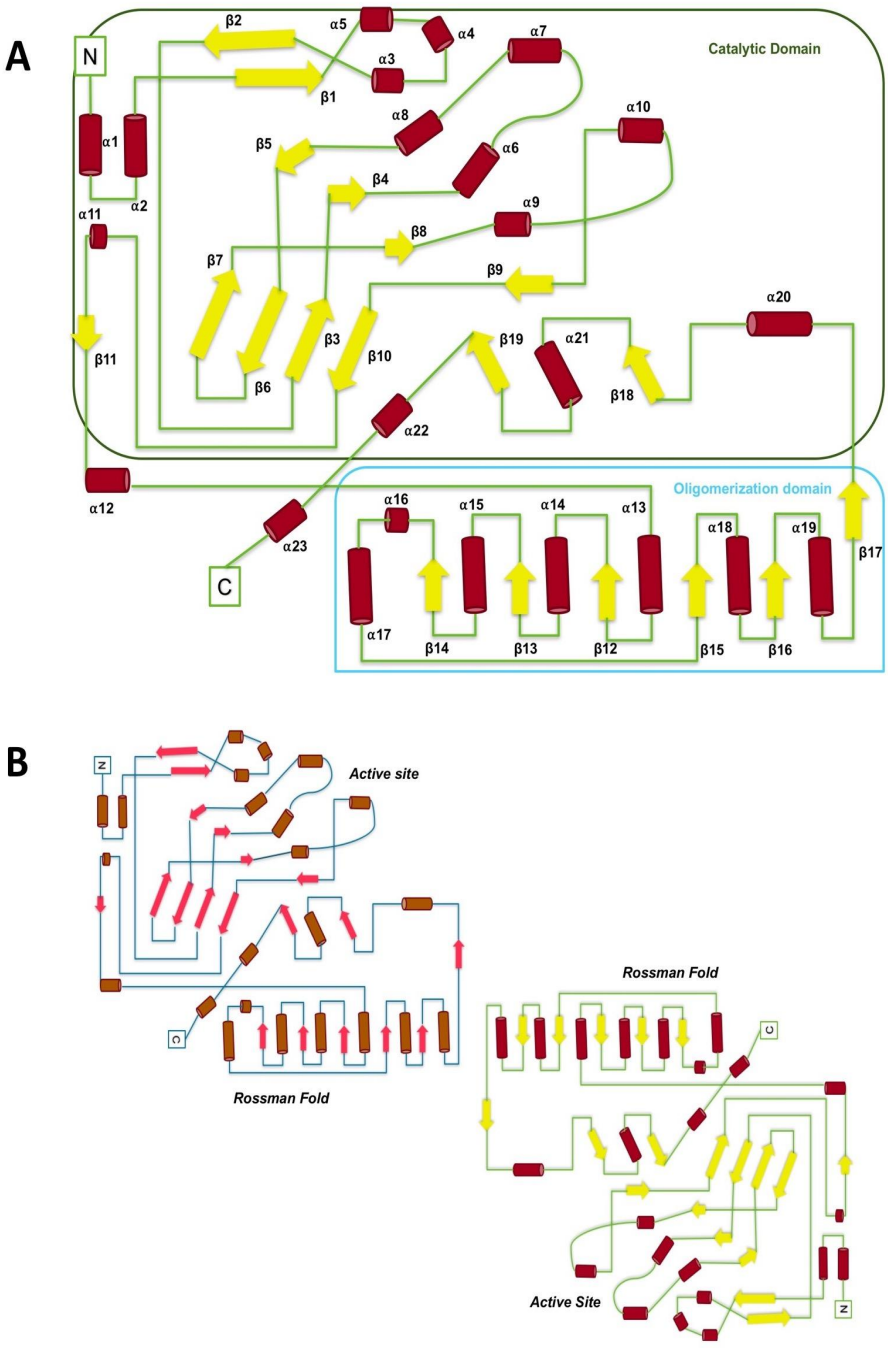
Supplementary Figure S1: Root-mean-squared deviations of the 4 subunits from the different ECR crystal structures. The Cα main chain atoms from each subunit were used for the alignments. The reported values in the left 4 tables are in Å. In the ternary complex, subunits A and B are open-form subunits having only cofactors while C and D are closed-form subunits which have the cofactors and butyryl-CoA bound. Top right: the arrangement of the 4 subunits in the apo and NADP+ bound ECR with dihedral D2 symmetry, 3 two-fold symmetry axes are shown with ovals and broken lines. Bottom right: the arrangement of the 4 subunits in the NADPH and butyryl-CoA/NADPH bound ECR, with vertical cyclic C2 axis shown by a set of oval and broken line. The viewing angles of these two are the same as in Figures 2a and 3a.



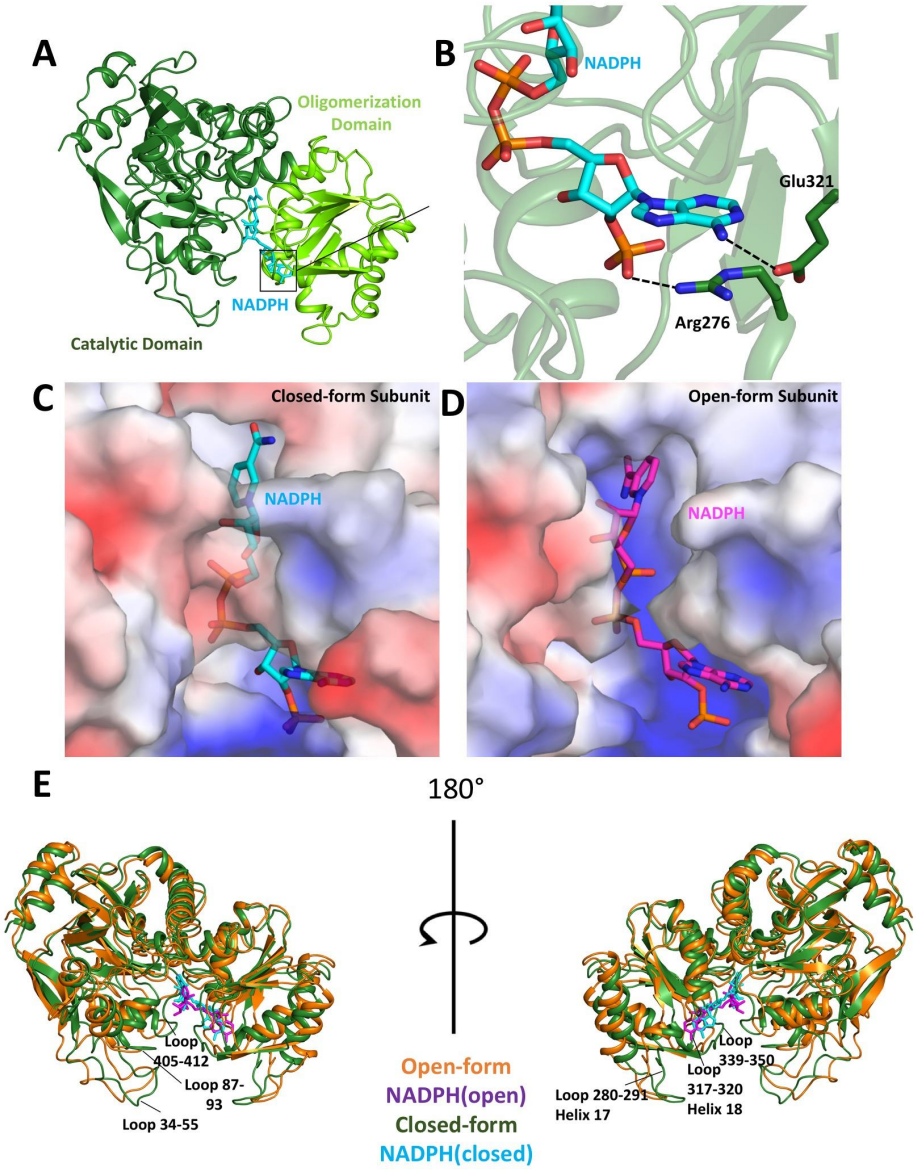
Supplementary Figure S2: This figure shows the disparity among subunits to reveal the broken symmetry. The left-hand side shows the various rotations that were used to test the symmetry. The right-hand side graphs the RMSD values of the post-rotation aligned with the reference position against each residue. The rotation axes are color coded as follows: x-axis = purple, y-axis = Red, z-axis = cyan.



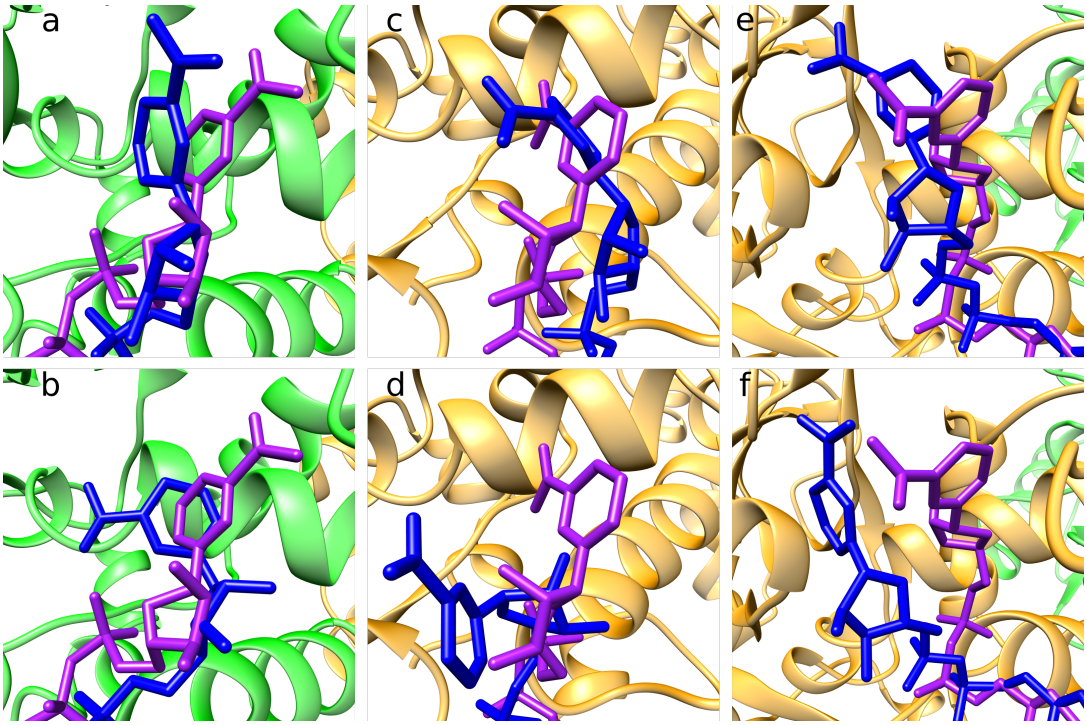
Supplementary Figure S3: A) Size-exclusion chromatography of *K. setae* ECR WT and variants. The WT protein and its variants elute at the same retention volume confirming the expected size of the protein. B) Blue native PAGE of *K. setae* ECR wildtype and variants. Oligomeric states according to numbering 1: tetramer 2: dimer 3: monomer.



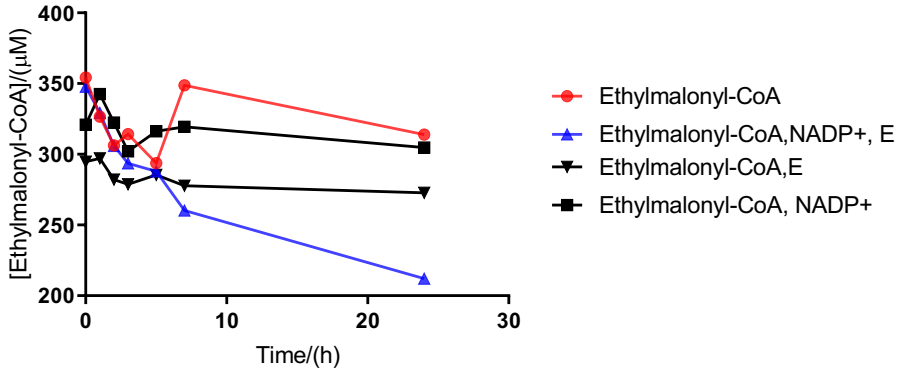
Supplementary Figure S4: Topology diagrams of both a monomer and a dimer which shows the central β-sheet network. A) The monomeric representation shows the disparity between the flexible loops of the catalytic domain and the rigid Rossmann fold in the oligomerization domain. B) This dimer representation shows how the 12 β-strand Rossmann fold network is formed which links the two subunits. The active site is also highlighted to show that is on the edge of the catalytic domain.



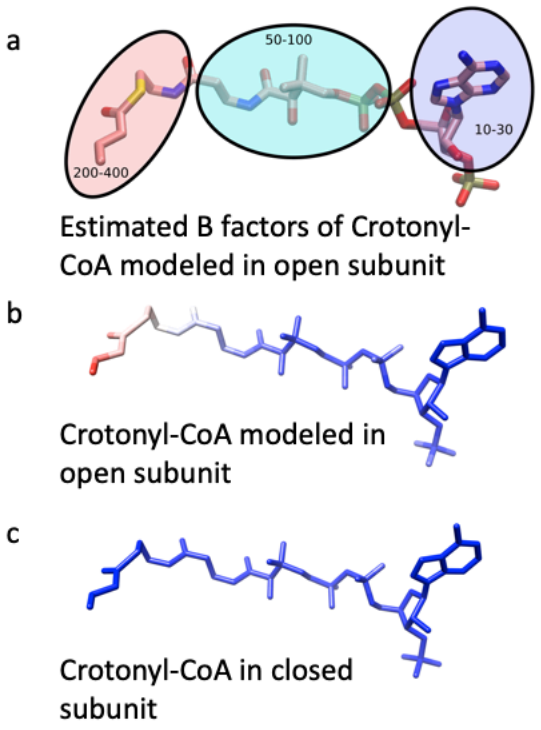
Supplementary Figure S5: A.) The NADPH spans from the adenine ring which is stabilized by the oligomerization domain to the catalytic domain, where the nicotinamide group is located. B.) The conserved stabilization of NADPH between the open and closed-form subunits is from Arg276 and Glu321. C.) The electrostatic surface of the co-factor binding pocket in the closed-form subunit. D.) The electrostatic surface of the co-factor binding pocket in the open-form subunit. E.) The conformational disparity between the open and closed-form subunits from two view angles, rotated 180°.



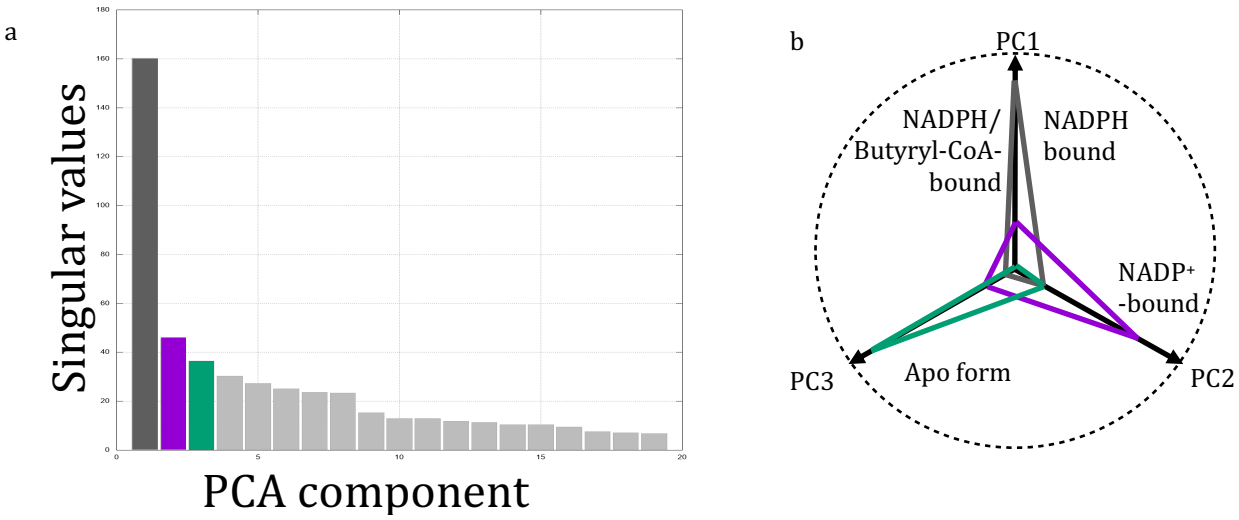
Supplementary Figure S6: Comparison between snapshots of molecular dynamics simulations (blue) and crystal structures (purple) of the binary complex in the closed and open subunits. For the closed subunits (a, b) the two predominant conformations sampled by NADPH fit very close to the crystal one. For the open subunits, the larger binding pocket allows the sampling of a wider range of conformations (d,f) including states similar to the one in the X-Ray structure (c,e).



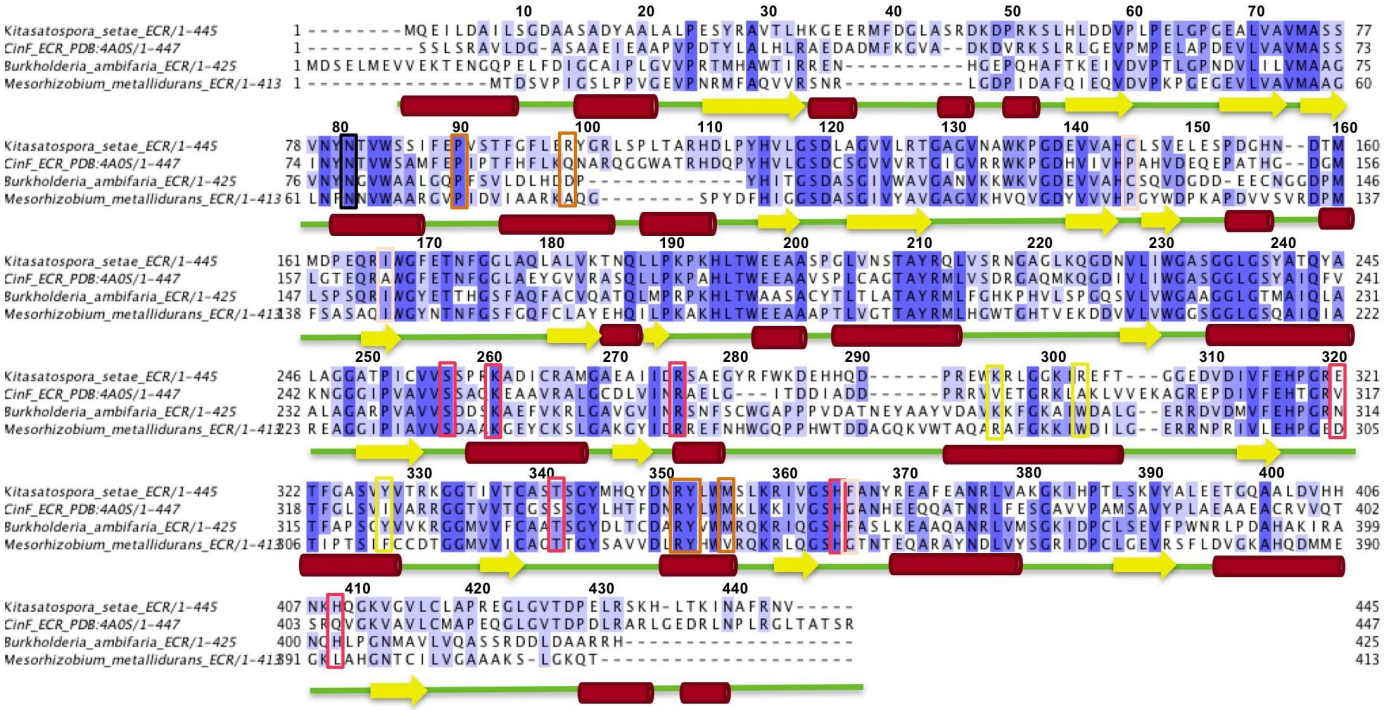
Supplementary Figure S7: Stability of ethylmalonyl-CoA stability measured over time at 19°C. The trace represented by blue triangles is the only condition in which ethylmalonyl-CoA is depleted over time. The enzyme catalyzed decarboxylation of ethylmalonyl-CoA to butyryl-CoA and CO_2 is the cause for the observed disappearance of ethylmalonyl-CoA. "E" corresponds to 1 µM *K. setae* ECR wt in the reaction mixture.



Supplementary Figure S8: QM/MM simulations of the substrate crotonyl-CoA and NADPH at the active sites. **a**. The differential flexibility of the Crotonyl-CoA substrate is reflected in the atomic fluctuations obtained from the dynamics. A. Range of B-factors (in $Å^2$) estimated from the squared atomic fluctuations (RMSF) and weighted by $(8/3)\pi^2$. **b**. Crotonyl-CoA molecule flexibility as modelled and simulated in the open subunit. **c**. Crotonyl-CoA molecule in the closed. The color reflects the estimated B-factor of each atom from low (blue) to high (red). (See also **Supplementary Videos 1&2**)

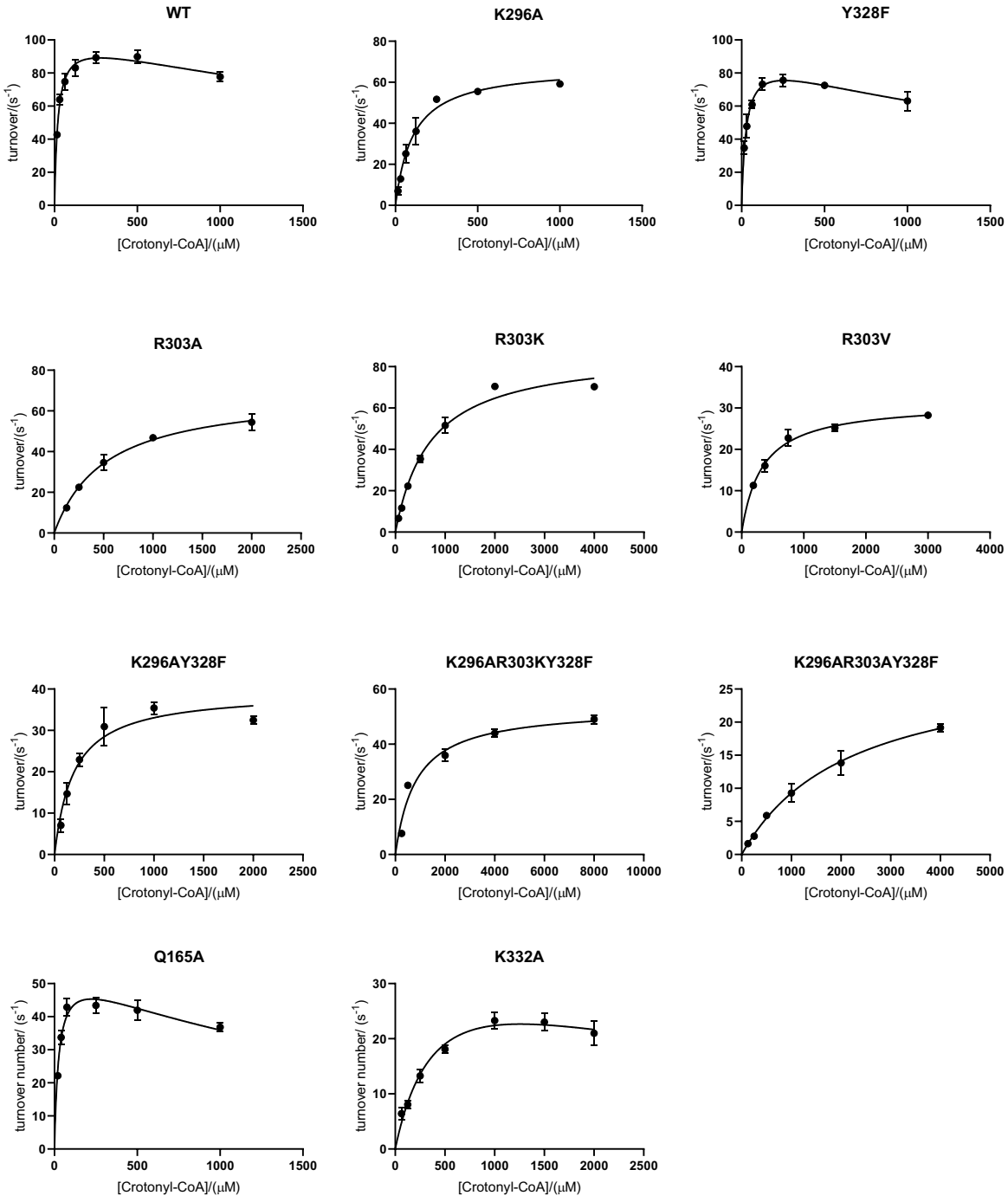


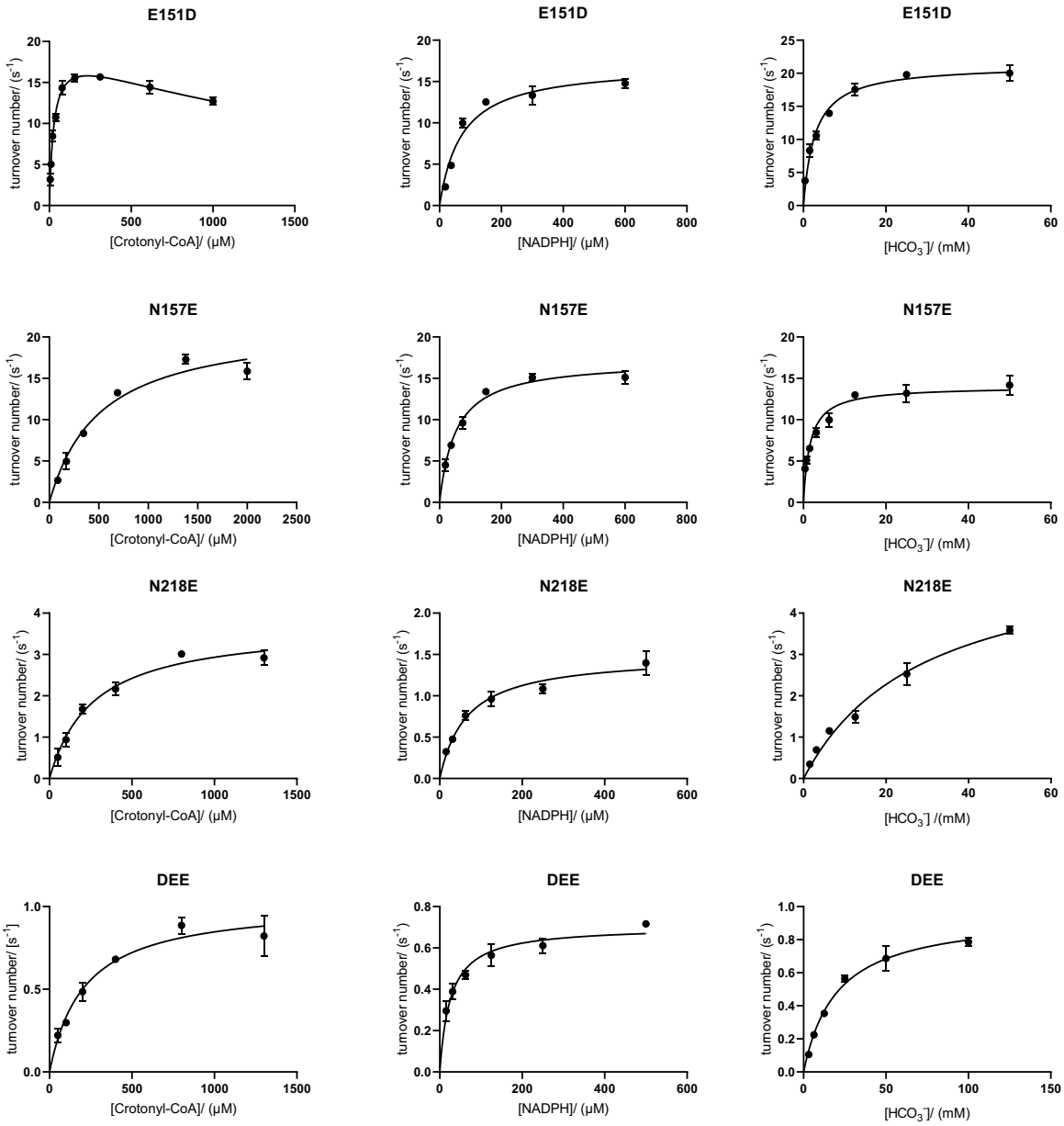
Supplementary Figure S9: Principal Component Analysis shows coupled motions of the catalytic domains to create alternating open and closed subunits. **a**. Singular values of the PCA components are plotted, showing the first component is the major contributor to the overall structural change. **b**. Contributions of PC1 to PC3 to the deviations of the NADPH-bound and NADPH/butyryl-CoA-bound structures from the average structure. See **Supplementary Videos 5** showing a morphing movie showing the structural changes between the apo form and ternary complex, and three most significant PCA motions.

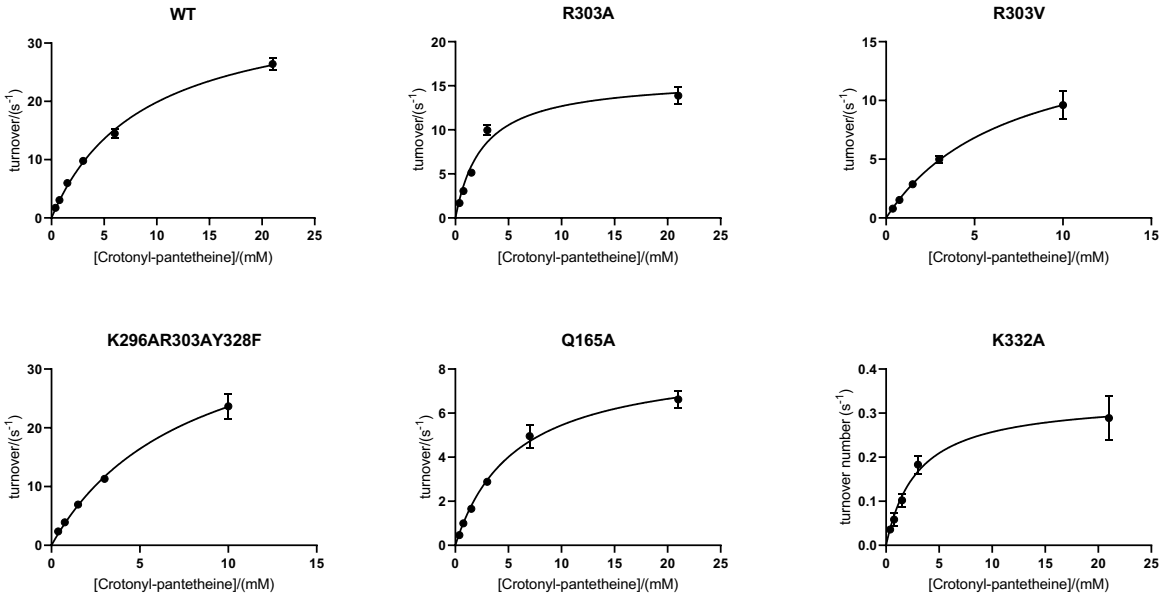


Supplementary Figure S10: The sequence alignment between two primary ECRs: *K. setae* & *M. metallidurans*, and two secondary ECRs: *B. ambifaria* & CinF from *S. collinus.* Two ECRs use crotonyl-CoA as their substrate: *K. setae* and *B. ambifaria*, while the others use octenoyl-CoA: *CinF* and *M. metallidurans.* The boxes in pink represent the residues which interact with NADP+/NADPH. The residues in orange represent the residues which interact with the bound CoA. The residues in salmon represent the residues involved in substrate specificity. The residues in yellow represent the novel adenine ring CoA binding pocket discovered in the *K.setae* ECR. The residue in black (Asn81) represents the residue in question for stabilizing the CO_2 molecule for reductive carboxylation.

Supplementary Figure S11: Steady state parameters of *K. seate* ECR and its variants. The data are summarized in Table 1 and Table 2.




E151D
turnover number/ (s^{-1})
[Crotonyl-CoA]/ (μM)
E151D
turnover number/ (s^{-1})
[NADPH]/ (μM)
E151D
turnover number/ (s^{-1})
[HCO_3^-]/ (mM)
N157E
turnover number/ (s^{-1})
[Crotonyl-CoA]/ (μM)
N157E
turnover number/ (s^{-1})
[NADPH]/ (μM)
N157E
turnover number/ (s^{-1})
[HCO_3^-]/ (mM)
N218E
turnover number/ (s^{-1})
[Crotonyl-CoA]/ (μM)
N218E
turnover number/ (s^{-1})
[NADPH]/ (μM)
N218E
turnover number/ (s^{-1})
[HCO_3^-] /(mM)
DEE
turnover number/ [s^{-1}]
[Crotonyl-CoA]/ (μM)
DEE
turnover number/ (s^{-1})
[NADPH]/ (μM)
DEE
turnover number/ (s^{-1})
[HCO_3^-]/ (mM)


WT
R303A
R303V
K296AR303AY328F
Q165A
K332A
turnover/(s-1)
turnover number (s-1)
[Crotonyl-pantetheine]/(mM)