CrossMark click for updates

Structure of NDP-forming Acetyl-CoA synthetase ACD1 reveals a large rearrangement for phosphoryl transfer

Renato H.-J. Weiße[a,b], Annette Faust[a,c], Marcel Schmidt[d], Peter Schönheit[d,1], and Axel J. Scheidig[a,c,1]

[a]Zoological Institute-Structural Biology, Kiel University, 24118 Kiel, Germany; [b]Centre for Biotechnology and Biomedicine, Structure Analysis of Biopolymers, Leipzig University, 04103 Leipzig, Germany; [c]Centre for Biochemistry and Molecular Biology, Kiel University, 24118 Kiel, Germany; and [d]Institut für Allgemeine Mikrobiologie, Kiel University, 24118 Kiel, Germany

Edited by Gregory A. Petsko, Weill Cornell Medical College, New York, NY, and approved December 16, 2015 (received for review September 18, 2015)

The NDP-forming acyl-CoA synthetases (ACDs) catalyze the conversion of various CoA thioesters to the corresponding acids, conserving their chemical energy in form of ATP. The ACDs are the major energy-conserving enzymes in sugar and peptide fermentation of hyperthermophilic archaea. They are considered to be primordial enzymes of ATP synthesis in the early evolution of life. We present the first crystal structures, to our knowledge, of an ACD from the hyperthermophilic archaeon *Candidatus Korachaeum cryptofilum*. These structures reveal a unique arrangement of the ACD subunits alpha and beta within an $\alpha_2\beta_2$-heterotetrameric complex. This arrangement significantly differs from other members of the superfamily. To transmit an activated phosphoryl moiety from the Ac-CoA binding site (within the alpha subunit) to the NDP-binding site (within the beta subunit), a distance of 51 Å has to be bridged. This transmission requires a larger rearrangement within the protein complex involving a 21-aa-long phosphohistidine-containing segment of the alpha subunit. Spatial restraints of the interaction of this segment with the beta subunit explain the necessity for a second highly conserved His residue within the beta subunit. The data support the proposed four-step reaction mechanism of ACDs, coupling acyl-CoA thioesters with ATP synthesis. Furthermore, the determined crystal structure of the complex with bound Ac-CoA allows first insight, to our knowledge, into the determinants for acyl-CoA substrate specificity. The composition and size of loops protruding into the binding pocket of acyl-CoA are determined by the individual arrangement of the characteristic subdomains.

X-ray structure | metabolic energy conversion | protein dynamics | acyl-coenzyme A thioester | superfamily

NDP-forming acyl-CoA synthetases (ACDs) catalyze the conversion of acyl-CoA thioesters to the corresponding acids and couple this reaction with the synthesis of ATP via the mechanism of substrate-level phosphorylation. ACDs have been studied in detail in hyperthermophilic archaea, where they function as the major energy-conserving enzymes in the course of anaerobic sugar and peptide fermentation (1–4). It is believed that ACDs represent a primordial mechanism of ATP synthesis in the early evolution of life. ACDs were found in all acetate (acid)-forming archaea (5, 6) and in the eukaryotic parasitic protists *Entamoeba histolytica* (7) and *Giardia lamblia* (8), but they have not been found in acetate-forming bacteria. In bacteria, with the exception of *Chloroflexus* (9), the conversion of inorganic phosphate and the thioester acetyl (Ac)-CoA to acetate and ATP is catalyzed by two enzymes, phosphate Ac-transferase and acetate kinase (10). Following the identification of ACD genes (5, 11) a novel protein superfamily of NDP-forming ACDs was proposed by a bioinformatics analysis (8). In addition to ACDs, this superfamily contains the well-characterized succinyl-CoA synthetases (SCSs) and ATP-citrate lyases (ACLYs) (8). Each ACD is composed of at least five subdomains with variable sequential arrangement (8). This phenomenon, termed "domain shuffling," is one of the key features of this superfamily (8). Superposition of several structures of SCS from *Escherichia coli* (*ec*SCS) (12–14), *Thermus aquaticus* (15), the mammalian GTP-specific SCS from pig (16), and a truncated form of human ACLY (17, 18) revealed that subdomains 1–5 share a common arrangement in these enzymes. From detailed studies of the reaction mechanism of *ec*SCS, a crucial enzyme tightly connected to the TCA cycle, a three-step mechanism was proposed, which involves the phosphorylation of a highly conserved His residue at the first active site (site I) as an intermediate step (12, 13). Subsequently, the phosphoryl moiety is transferred onto an NDP that is bound at the second active site (site II). The distance of around 36 Å between site I and site II is assumed to be bridged by a so-called "swinging loop" (12, 14, 19). A mechanism comparable to the mechanism of the SCSs was presumed for the ACDs (20). A study based on sequence, biochemical, and mutational analyses identified a highly conserved and functional relevant His residue within the beta subunit of the ACDs (20). Of note, the SCS enzymes do not contain a comparable His residue. Thus, an extension of the mechanism by a fourth step was suggested for the ACD1 from *Pyrococcus furiosus* (20). Here, the second His residue is thought to serve as an additional intermediate, which gets phosphorylated transiently during the enzymatic reaction and subsequently transmits the phosphoryl moiety onto the bound NDP. The necessity for a four-step mechanism may originate from a shortening of the proposed swinging loop, which facilitates the phosphoryl transfer between the subunits (20). However, due to the completely different arrangement of the subdomains within ACD [alpha(1-2-5)/beta(3-4)] in comparison to *ec*SCS [alpha(1-2)/beta(3-4-5)], a different 3D arrangement of

Significance

Acyl-CoA thioesters are key substrates for energy conversion. Related ATP/GTP-producing synthetases form a large superfamily with members in all kingdoms of life. In contrast to their general importance, the underlying reaction mechanism of these enzymes is still not understood in all steps. Here, we describe various structures of a nucleoside diphosphate-forming acetyl–CoA synthetase from an evolutionary very old archaeon. A large conformational rearrangement within the enzyme is observed. The structures reveal a partial unwinding and reorientation by 120° of a phosphohistidine-containing segment. This conformational rearrangement couples the acyl-CoA binding site with the nucleoside diphosphate binding site. The presented structures prove a long-standing hypothesis and provide insight into the determinants for substrate selectivity.

Author contributions: R.H.-J.W., P.S., and A.J.S. designed research; R.H.-J.W. and M.S. performed research; R.H.-J.W., A.F., M.S., P.S., and A.J.S. analyzed data; and R.H.-J.W., M.S., P.S., and A.J.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4XYL, 4XYM, 4XZ3, 4Y8V, 4YAK, 4YAJ, 4YB8, 4YBZ, and 5HBR).

[1]To whom correspondence may be addressed. Email: peter.schoenheit@ifam.uni-kiel.de or axel.scheidig@strubio.uni-kiel.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518614113/-/DCSupplemental.

the subdomains in ACD and, as a consequence, a variation of the proposed reaction steps has to be considered.

ACDs are versatile enzymes that are capable of metabolizing a variety of CoA thioesters generated in the course of sugar and peptide fermentation (3, 4, 21). In contrast, the *ec*SCS specifically binds succinyl-CoA, accepting only small aliphatic CoA thioesters as additional substrates (22). The ACLY is even more specific for Ac-CoA (23). Because ACDs are involved in sugar and amino acid metabolism, they are ambivalent regarding their converted substrates (3, 4, 24, 25). In addition, domain shuffling creates a very diverse pattern of subdomain organization within members of the ACD family (8). Some ACD enzymes are even built up as single-chain proteins with fused alpha-beta or beta-alpha subunits. The linker region between the fused subunits is usually very short (8). Therefore, a structural model for ACDs based on the structure of *ec*SCS as presented by Bräsen et al. (20) cannot be fully compatible with those single-chain ACDs, because distances of more than 60 Å between the termini of individual alpha and beta subunits must be bridged.

So far, there are no structural data to explain these described peculiarities (necessity of a second His, different arrangement of the subdomains, and broader substrate selectivity). In this study, we present a comprehensive analysis of various crystal structures of the functional complex of the ACD isoform 1 from the hyperthermophilic archaeon *Candidatus Korarchaeum cryptofilum* (*ckc*ACD1). *Ca. K. cryptofilum* belongs to the most ancestral archaea (26). Based on genome analyses, *Ca. K. cryptofilum* has been proposed to ferment amino acids using ACDs as major energy-conserving enzymes (26). Three ACD isoenzymes with different substrate specificities have been characterized (27). *ckc*ACD1 converts small aliphatic CoA thioesters. It is composed of four protein chains (two alpha subunits and two beta subunits) that form an $\alpha_2\beta_2$-heterotetramer. Each alpha subunit has one active site for acyl-CoA binding, and each beta subunit has one active site for NDP binding. Due to its similar molecular composition and kinetic properties as ACD1 from the hyperthermophilic archaeon *P. furiosus* (*pf*ACD1), it is likely that *ckc*ACD1 follows the same four-step catalytic reaction mechanism as has been proposed for the homologous *pf*ACD1 (20). In particular, the structural rearrangements accompanying the phosphate shuttle process between alpha and beta subunits of ACD will be described for the first time to our knowledge. This work provides direct evidence for the proposed loop swinging, which, in addition, can explain the necessity for a second active-site His located in the

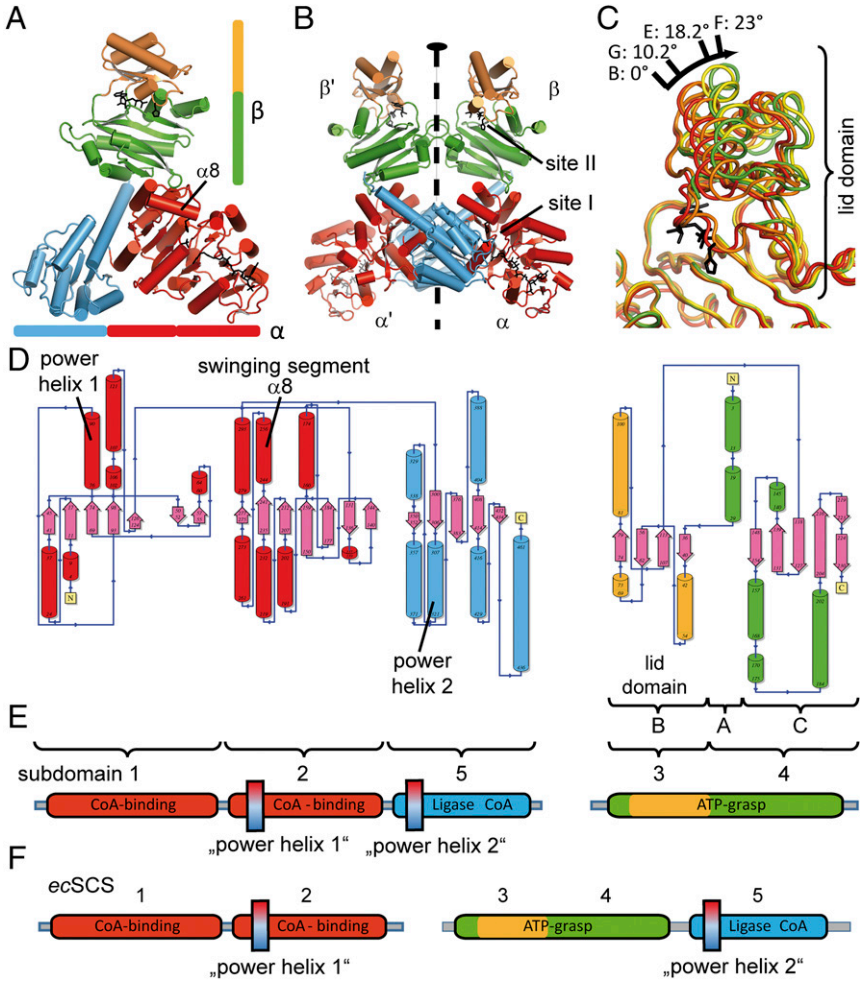

Fig. 1. Overall structure and domain arrangement of *ckc*ACD1. Ribbon representation of the protomer composed of one alpha subunit and one beta subunit (*A*) and the complete $\alpha_2\beta_2$-heterotetramer (*B*). The twofold symmetry is indicated with a black oval. (*C*) Conformational differences of the beta subunit in various crystal structures (*ckc*ACD1-B, *ckc*ACD1-E, *ckc*ACD1-F, and *ckc*ACD1-G) (Table 1). The relative movement of subdomain 3 compared with subdomain 4 is noted in degrees ("opening angle"). Because motion of subdomain 3 is required for catalysis, this region is also called the "lid" (32). (*D*) Topology diagram generated with the program PDBsum (51). The colors of the helixes are in accordance with the color scheme used for the individual subdomains as outlined in *E*. (*F*) Domain arrangement for the heterodimeric *ec*SCS (12). The positions of the two power helices are indicated.

beta subunit near site II. Further detailed analysis of the binding mode of Ac-CoA to *ckc*ACD1 provides insight into substrate specificity of the enzyme.

Results

The 3D assembly of the two alpha and two beta subunits of the *ckc*ACD1 complex revealed a unique $\alpha_2\beta_2$-heterotetramer (Fig. 1). The protomer is composed of one alpha subunit and one beta subunit (Fig. 1 *A* and *E* and *SI Appendix*, Fig. S1). Two protomers are arranged with a twofold symmetry (Fig. 1*B*). The alpha subunit comprises three subdomains (1-2-5), and the beta subunit comprises two subdomains (3-4). The numbering is according to the nomenclature established for *ec*SCS (12, 19, 20) (Fig. 1*F*). All subdomains have an α/β secondary structure topology. With an interface of ~2,000 Å^2 between both alpha subunits, 740–770 Å^2 between alpha and beta subunits, and 540 Å^2 between both beta subunits, the assessment of the *ckc*ACD1 complex was assigned as stable based on the free energy of assembly dissociation (ΔG^{diss}) as calculated by the Proteins, Interfaces, Structures and Assemblies (PISA) server (28, 29).

In the alpha subunit, the N-terminal subdomain (subdomain 1: Met1α–Phe131α) is responsible for CoA binding and contains a Rossmann-like fold (30), which is succeeded by two domains with a flavodoxin-like fold (31). The middle domain (subdomain 2: Gly132α–Asn293α) contains the conserved His (His254α), which is phosphorylated during the first part of the enzymatic reaction cycle. His254α is located within an array rich in small amino acids, providing this region an extraordinary mobility. The C-terminal domain of the alpha subunit (subdomain 5) consists of the amino acids Arg300α–Arg464α. According to the subdomain nomenclature for SCS and ACD enzymes, the beta subunit of *ckc*ACD1 can be described as being built up of two different structural domains, subdomains 3 and 4 (8). Alternatively, the structure of the beta subunit can be referred to as an ATP-grasp fold, consisting of three subdomains A, B, and C (32). In *ckc*ACD1 the subdomain A (Met1β–Pro33β) consists only of two helices and interacts tightly with subdomain C (Gly117β–Arg230β). Considering their similar atomic displacement parameters, both subdomains A and C can be regarded as one structural unit (subdomain 4). The substrate binding site (site II) is usually formed by the assembled subdomains A and C (32). In the context of ACD, this substrate is the phosphate-carrying moiety, as will be described later on. Residues Thr35β–Phe113β form subdomain B (subdomain 3). It harbors the conserved phosphate-binding T-loop (33) (residues Lys60β–Val75β), as well as a highly conserved HK(S/T)(D/E) motif (residues His68β–Asp71β) found in nearly all archaeal and bacterial ACDs (8). Interaction of subdomain 3 and subdomain 4 creates the nucleotide binding site.

ATP-grasp enzymes are known to carry out opening and closure motions to allow binding of the nucleotide ("open") and to orient residues involved in catalysis properly ("closed") (34). In the crystal structures of *ckc*ACD1, the beta subunit was found in different states of "opening" (Fig. 1*C*). The most closed state was observed with bound Ca^{2+}/ado-5′-β,γ-methylene triphosphate (AMPPCP; *ckc*ACD1-B), and the most opened conformation was found in the structure *ckc*ACD1-I with a rotation angle of nearly 25° (Fig. 1*C* and *SI Appendix*, Table S2).

***ckc*ACD1 Binds its Acyl-CoA Substrates Comparable to Other Members of the ACD Superfamily: Binding Mode for CoA.** To understand the functionality of the *ckc*ACD1 complex, we prepared crystals of the enzyme in complex with its substrates CoA, Ac-CoA, phosphate, and ADP, as well as the nonhydrolyzable ADP analog ado-5′-α,β-methylene diphosphate (AMPCP) and the nonhydrolyzable ATP analog AMPPCP. In all cases, CoA was bound to subdomain 1 in an elongated form, pointing its sulfur atom toward the site of the His phosphorylation (His254α) (Figs. 2 and 3 and *SI Appendix*, Fig. S2), in accordance with *ec*SCS referred to as site I. The Ade moiety is inserted into a hydrophobic cleft, and its position is additionally adjusted by two water molecules interacting with either the N6A or N1A and the carbonyl oxygen atoms of the residues Lys79α or Pro59α, respectively (*SI Appendix*, Fig. S2*E*). Upon CoA binding the highly flexible side chain of Lys24α becomes fixed via an interaction of its side-chain amino group with the oxygen atoms of the 3′ and 5′ alpha-phosphate groups (*SI Appendix*, Fig. S2*D*). Interestingly, the *Sus scrofa* SCS (*ss*SCS) features the Lys residue Lys26α at a structurally equivalent position. The side chain of this residue, however, is not involved in CoA binding (35).

Binding Mode of the Ac Moiety Within Ac-CoA. In one structure of *ckc*ACD1 (*ckc*ACD1-E), cocrystallized with Ac-CoA, the electron

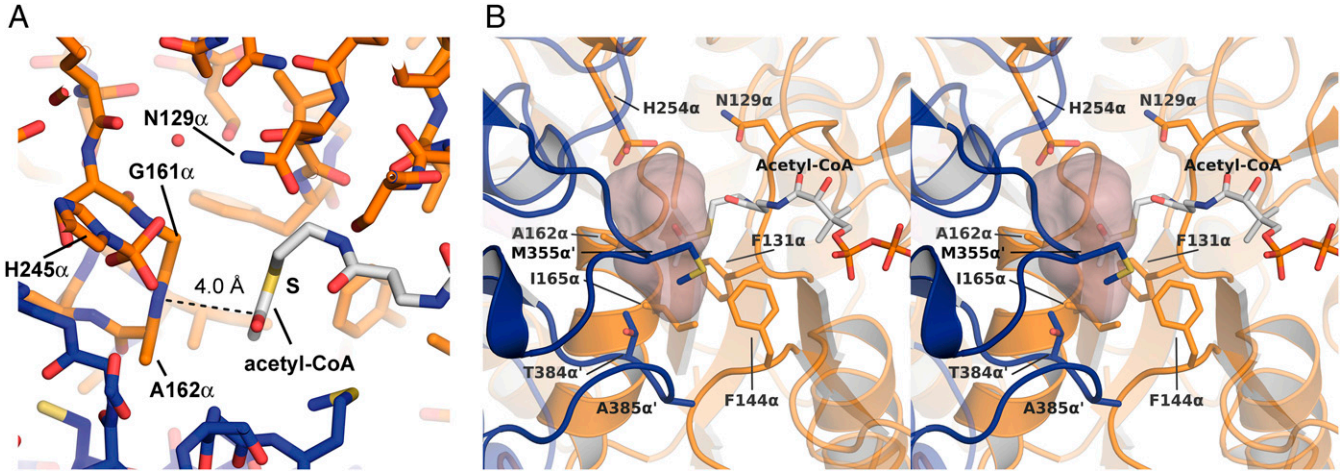

Fig. 2. Binding site for the activated acyl moiety. (*A*) Orientation of the Ac group of Ac-CoA. A dashed line indicates its distance of 4.0 Å to the peptide bond between residues Gly161α and Ala162α. (*B*) Stereo representations of the estimated binding pocket for larger acyl groups. The space available for the acyl moiety was calculated with the program HOLLOW (52) using the crystal structure *ckc*ACD1-E, a sphere radius of 1.4 Å, and CoA derived from Ac-CoA. Amino acid residues shaping the potential binding pocket are represented as sticks and are labeled. Carbon atoms in Ac-CoA are shown in white, nitrogen in blue, oxygen in red, and sulfur in yellow. The carbon atoms of the surrounding protein are colored in orange (alpha subunit) and blue (alpha subunit of the symmetry-related protomer), respectively.

density allowed us to model the Ac group connected to CoA (*SI Appendix*, Fig. S2*B*). Thus, for the first time to our knowledge, an ACD with its CoA thioester substrate bound can be described. The Ac moiety appears to be nearly coplanar to the peptide bond between Gly161α and Ala162α, with a distance of 4.0 Å between the nitrogen atom of Gly161α and the carboxyl carbon of the Ac group (Fig. 2*A*). This distance is in the range of typical π-π interactions (36). Based on the chemical environment, the Ac group was arranged such that the methyl group was positioned to be accommodated by a more hydrophobic environment built up from the side chains of the residues Phe131α, Phe144α, Ala162α, Ile165α, Met355α, and Thr384α′ (the prime symbol indicates the "other" subunit) (Fig. 2*B* and *SI Appendix*, Fig. S2*C*). The phosphohistidine pocket at site I also comprises a highly conserved Asn residue (Asn173α). In the case of Ac-CoA, the distance to the thiocarboxyl carbon at 6.6 Å is rather long; however, due to its position, it might be involved in the cleavage of the thioester bond by inorganic phosphate.

Binding Mode of Phosphate Within Site I of the Alpha Subunit. Crystals of *ckc*ACD1 soaked with 50 mM Na_2HPO_4 revealed the phosphate ion positioned between both alpha subunits at the amino termini of the so-called "power helices" α5 and α11′ (*ckc*ACD1-G and *ckc*ACD1-I; Fig. 3*A*). Interestingly, interactions of the phosphate anion with the protein were mainly facilitated by the amino functions of the peptide bonds of four residues: Gly161α and Ala162α from one alpha subunit and Gly308α′ and Gly309α′ from the other alpha subunit. In addition, a strong hydrogen bond was formed to the side-chain hydroxyl group of Ser160α. The distance between the carbonyl carbon of the Ac group of Ac-CoA and the closest oxygen atom of the phosphate is ∼6 Å. The electron density surrounding the phosphate was interpreted as two metal ions. In *ec*SCS and *ss*SCS, phosphate ions are located in a very similar position (14, 16). As for the described SCS, the N termini of the two equivalent power helices participate in the compensation of the negative charge of the phosphate ion; however, no additional metal ions have been observed in SCS.

Environment of the Phosphorylated His Residue His254α. We obtained crystals of the *ckc*ACD1 complex with phosphorylated His254α (*ckc*ACD1-E and *ckc*ACD1-H; Fig. 3*B*). The conformation of the phosphorylated His was stabilized by ionic interaction with the side-chain carboxyl group of Glu213α. A hydrogen bridge (distance of 2.7 Å) provides an ideal geometry for the phosphorylation reaction by favoring the protonation state of the imidazole group. In addition, the phosphoryl moiety forms a hydrogen bond to the hydroxyl group of Ser160α. In the crystal structure *ckc*ACD1-H, the phosphate moiety also interacts with a bound metal ion, which is coordinated by the side chain of Asp351α′. The interaction with the metal ion results in a further movement of the phosphohistidine and adjoining residues toward site I, and subsequent stabilization of the phosphorylation of His254α by forming several hydrogen bridges to the amino groups of the peptide bonds between the amino acids Ser160α–Gly161α, Gly161α–Ala162α, Gly307α′–Gly308α′, and Gly308α′–Gly309α′. Interestingly, the residues Ser160α and Asp351α′ are highly conserved throughout the ACD superfamily. Only Ser160α is sometimes exchanged to Thr, an amino acid with comparable side-chain characteristics.

Binding Mode of Ado Nucleotides Within Site II Located in the Beta Subunit. The binding site for the nucleotide substrates of *ckc*ACD1 is within the beta subunit (ATP-grasp domain). In *ckc*ACD1, the Ade moiety is sandwiched in a hydrophobic environment created by residues of subdomain 3 and subdomain 4, with additional hydrogen bridges to the side chains of Gln111β and Lys60β, as well as the protein backbone (*SI Appendix*, Fig. S3). Furthermore, Lys60β, a residue conserved in nearly all ACDs,

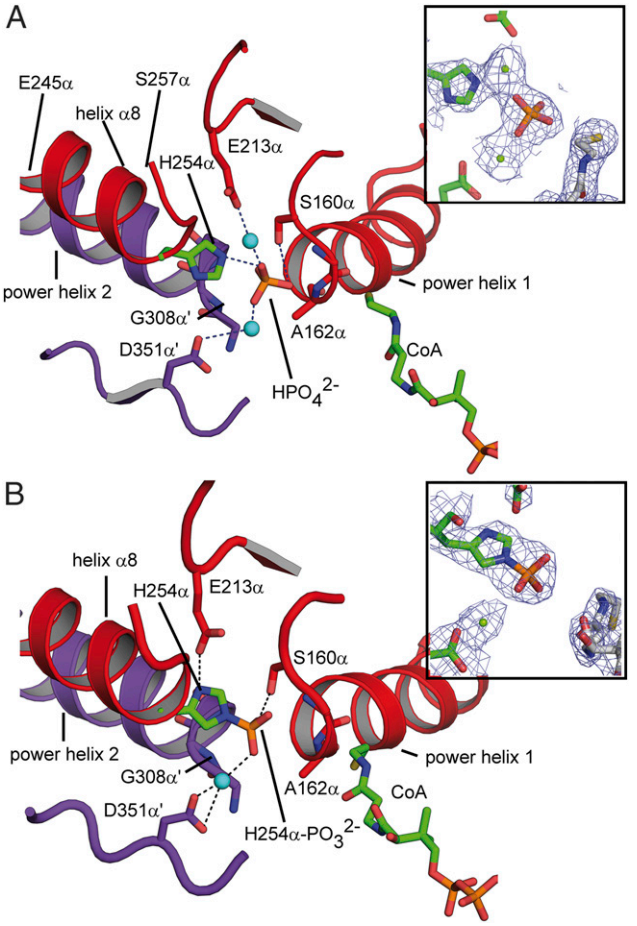

Fig. 3. Phosphate binding site of *ckc*ACD1. Environment of bound inorganic phosphate (*A*) and of the phosphohistidine residue His254α-P (*B*) within site I of the alpha subunit. The two so-called power helices (*ckc*ACD1 162α–173α/308α′–319α′) and helix α8 of the phosphohistidine segment are displayed in ribbon representation, and the phosphate group and His254α, as well as interacting residues, are represented as sticks. Carbon atoms of His254α and the CoA are colored in green, and other carbon atoms of the surrounding protein are colored in red (alpha subunit, subdomain 2) and blue (alpha subunit of the symmetry-related protomer, subdomain 5), respectively. Magnesium ions are shown as cyan spheres. (*Insets*) Electron density distribution around the phosphate group with a contouring level of 1σ (electron density distribution is shown as blue mesh; carbon atoms of protein are shown in green, carbon atoms of CoA in white, nitrogen atoms in blue, oxygen atoms in red, sulfur atoms in yellow, and phosphor atoms in orange).

interacts with the alpha-phosphate group of ADP. A dual role of this particular Lys residue was also observed for the corresponding Lys46β of *ec*SCS (14). Furthermore, the phosphate groups of ADP in the crystal structures *ckc*ACD1-B, *ckc*ACD1-D, *ckc*ACD1-G, and *ckc*ACD1-H are in contact with the side-chain amine of Lys69β. Additionally, several interactions facilitate tight binding of the beta-phosphate moiety of the nucleotide to the protein. Hydrogen bridges were observed between the phosphate and the T-loop residues involving the side-chain hydroxyl group of Ser70β and the backbone nitrogens of Lys69 and Ser70β, respectively (*SI Appendix*, Fig. S3*E*). A magnesium ion was complexed by both alpha- and beta-phosphates in the structures *ckc*ACD1-D and *ckc*ACD1-G. In the case of the *ckc*ACD1 structure with a closed ATP-grasp conformation (*ckc*ACD1-B), a calcium ion is complexed by oxygen atoms of both alpha-phosphate and beta-phosphate, as well as the side-chain carboxyl group of

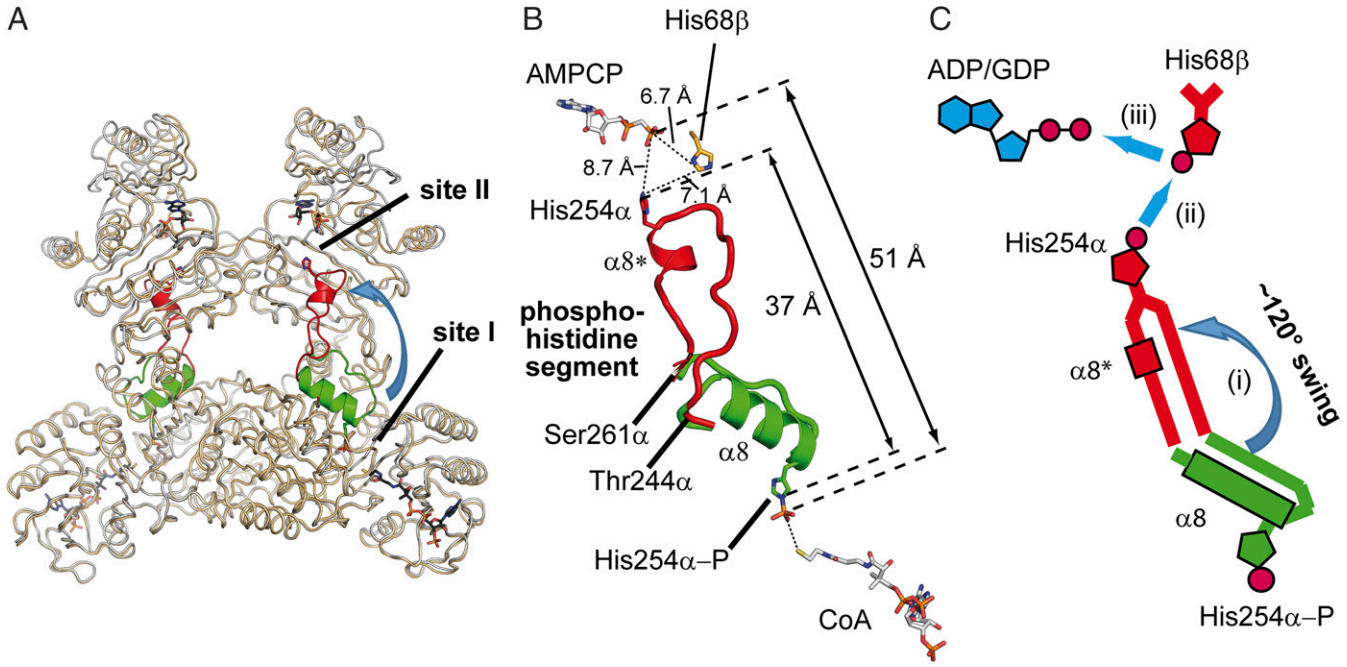

Fig. 4. Rearrangement of the His254α-P containing segment upon phosphorylation. (*A*) Composition representation of two *ckc*ACD1 structures to visualize the movement (blue arrow) of the phosphohistidine segment from site I to site II. Superimposed are the crystal structures *ckc*ACD1-H (gray ribbons) and *ckc*ACD1-C (orange ribbons). The phosphohistidine segment is colored in green if oriented toward site I (structure *ckc*ACD1-H) and in red if oriented toward site II (*ckc*ACD1-C). (*B*) Distances between site I and site II to highlight the gap that the activated phosphoryl moiety has to travel when attached to His254α of the phosphohistidine segment. Carbon atoms of CoA and AMPCP are colored in white, nitrogen atoms in blue, oxygen atoms in red, phosphor atoms in orange, and sulfur atoms in yellow. (*C*) Schematic representation of the reorientation of the phosphohistidine segment. After phosphorylation of His254α the phosphoryl moiety is transferred to the nucleotide binding site located in the beta subunit. For this step, reorientation and reorganization of the phosphohistine segment are needed (*i*) for transferring the phosphoryl moiety to His68β-forming phospho-His68β (*ii*), which, in turn, phosphorylates bound ADP/GDP (*iii*).

Asp224β. All mentioned residues belong to the key fingerprint motif of ATP-grasp enzymes (32). The beta-phosphate additionally interacts with the guanidinyl group of Arg226β. However, this interaction is not observed for all beta subunits, which suggests that the side chain of Arg226β is flexible.

Due to the strong interactions of the T-loop with the phosphate groups of ADP, the opening and closing of the lid domain causes a stretching of the bound nucleotide (*SI Appendix*, Fig. S3*D*). Therefore, the movement of the lid domain appears to be involved in moving in and pulling out of the nucleotide substrate from the ATP-grasp domain. During this process, the magnesium ion remains complexed to the phosphate groups and is withdrawn from an interaction with the side-chain carboxyl group of Asp224β.

We also obtained one crystal structure with the nonhydrolyzable ATP analog AMPPCP (*ckc*ACD1-C). In contrast to ADP, the Ade and sugar moiety of AMPPCP were shifted by 1–2 Å out of the nucleotide binding pocket. The beta-phosphate is rotated toward the surrounding solvent, with a concomitant change of the dihedral angle C5′-O5′-PA-O3A of ∼130° (*SI Appendix*, Fig. S3 *C* and *E*). Interestingly, the angle between the three phosphor atoms constitutes nearly 90°. A magnesium ion is complexed by oxygen atoms of all three phosphate groups, as well as the carboxyl group of Asp224β. There is also an interaction of the gamma-phosphate with the guanidinyl group of Arg226β. At the corresponding sequence position, other members of the ACD family feature either Arg or Lys residues. In the case of GTP bound to *ss*SCS, the amino function of the side chain of Lys222β interacts with the gamma-phosphate in a similar manner (16), which might indicate a specific role for this residue.

***ckc*ACD1 Bridges a 51-Å Gap During its Enzymatic Cycle.** The beta subunit harbors two binding sites, one for the nucleotide and one for the activated phosphate group. Whereas the catalytic site within the alpha subunit is referred to as site I (binding of CoA and inorganic phosphate), the catalytic site within the beta subunit (binding pocket for NDP and activated phosphate) is referred to as site II (13). The distance between these two sites is very large. In *ckc*ACD1, the space between the phosphor atom of the phosphate ion (bound to His254α) and the beta-phosphate of AMPCP (bound within the beta subunit) constitutes around 51 Å. Therefore, the transfer of the phosphoryl moiety between the alpha and beta subunits requires large rearrangements. For *ec*SCS, a swinging loop mechanism for the phosphohistidine-containing loop has been proposed (14). Here, we present, for the first time to our knowledge, detailed structural data of the "phosphohistidine segment" pointing toward the nucleotide-binding site of the beta subunit (crystal structure *ckc*ACD1-C) (Fig. 4). This segment consists of the amino acids Gly242α to Val262α of the alpha subunit, and also contains the His residue His254α, which is phosphorylated during the enzymatic cycle (*SI Appendix*, Fig. S1). These residues undergo a major rearrangement in space, as well as in secondary structure, during movement from site I to site II (Fig. 4). Hence, we henceforth use the term phosphohistidine segment instead of the usual expression, swinging loop. This segment in its site I orientation is composed of an α-helix from residue Glu245α to Thr255α (helix α8), directing His254α toward the CoA binding site (Fig. 3). The amino acid residues Gly256α to Gly260α act as a flexible loop connecting the helix α8 with the subsequent helix α9. By passaging the (phosphorylated) His254α to the beta subunit of *ckc*ACD1, the phosphohistidine segment loses most of its helical structure. Accompanying the partial unwinding of helix α8, the residues Thr244α to Thr255α are positioned in a nearly straight array bridging a distance of more than 10 Å. The upstream residues are assembled into a new short helix α8* (the asterisk indicates the rearrangement of helix α8) formed by the amino acids His254α–Ile258α. After the rearrangement of this segment, the τ-nitrogen of

His254α travels around 37 Å between sites I and II (Fig. 4*B*). Interaction of the phosphohistidine segment with the beta domain also causes substantial changes in nearby residues. The formation of a hydrophobic pocket comprising residues Ile258α, Ala259α, Ile121β, Met135β, Val143β, Val149β, and Phe151β results in partial unwinding and reorientation of helix α5 (residues Gly140β–Phe146β), which, in turn, alters the arrangement of the hydrophobic interface between both beta domains (*SI Appendix*, Fig. S4*A*). As a further consequence, the residues between α6 and α7 are shifted in their position. The largest movement observed is 5.9 Å (maximal for Gly179α′) (*SI Appendix*, Fig. S4*B*).

Discussion

Substrate Specificity. The *ec*SCS has a high specificity for succinyl-CoA (22) but converts small aliphatic CoA esters (22) like *ckc*ACD1 and other ACDs (2, 3, 5–7, 20, 21, 24, 37–39) as well. Interestingly, members of the ACD family are also able to transform aromatic CoA thioesters (3, 24, 39), although catalytic activity toward succinyl-CoA has only been described a few times (3, 25, 40). The high specificity of SCS enzymes for succinyl-CoA does make a lot of sense because this enzyme family is a central and, as such, an optimized element within the TCA cycle. In contrast, *ckc*ACD1 and other members of this family are not integrated into a well-tuned metabolic pathway but, instead, have to conserve energy [e.g., during glycolysis or the catabolic degradation of amino acids (3, 7, 20, 24)]. In this respect, a broader substrate spectrum of ACDs appears plausible.

***ckc*ACD1 and *ec*SCS Display Very Similar Features for the Catalytic Active Site I.** In the crystal structure of *ckc*ACD1-E, Ac-CoA was bound to the enzyme. However, the resolution of this crystal was not good enough to allow the differentiation between its methyl and carbonyl moieties unambiguously. Therefore, we compared the binding pocket for the Ac moiety within *ckc*ADC1 with the putative succinyl binding pocket of *ec*SCS to understand how the aliphatic moiety of the acid component within acyl-CoA should be positioned in ACDs. In both, *ckc*ACD1 and SCS, the available space is large enough to harbor even bulkier substrates. The capacious volume is somewhat surprising, given the limited spectrum of transformed substances. In the case of *ckc*ACD1, there is a small hydrophobic pocket next to the sulfhydryl group of CoA, which can be used to explain the substrate specificity of *ckc*ACD1 (Fig. 2*B*). The size of this pocket is restricted by the side chains of the amino acids Ala162α, Ile165α, Phe144α, Met355α′, Thr384α′, and Ala385α′. Due to this limited size, this pocket permits only short-chained acyl-CoA thioesters to fit. The same region in *ec*SCS features a larger pocket surrounded by polar amino acids, which could accept the carboxy group of succinate of succinyl-CoA (*SI Appendix*, Fig. S5). The enzyme ACLY converts Ac-CoA (23). A comparable mode of interaction of the Ac group of Ac-CoA interacting with ACLY or *ckc*ACD1, respectively, might be assumed. Hence, we superimposed *ckc*ACD1 onto the crystal structure of human ACLY with bound citrate (17). The resulting distance between the carbonyl carbon of Ac-CoA of *ckc*ACD1 and the carbon C1 of citrate in ACLY is ∼2.8 Å. Thus, we expect the binding site for the aliphatic moiety of acyl-CoA for ACDs in a similar location as observed in ACLY. Based on this comparison, the orientation of the Ac group of Ac-CoA is as shown in Fig. 2*A*. The thioate function of succinyl-CoA bound to *ec*SCS might, by analogy with *ckc*ACD1, be situated coplanar to the peptide bond between the corresponding amino acids Cys123α and Pro124α. Interestingly, these residues were significantly distorted in the crystal structure of an *ec*SCS mutant with absent activity (13). The Cys residue itself, however, was shown not to be part of the catalytic mechanism (41). Furthermore, in the same study, it was shown that variations of the size of the side chain at the position of Cys123α deteriorate the enzyme activity. It appears that slight differences in the orientation of the Cys-Pro peptide bond may result in misalignment of the thioester moiety, and therefore a decrease in the enzymatic function.

Site I, Acyl-CoA Binding Site. The CoA binding site I is situated at a highly conserved interface formed by residues of subdomains 1, 2, and 5. These subdomains are packed in a similar way in SCS, ACLY, and *ckc*ACD1. A sequence alignment of several ACD proteins revealed that the second ligase-CoA domain (subdomain 5) has only moderate conservation in its primary structure (8), which is in strong contrast to the situation observed for SCS enzymes. Most interestingly, the sequence alignment revealed that *ckc*ACD1 and *pf*ACD1 not only share a similar substrate spectrum but also feature identical amino acids lining the potential substrate pocket (50% overall sequence identity). Those residues were dissimilar in other *P. furiosus* isoenzymes (35–42% overall sequence identity) with a different substrate spectrum (*SI Appendix*, Fig. S6). Hence, we propose that alterations of this cleft are responsible for different substrate specificity.

Site II, Purine Nucleotide Specificity. ACD and SCS enzymes are able to bind and convert all purine nucleotides at site II, but some enzymes exhibit specificity for either the Ado or Guo nucleotide (2, 21, 24, 38, 40). Kinetic characterization of *ckc*ACD1 revealed that both nucleotides are accepted as substrates (*SI Appendix*, Fig. S7 *A* and *B*). In this work, we crystallized the *ckc*ACD1 complex solely with the Ado nucleotide ADP and its nonhydrolyzable analog AMPCP, as well as with the nonhydrolyzable ATP analog AMPPCP. The purine moiety is situated in a hydrophobic environment maintaining hydrogen bonds to the backbone nitrogen of Ala114β, the backbone oxygen of Glu112β, the side-chain carbonyl group of Gln111β, and the side-chain amino moiety of Lys60β. To gain insight into the potential binding mode of Gua nucleotides, we superimposed *ckc*ACD1 with the ATP-grasp domain containing enzyme CK2 from *Zea mays*, which is known to metabolize Ade and Gua nucleotides (42–44). We found that (*i*) the Ade moiety of ATP bound to CK2 [Protein Data Bank (PDB) ID code 1DAW] is oriented comparable to *ckc*ACD1 and (*ii*) the Gua moiety of GTP bound to CK2 (PDB ID code 1DAY) is shifted in its position and compares quite well with the binding mode found in the structure of GTP-specific *ss*SCS (16). Hence, we expect a comparable binding mode for Guo nucleotides in *ckc*ACD1.

Consequences of Different Domain Arrangements. The signature motifs for the superfamily of ACDs (NDP-forming) have been extensively studied by comparison of the sequences of bacterial, archaeal, and eukaryotic members (8). The same study highlights the different domain arrangements and connectivities observed in the various kingdoms of life. Based on the structure of the *ec*SCS and sequence comparison with related enzymes, it can be derived that a functional ACD is composed of five subdomains (subdomains 1–5), with *ec*SCS displaying the arrangement [alpha(1-2)/beta(3-4-5)]. In the context of domain shuffling, it has to be considered that the respective subdomains 3 and 4 should not be regarded as individual domains. The sequence region for these two subdomains is best described as an ATP-grasp domain, defining the binding site for the NDP and the binding site for the phosphate-carrying substrate (site II). In case of ACLY, the subdomain composition within a single protein chain is 3-4-5-1-2-CS with an additional domain (CS) fused to its carboxyl terminus. Due to the domain shuffling, the overall arrangement of the subdomains is significantly different between *ec*SCS and *ckc*ACD1 (Fig. 1 and *SI Appendix*, Fig. S1). The existence of ACD as a $\alpha_2\beta_2$-heterotetramer is a direct consequence of its specific domain arrangement [alpha(1-2-5)/beta(3-4)]. In contrast, SCS enzymes with the domain arrangement alpha(1-2)/beta(3-4-5) have

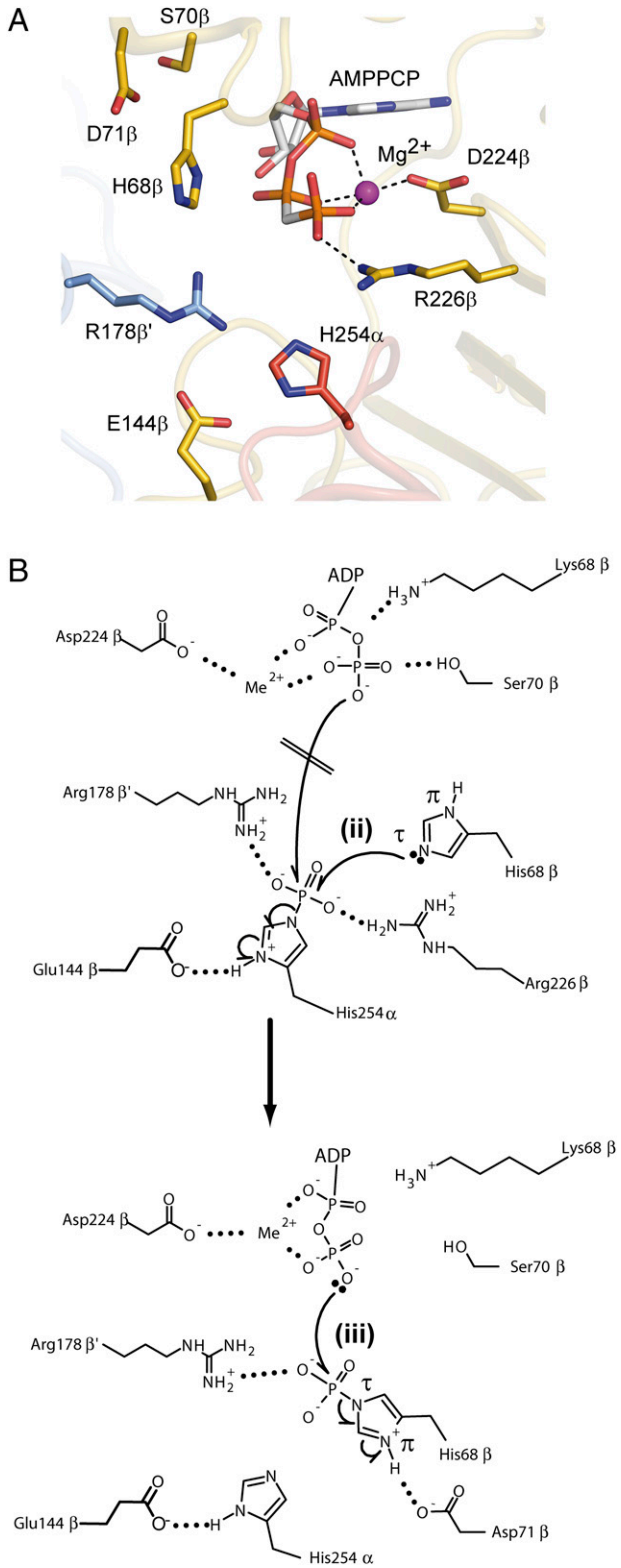

Fig. 5. Rearrangement of the phosphohistidine segment and proposed involvement of His68β for phosphorylation of ADP. (*A*) Representation of the region around the phosphate-carrying His residue His254α pointing toward site II within the beta subunit (ATP-grasp domain) after delivery of the phosphate moiety (crystal structure *ckc*ACD1-C). Residues involved in reaction steps *ii* and *iii* in Fig. 4*C* are labeled. Side-chain atoms are shown as carbon in magenta (alpha subunit), carbon in yellow (beta subunit), carbon in light blue (beta′ subunit), nitrogen in blue, and oxygen in red. Carbon atoms in AMPPCP are shown in white, nitrogen in blue, oxygen in red, and phosphor in orange. (*B*) Scheme for the transfer of the phosphoryl moiety from His254α to ADP via His68β within site II. Phosphorylated His254α is positioned in an ideal orientation for transfer of the phosphoryl moiety to His68β (reaction step *ii* in Fig. 4*C*). The negative charge from the phosphoryl moiety is neutralized by interaction with the guanidinyl group of Arg226β and Arg178β′. This makes phospho-His254α more susceptible to nucleophilic attack by His68β. Only the nitrogen atom N^{ε} is in the correct position for an in-line attack on the phosphor atom. Phosphorylation of His68β must result in a displacement of the beta-phosphate group of ADP to bring one of its oxygen atoms into the optimal position for nucleophilic attack (reaction step *iii* in Fig. 4*C*). This reorientation needs rotation of the bond O5′-PA of ADP, as well as rotation of phospho-His68β toward the nucleotide.

representatives with an $\alpha_2\beta_2$-heterotetrameric structure, as well as with αβ-heterodimeric structures, as observed for the SCS from *E. coli* (19), *T. aquaticus* (15), pig (16), and human (45). For *ec*SCS, two very important α-helices have been highlighted and termed power helices because they stabilize the phosphoryl moiety within the catalytic site I by their helix dipole moment (19). Whereas in *ec*SCS, the two power helices of one active site I are contributed from one alpha subunit and one beta subunit, in *ckc*ACD1, the two power helices for each active site I are positioned on two different alpha subunits. To form a functional unit in *ckc*ACD1, two alpha subunits have to come together, thus generating a heterotetramer, whereas a heterodimer is already the minimal composition for an active SCS (*SI Appendix*, Fig. S8).

As a further consequence of the domain arrangement, the distance between the phosphorylated His residue His254α and ADP as the final acceptor of the phosphoryl moiety bound to the beta subunit is significantly increased from around 35 Å (*ec*SCS) (*SI Appendix*, Fig. S9) to around 51 Å as observed in the presented complexes of *ckc*ACD1 (Fig. 4). In one of our structures (*ckc*ACD1-C), we observed an orientation of the phosphohistidine segment with the nonphosphorylated His254α pointing toward site II. This movement of the His254α necessitates a partial unwinding of helix α8 (site I), reorientation of the extended loop, and rewinding to helix α8* (site II) (Fig. 4 *B* and *C*). In this new orientation, His254α is in close proximity to the beta-phosphate group of ADP (~9 Å) and His68β (~7 Å), the second, highly conserved His residue in the ACD subfamily. The distance of His254α to ADP appears to be too large for a direct transfer of the activated phosphate but requires an intermediate mediator. Because His68β was indeed observed to be functionally relevant in an in vitro experiment with an only marginally active *ckc*ACD1 variant (*SI Appendix*, Fig. S7*C*), we postulate that His68β is an essential part of the overall reaction. In accordance to *ckc*ACD1, exchange of the corresponding residue of *pf*ACD1 results in inactivation of the enzyme (20). Thus, we propose an additional, fourth reaction step for the overall reaction with phosphorylated His68β as an intermediate.

Postulated Reaction Steps. Based on our various crystal structures and supplemental biochemical studies, we are able to explain the necessity of the additional fourth reaction step as suggested by Bräsen et al. (20) in comparison to the three-step reaction mechanism as formulated for SCS enzymes (12, 13, 19, 46). The first part of the overall reaction starts at active site I with the binding of Ac-CoA (structure *ckc*ACD1-E) (Fig. 2*A*) and inorganic phosphate (structure *ckc*ACD1-G and *ckc*ACD1-I) (Fig. 3*A*). In the first reaction step, the cleavage of Ac-CoA into Ac-phosphate and CoA is coupled to the phosphorylation of His254α (structures *ckc*ACD1-E and *ckc*ACD1-H) (Fig. 3*B*) in the second step. These first two reaction steps are similar to those steps in *pf*ACD-I and *ec*SCS (20). The activated phosphoryl moiety is transferred onto the NDP bound at the active site II. To facilitate this transfer, Bailey et al. (47) and Joyce et al. (14) postulated that the loop containing the phosphorylated His swings

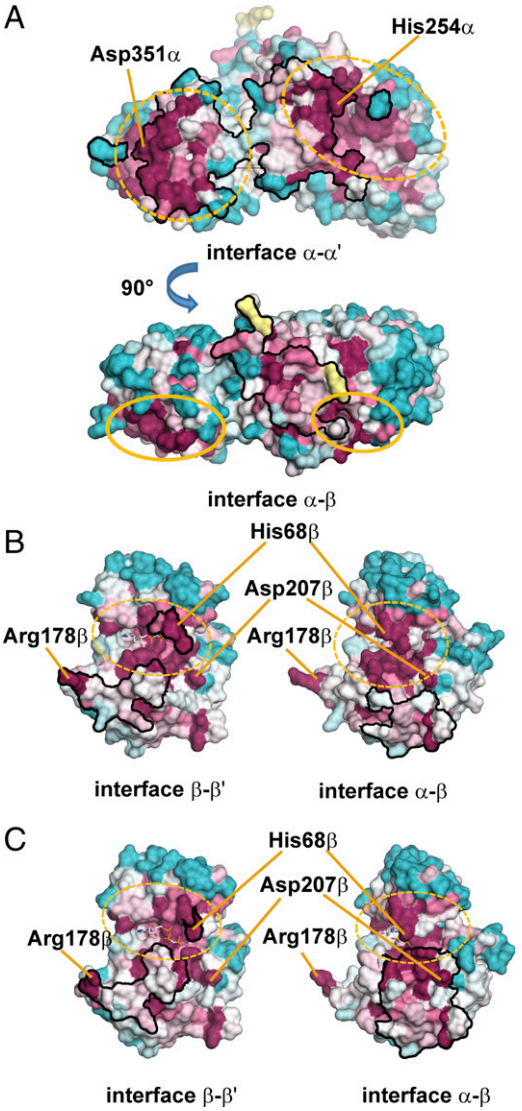

Fig. 6. Conservation pattern for surface residues of the alpha and beta subunits of *ckc*ACD1. Sequence alignment and calculation of the level of conservation were performed with the program ConSurf (48). Residues at the protein surface are colored according to their level of conservation (blue and light blue, less conserved; white, residues of average conservation; violet and lilac, highly conserved; yellow, insufficient data). Framed black are regions that are involved in interface formation (type of interface is indicated below the corresponding picture). (*A*) Interface between the two alpha subunits (derived from the crystal structure of *ckc*ACD1-G). (*Upper*) Side view of the interface between both alpha subunits. Ellipses in yellow are used to depict both parts of site I (N-terminal part, broken line; C-terminal part, dashed line). (*Lower*) View was rotated by 90° as indicated to show the amino acids forming the interface to the beta domain. (*B* and *C*) Interface formed by the two beta subunits. (*B*) Interface of the beta subunit as observed in the crystal structure *ckc*ACD1-B, with the phosphohistidine segment oriented toward active site I. (*C*) Interface of the beta subunit as observed in the crystal structure *ckc*ACD1-C, with the phosphohistidine segment oriented toward active site II. (*B* and *C*, *Left*) Interface contact area between the two symmetry-related beta subunits. (*B* and *C*, *Right*) Interface contact area between one beta subunit and its corresponding alpha subunit. (*Left* and *Right*) Views are in slightly different orientation.

from site I to site II (*SI Appendix*, Fig. S9). With the structure of *ckc*ACD1-C, we could prove this postulated rearrangement for the first time to our knowledge (Fig. 4). However, in contrast to the three partial reactions proposed for SCS, the ACD enzymes do not directly transfer the phosphoryl moiety from the phosphorylated His of the rearranged loop onto the NDP bound within the ATP-grasp domain [structures *ckc*ACD1-B, *ckc*ACD1-C, *ckc*ACD1-D, *ckc*ACD1-G, and *ckc*ACD1-H (20); *SI Appendix*, Fig. S3]. After the rearrangement of the phosphohistidine-containing segment (step *i* in Fig. 4*C*), the ACD enzymes transfer in an additional step the phosphoryl moiety from phospho-His254α onto His68β to form phospho-His68β (step *ii* in Fig. 4*C*). In the fourth reaction step, the phosphoryl moiety is used to phosphorylate ADP or GDP (step *iii* in Figs. 4*C* and 5). The crucial role of His68β is underlined by the observation that mutation of His68β causes loss of activity (*SI Appendix*, Fig. S7*C*).

We could not trap the phosphorylated His68β in a structure of *ckc*ACD1. Based on the spatial proximity of His254α positioned toward site II and His68β (structure *ckc*ACD1-C) (Fig. 5), we propose a nucleophilic attack by nitrogen N^{ε} (τ-position) of the imidazole group of the His68β via an in-line mechanism. Due to steric reasons, the reaction would not be possible for the N^{δ} atom (π-position). His68β is situated in the T-loop of the ATP-grasp domain. For the final reaction step, the transfer of the phosphoryl moiety from phospho-His68β onto the NDP, a relocation of the diphosphate chain of the nucleotide will be necessary. In the crystal structure of *ckc*ACD1-C with bound AMPPCP at site II, we observed such a conformation for the arrangement of the alpha-phosphate and beta-phosphate moieties of AMPPCP (*SI Appendix*, Fig. S3). In *ec*SCS, Glu197β is thought to mediate the optimal protonation state for His246α to facilitate direct phosphorylation of ADP (12, 14). In *ckc*ACD1, the structurally equivalent residue Asp209β interacts with the main chain nitrogens of His254α and Thr255α and the hydroxyl group of the side chain of Thr255α (Fig. 5 and *SI Appendix*, Fig. S8*C*). Thus, it stabilizes completely the phosphohistidine segment in its optimal position for phosphoryl transfer onto the second His, His68β. In addition, from mutational studies of *pf*ACD1, it is known that beside its charge, the length of this Asp residue (Asp212β in *P. furiosus*) is essential for enzymatic activity (20). On the contrary, the length of the side chain of the corresponding residue in ecSCS, Glu197β, only marginally influences the properties of the enzyme (13). Therefore, the role of *ec*SCS-Glu197β is different compared with the corresponding residue Asp209β in *ckc*ACD1. In the crystal structure of *ckc*ACD1-C, the side chain of the Glu residue Glu144β was found in a hydrogen bridge with N^{δ} of His His254α (*SI Appendix*, Fig. S4*C*). Interestingly, sequences of other ACDs feature a Glu residue corresponding to Glu144β in *ckc*ACD1 (8). We believe that this residue maintains the ideal protonation state of phospho-His254α during the phosphoryl transfer to His68β in *ckc*ACD1 and likely in other ACDs. The proposed mechanism for the unique fourth partial reaction is depicted in Fig. 5*C*.

***ckc*ACD1 Domain Assembly Seems to be Prototypic for ACDs.** Alignment of the *ckc*ACD1 with sequences of several members of the ACD family showed a high degree of conservation for the CoA binding subdomains (subdomains 1 and 2) and the ATP-grasp domain (subdomains 3 and 4). Larger variations are only seen for the ligase-CoA subdomain 5 (8, 20), which is observed for members of the class of two-component [composition alpha(1-2-5)/beta(3-4)], as well as for fusion-type ACDs (composition 1-2-5-3-4 or 3-4-1-2-5). To assess whether the complex structure of *ckc*ACD1 might be archetypical for the family of ACDs, we compared the conservation of residues participating in the formation of interdomain interfaces (i.e., contacts between α-α′, α-β, and β-β′, respectively) with residues not involved in multimerization using the program ConSurf (48). The results are depicted in Fig. 6. Most of the amino acids within the interfaces are conserved, which might be a hint for common epitopes between the alpha-alpha interface (Fig. 6*A*, *Upper*) and the interface between alpha and beta subunits (Fig. 6*A*, *Lower*), as well as the heterotypic interaction to the alpha subunit or the homotypic interaction to the second beta subunit (Fig. 6 *B* and *C*). Based on

Table 1. Data collection and refinement statistics

*ckc*ACD1-#	A	B	C	D	E	F	G	H	I
PDB entry	4XYL	4XYM	4XZ3	4Y8V	4YAK	4YAJ	4YB8	4YBZ	5HBR
Resolution range (Å)	83.46–1.95 (1.98–1.95)	127.04–1.90 (1.93–1.90)	48.84–2.40 (2.47–2.40)	49.21–2.10 (2.14–2.10)	83.57–2.46 (2.53–2.46)	125.81–2.20 (2.25–2.2)	110.61–1.90 (1.93–1.90)	83.03–2.1 (2.14–2.1)	126.03–1.99 (2.03–1.99)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell a, b, c (Å)	102.5, 112.2, 124.9	99.6, 114.4, 127.0	100.2, 111.9, 127.6	106.5, 111.0, 126.7	106.4, 111.7, 126.0	105.8, 111.2, 125.8	106.1, 110.6, 125.7	105.3, 109.9, 126.8	106.1, 110.7, 126.0
Total reflections	684,725	706,193	711,978	548,838	591,856	570,073	1,515,778	573,438	672,786
Multiplicity	6.5 (6.6)	6.3 (2.7)	12.5 (12.5)	6.4 (4.4)	10.9 (6.8)	7.7 (3.1)	13.0 (9.8)	6.7 (6.9)	6.6 (5.9)
Completeness (%)	99.8 (99.7)	97.8 (58.1)	99.9 (98.8)	97.6 (77.9)	98.1 (77.3)	97.3 (57.0)	100.0 (99.5)	99.5 (99.1)	99.6 (93.1)
Mean I/σ(I)	6.7 (0.6)	8.1 (0.3)	9.9 (0.8)	10.1 (1.1)	9.7 (0.6)	6.4 (0.4)	9.4 (0.5)	7.5 (0.8)	8.4 (0.5)
CC(1/2)	0.992 (0.169)	0.996 (0.103)	0.995 (0.229)	0.995 (0.422)	0.996 (0.231)	0.990 (0.105)	0.998 (0.159)	0.993 (0.312)	0.998 (0.124)
Wilson *B*-factor (53) ($Å^2$)	25.5	24.9	40.2	20.4	45.3	30.8	37.1	26.2	37.2
R-work (%)	20.64	19.53	19.65	18.44	19.36	21.45	18.16	21.58	17.84
R-free (%)	24.64	23.71	24.44	22.55	23.82	24.13	22.05	25.34	22.22

Parameters for the outermost shell are shown in parentheses. CC(1/2), percentage of correlation between intensities from random half datasets. Correlation significant at the 0.1 % level (54); mean intensity over sigma, I/σ(I); R-work, $R = 100 \times \Sigma_{hkl} \mid |F_{obs}| - |F_{calc}| \mid / \Sigma_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively; R-free is equivalent to R-work but is calculated from reflections (5%) that were omitted from the refinement process (55, 56).

these results, we propose that the *ckc*ACD1 complex represents the prototypic assembly of at least all heterotetrameric ACDs with an $\alpha_2\beta_2$ composition.

Conclusions

We have presented the first, to our knowledge, crystal structures of a functional nucleoside diphosphate-dependent Ac-CoA synthetase. The protein displays a so far unknown 3D arrangement of the five subdomains, which are characteristic for the underlying enzyme superfamily. The different structures represent sequential snapshots along the four-step reaction path, which starts with Ac-CoA, involves two phosphorylated His residues, and ends with a nucleoside-triphosphate. We could describe in detail the binding pockets for the Ac moiety of Ac-CoA, phosphate, and the first phosphorylated His residue. We describe the binding mode for the nucleoside diphosphate, as well as for the final product, nucleoside triphosphate. These structures provide insight into the determinants of the broad-substrate spectrum observed for *ckc*ACD1. Most interestingly, we could present structural proof for the swinging of a protein segment containing the first phosphorylated His in close proximity to the second phosphorylated His. Our structures provide evidence for the necessity of the second His as a phosphoryl intermediate. Furthermore, we propose that the observed structure is archetypical for this group of heterotetrameric ACD enzymes with alpha(1-2-5)/beta(3-4) composition. Our study explains on the structural level the various differences described for ACD in comparison to SCS, the best-studied member of this enzyme superfamily. Overall, this study provides global information for enzymes composed of multiple subdomains and allows us to understand the transmission of activated substrates (e.g., transfer of phosphoryl group) between two distantly separated active sites within the same enzyme.

Materials and Methods

Protein Expression, Purification, and Crystallization. *Ca. K. cryptofilum* ACD1 was expressed in *E. coli* and purified to homogeneity using heat precipitation and gel filtration. Detailed descriptions are provided in *SI Appendix*. Crystals of full-length *ckc*ACD1 were grown from solutions containing 100 mM Tris·HCl (pH 8.3–8.7) and 14–18% (wt/vol) PEG 6000 at 278 K or 18–22% (wt/vol) PEG 6000 at 291 K, respectively. Supplementation of the solutions with 10–30 mM $MgCl_2$ or $CaCl_2$ resulted in better crystal growth. Crystals of *ckc*ACD1 were indexed as orthorhombic space group $P2_12_12_1$, with the approximate cell constants for each crystal in the range of a = 100–106 Å, b = 111–114 Å, and c = 125–127 Å and diffracted the X-ray beam up to a resolution of 2 Å (details are provided in Table 1 and *SI Appendix*, Table S1).

Structure Determination. With a calculated Matthews coefficient of 2.4 $Å^3$/Da (49) two alpha/beta subunits per asymmetrical unit were expected. The phase problem was solved using the molecular replacement method with models prepared from PDB ID codes 2CSU (alpha subunit) and 1WR2 (beta subunit). Positioning of two entities of the alpha subunit model yielded a clear dimeric arrangement, which was in accordance with a calculated self-rotation function indicating a twofold noncrystallographic symmetry axis. Assuming a heterotetrameric $\alpha_2\beta_2$ complex, two molecular replacement solutions were also anticipated for the beta subunit. However, the molecular replacement search provided only one clear solution for the expected two molecules of the beta subunit with a reliable signal-to-noise ratio. At this stage, no reliable molecular replacement solutions could be obtained. Hence, the transformation matrix between both alpha subunits was calculated and applied on the first beta subunit to generate the symmetry-related second instance. The complete model was subjected to several rebuild cycles. Missing parts of the model became well visible in the electron density. Moreover, CoA, a component of the crystallization droplet, could clearly be identified in the difference electron density map. In subsequent steps, we prepared several crystals of *ckc*ACD1 in complex with various cofactors. Data collection and refinement statistics for these crystals are summarized in Table 1 and *SI Appendix*, Table S1.

Detailed descriptions for site-directed mutagenesis, kinetic characterization, and additional structural interpretations can be found in *SI Appendix*.

ACKNOWLEDGMENTS. We acknowledge access to the core facilities of the Centre for Biochemistry and Molecular Biology (BiMo)/Labor für Molekulare Biowissenschaften (LMB) of Kiel University. We thank Sebastian Krossa for helpful discussions and support during data collection. Diffraction data were collected on BL14.1, BL14.2, and BL14.3 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSYII storage ring [Berlin-Adlershof, Germany (50)], on the ID23-1 beamline at the European Synchrotron Radiation Facility (ESRF), and on P14 operated by the European Molecular Biology Laboratory (EMBL) at the Positron-Electron Tandem Ring Accelerator (PETRA) III storage ring (Deutsches Elektronen Synchrotron). We thank the beamline staff at all beamlines for providing assistance in using the beamlines. We thank the HZB for the allocation of synchrotron radiation beam time. We acknowledge financial support from the HZB. The research leading to these results has received funding from the European Community's Seventh Framework Programme (Grant FP7/2007-2013) for Biological Structure determination at synchrotron X-ray radiation facilities (BioStruct-X) (Grant Agreement 283570). Additionally, we are grateful for access to the High-Throughput Crystallisation (HTX) facility with a grant from P-Cube and BioStruct-X at the EMBL outstation. Beam time at P14 at the EMBL outstation was also funded by a BioStruct-X grant. This work was supported by Grant SCHO 316/10-1 from the Deutsche Forschungsgemeinschaft (to P.S.).

1. Schäfer T, Selig M, Schönheit P (1993) Acetyl-CoA synthetase (ADP forming) in archaea, a novel enzyme involved in acetate formation and ATP synthesis. *Arch Microbiol* 159(1):72–83.
2. Glasemacher J, Bock AK, Schmid R, Schönheit P (1997) Purification and properties of acetyl-CoA synthetase (ADP-forming), an archaeal enzyme of acetate formation and ATP synthesis, from the hyperthermophile *Pyrococcus furiosus*. *Eur J Biochem* 244(2):561–567.
3. Scott JW, Poole FL, Adams MW (2014) Characterization of ten heterotetrameric NDP-dependent acyl-CoA synthetases of the hyperthermophilic archaeon *Pyrococcus furiosus*. *Archaea* 2014:176863.
4. Awano T, et al. (2014) Characterization of two members among the five ADP-forming acyl coenzyme A (Acyl-CoA) synthetases reveals the presence of a 2-(Imidazol-4-yl)acetyl-CoA synthetase in *Thermococcus kodakarensis*. *J Bacteriol* 196(1): 140–147.
5. Musfeldt M, Selig M, Schönheit P (1999) Acetyl coenzyme A synthetase (ADP forming) from the hyperthermophilic Archaeon *pyrococcus furiosus*: Identification, cloning, separate expression of the encoding genes, acdAI and acdBI, in *Escherichia coli*, and in vitro reconstitution of the active heterotetrameric enzyme from its recombinant subunits. *J Bacteriol* 181(18):5885–5888.
6. Bräsen C, Schönheit P (2001) Mechanisms of acetate formation and acetate activation in halophilic archaea. *Arch Microbiol* 175(5):360–368.
7. Jones CP, Ingram-Smith C (2014) Biochemical and kinetic characterization of the recombinant ADP-forming acetyl coenzyme A synthetase from the amitochondriate protozoan *Entamoeba histolytica*. *Eukaryot Cell* 13(12):1530–1537.
8. Sánchez LB, Galperin MY, Müller M (2000) Acetyl-CoA synthetase from the amitochondriate eukaryote *Giardia lamblia* belongs to the newly recognized superfamily of acyl-CoA synthetases (Nucleoside diphosphate-forming). *J Biol Chem* 275(8): 5794–5803.
9. Schmidt M, Schönheit P (2013) Acetate formation in the photoheterotrophic bacterium *Chloroflexus aurantiacus* involves an archaeal type ADP-forming acetyl-CoA synthetase isoenzyme I. *FEMS Microbiol Lett* 349(2):171–179.
10. Bock AK, Glasemacher J, Schmidt R, Schönheit P (1999) Purification and characterization of two extremely thermostable enzymes, phosphate acetyltransferase and acetate kinase, from the hyperthermophilic eubacterium *Thermotoga maritima*. *J Bacteriol* 181(6):1861–1867.
11. Sánchez LB, Morrison HG, Sogin ML, Müller M (1999) Cloning and sequencing of an acetyl-CoA synthetase (ADP-forming) gene from the amitochondriate protist, *Giardia lamblia*. *Gene* 233(1-2):225–231.
12. Fraser ME, James MN, Bridger WA, Wolodko WT (1999) A detailed structural description of *Escherichia coli* succinyl-CoA synthetase. *J Mol Biol* 285(4):1633–1653.
13. Fraser ME, Joyce MA, Ryan DG, Wolodko WT (2002) Two glutamate residues, Glu 208 alpha and Glu 197 beta, are crucial for phosphorylation and dephosphorylation of the active-site histidine residue in succinyl-CoA synthetase. *Biochemistry* 41(2):537–546.
14. Joyce MA, Fraser ME, James MN, Bridger WA, Wolodko WT (2000) ADP-binding site of *Escherichia coli* succinyl-CoA synthetase revealed by x-ray crystallography. *Biochemistry* 39(1):17–25.
15. Joyce MA, Hayakawa K, Wolodko WT, Fraser ME (2012) Biochemical and structural characterization of the GTP-preferring succinyl-CoA synthetase from *Thermus aquaticus*. *Acta Crystallogr D Biol Crystallogr* 68(Pt 7):751–762.
16. Fraser ME, Hayakawa K, Hume MS, Ryan DG, Brownie ER (2006) Interactions of GTP with the ATP-grasp domain of GTP-specific succinyl-CoA synthetase. *J Biol Chem* 281(16):11058–11065.
17. Sun T, Hayakawa K, Bateman KS, Fraser ME (2010) Identification of the citrate-binding site of human ATP-citrate lyase using X-ray crystallography. *J Biol Chem* 285(35):27418–27428.
18. Sun T, Hayakawa K, Fraser ME (2011) ADP-Mg2+ bound to the ATP-grasp domain of ATP-citrate lyase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 67(Pt 10): 1168–1172.
19. Wolodko WT, Fraser ME, James MN, Bridger WA (1994) The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5-A resolution. *J Biol Chem* 269(14): 10883–10890.
20. Bräsen C, Schmidt M, Grötzinger J, Schönheit P (2008) Reaction mechanism and structural model of ADP-forming Acetyl-CoA synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*: Evidence for a second active site histidine residue. *J Biol Chem* 283(22):15409–15418.
21. Musfeldt M, Schönheit P (2002) Novel type of ADP-forming acetyl coenzyme A synthetase in hyperthermophilic archaea: Heterologous expression and characterization of isoenzymes from the sulfate reducer *Archaeoglobus fulgidus* and the methanogen *Methanococcus jannaschii*. *J Bacteriol* 184(3):636–644.
22. Hildebrand JG, Spector LB (1969) Succinyl phosphate and the succinyl coenzyme A synthetase reaction. *J Biol Chem* 244(10):2606–2613.
23. Srere PA (1959) The citrate cleavage enzyme. I. Distribution and purification. *J Biol Chem* 234:2544–2547.
24. Mai X, Adams MW (1996) Purification and characterization of two reversible and ADP-dependent acetyl coenzyme A synthetases from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* 178(20):5897–5903.
25. Shikata K, Fukui T, Atomi H, Imanaka T (2007) A novel ADP-forming succinyl-CoA synthetase in *Thermococcus kodakaraensis* structurally related to the archaeal nucleoside diphosphate-forming acetyl-CoA synthetases. *J Biol Chem* 282(37): 26963–26970.
26. Elkins JG, et al. (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci USA* 105(23):8102–8107.
27. Schmidt M (2013) ADP-abhängige Acyl-CoA Synthetasen in Archaea, Bacteria und Eukarya: Charakterisierung, Funktion und Phylogenie. PhD thesis (Kiel University, Kiel, Germany). German.
28. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.
29. Krissinel E (2010) Crystal contacts as nature's docking solutions. *J Comput Chem* 31(1): 133–143.
30. Rossmann MG, Moras D, Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* 250(463):194–199.
31. von Heijne G, Blomberg C (1978) Early evolution of cellular electron transport: Molecular models for the ferredoxin-rubredoxin-flavodoxin region. *Orig Life* 9(1):27–37.
32. Fawaz MV, Topper ME, Firestine SM (2011) The ATP-grasp enzymes. *Bioorg Chem* 39(5-6):185–191.
33. Thoden JB, Firestine SM, Benkovic SJ, Holden HM (2002) PurT-encoded glycinamide ribonucleotide transformylase. Accommodation of adenosine nucleotide analogs within the active site. *J Biol Chem* 277(26):23898–23908.
34. Polekhina G, Board PG, Gali RR, Rossjohn J, Parker MW (1999) Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. *EMBO J* 18(12): 3204–3213.
35. Huang J, Malhi M, Deneke J, Fraser ME (2015) Structure of GTP-specific succinyl-CoA synthetase in complex with CoA. *Acta Crystallogr F Struct Biol Commun* 71(Pt 8): 1067–1071.
36. Deane CM, Allen FH, Taylor R, Blundell TL (1999) Carbonyl-carbonyl interactions stabilize the partially allowed Ramachandran conformations of asparagine and aspartic acid. *Protein Eng* 12(12):1025–1028.
37. Reeves RE, Warren LG, Susskind B, Lo HS (1977) An energy-conserving pyruvate-to-acetate pathway in *Entamoeba histolytica*. Pyruvate synthase and a new acetate thiokinase. *J Biol Chem* 252(2):726–731.
38. Bräsen C, Schönheit P (2004) Unusual ADP-forming acetyl-coenzyme A synthetases from the mesophilic halophilic euryarchaeon *Haloarcula marismortui* and from the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *Arch Microbiol* 182(4): 277–287.
39. Bräsen C, Schönheit P (2004) Regulation of acetate and acetyl-CoA converting enzymes during growth on acetate and/or glucose in the halophilic archaeon *Haloarcula marismortui*. *FEMS Microbiol Lett* 241(1):21–26.
40. Sanchez LB, Müller M (1996) Purification and characterization of the acetate forming enzyme, acetyl-CoA synthetase (ADP-forming) from the amitochondriate protist, *Giardia lamblia*. *FEBS Lett* 378(3):240–244.
41. Hidber E, Brownie ER, Hayakawa K, Fraser ME (2007) Participation of Cys123alpha of *Escherichia coli* succinyl-CoA synthetase in catalysis. *Acta Crystallogr D Biol Crystallogr* 63(Pt 8):876–884.
42. Allende JE, Allende CC (1995) Protein kinases. 4. Protein kinase CK2: An enzyme with multiple substrates and a puzzling regulation. *FASEB J* 9(5):313–323.
43. Niefind K, Pütter M, Guerra B, Issinger OG, Schomburg D (1999) GTP plus water mimic ATP in the active site of protein kinase CK2. *Nat Struct Biol* 6(12):1100–1103.
44. Issinger OG (1993) Casein kinases: Pleiotropic mediators of cellular regulation. *Pharmacol Ther* 59(1):1–30.
45. Lambeth DO, Tews KN, Adkins S, Frohlich D, Milavetz BI (2004) Expression of two succinyl-CoA synthetases with different nucleotide specificities in mammalian tissues. *J Biol Chem* 279(35):36621–36624.
46. Majumdar R, Guest JR, Bridger WA (1991) Functional consequences of substitution of the active site (phospho)histidine residue of Escherichia coli succinyl-CoA synthetase. *Biochim Biophys Acta* 1076(1):86–90.
47. Bailey DL, Fraser ME, Bridger WA, James MN, Wolodko WT (1999) A dimeric form of Escherichia coli succinyl-CoA synthetase produced by site-directed mutagenesis. *J Mol Biol* 285(4):1655–1666.
48. Glaser F, et al. (2003) ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* 19(1):163–164.
49. Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* 33(2):491–497.
50. Mueller U, et al. (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J Synchrotron Radiat* 19(Pt 3):442–449.
51. Laskowski RA (2009) PDBsum new things. *Nucleic Acids Res* 37(Database issue): D355–D359.
52. Ho BK, Gruswitz F (2008) HOLLOW: Generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct Biol* 8:49.
53. Popov AN, Bourenkov GP (2003) Choice of data-collection parameters based on statistic modelling. *Acta Crystallogr D* 59(Pt 7):1145–1153.
54. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.
55. Brünger AT (1992) Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355:472–475.
56. Tickle IJ, Laskowski RA, Moss DS (2000) R_{free} and the R_{free} ratio. II. Calculation of the expected values and variances of cross-validation statistics in macromolecular least-squares refinement. *Acta Crystallogr D* 59(Pt 4):442–450.

Supporting Information

Weiße et al.

SI Materials and Methods

Cloning. Genomic DNA of *Candidatus Korarchaeum cryptofilum* was kindly provided by Dr. James G. Elkins (Oak Ridge National Laboratory, USA). The open reading frames (ORF) coding for a putative ACD alpha subunit (kcr_0198) and beta subunit (kcr_0115) were amplified by PCR. For kcr_0198 the primer pair 5′-CTTGGGATGCCATATGAACG-ACCTAGAGAGGC-3′ and 5′-GATGAACGTCTCGAGATCACCTCACGGCCAGG-3′ was used. For kcr_0115 the primer pair 5′-CGGTGATCATATGAGCTCAAGGGACC-3′ and 5′-CTTCCTCGAGAAGCCCTCACCTCAGG-3′ was used. Each primer pair contained restriction sites for the endonucleases NdeI and XhoI (underlined), which were used for insertion of the amplified ORF into pET17b to yield the expression vectors pET17bkcr_0198 and pET17bkcr_0115, respectively. Each vector was subsequently transformed into *E. coli* strain BL21(DE3) CodonPlus RIL (Stratagene).

Expression and purification. Recombinant *ckc*ACD1 was obtained by heterologous expression in *E. coli* BL21(DE3) CodonPlus RIL cells. Bacterial strains bearing either the expression plasmids encoding the alpha subunit (pET17bkcr_0198) or the beta subunit (pET17bkcr_0115) were independently cultivated in lysogenic broth (LB) media at 37 °C in presence of ampicillin and chloramphenicol. The following expression protocol was applied to each culture. At an optical density of 0.4 to 0.8 at 600 nm, expression was induced using IPTG to a final concentration of 0.4 mM. After a further cultivation period of 5 hours, cells were harvested by centrifugation and re-suspended in ACD buffer (100 mM TRIS/HCl, 150 mM NaCl and 5 mM $MgCl_2$ (final pH 7.5)). Cell lysis was performed by sonification. Cell debris was removed via centrifugation. The cleared lysates were incubated at 68 °C for 20 min to precipitate heat-labile proteins, which were pelleted by centrifugation. SDS-PAGE analysis was used to estimate the amount and purity of the target proteins. Reconstitution of the active *ckc*ACD1 heterotetramer was performed by mixing supernatants from both subunits, with the beta subunit in stoichiometric excess. The reconstituted *ckc*ACD1 complex was purified via gel filtration (Superdex 200 (GE Healthcare)) using ACD buffer for equilibration and protein elution. Due to a massive release of endogenous DNA during the lysis step, the eluted protein was contaminated with a high concentration of oligonucleotides. Multiple repeats of

purification via gel filtration were performed until UV spectra showed only protein specific absorption. Prior the final gel filtration step, the protein solution was incubated at 78 °C for 15 min. Purified *ckc*ACD1 was concentrated to 12.9 mg/ml as determined via UV absorption using a theoretical extinction coefficient of 49,280 $M^{-1}cm^{-1}$ calculated with the software tool PROTPARAM (1). Seleno-methionine labelled *ckc*ACD1 was produced with the following modifications for expression and purification. SelenoMethionine Medium Base supplemented with SelenoMethionine Medium Nutrient Mix (both Molecular Dimensions, Newmarket, United Kingdom) was used according to manufacturer's instructions (referred to as minimal medium in the following text). *E. coli* cells were cultured in this minimal medium with 50 mM methionine until an optical density of 0.4 at 600 nm was reached. At that point the cells were pelleted and re-suspended in pre-warmed minimal medium supplemented with 25 mM seleno-methionine. After a 20 min cultivation period, expression was induced with IPTG to a final concentration of 0.4 mM. The bacterial cells were harvested after 3 hours of cultivation. Essentially the same purification procedure as described for wild type *ckc*ACD1 was applied, with buffers additionally supplemented with 20 mM DTT.

Crystallisation. Initial crystallization experiments were performed at the HTX facility at the EMBL outstation (Hamburg, Germany) in sitting drop geometry. Crystals were obtained in various conditions, indicating two crystal forms according to their different optical appearance. Optimization of various parameters including pH, precipitation agent and concentration, crystallization temperature and droplet geometry were carried out to finally yield larger and well-formed crystals suitable for X-ray diffraction experiments. In order to obtain enzyme in complex with various substrates, the protein solution was usually supplemented with the appropriate cofactors. A magnesium ion-reduced protein solution was prepared by dialyzing against the 100-fold volume of 100 mM TRIS/HCl, 150 mM sodium chloride (final pH 7.5) using a membrane with 12 kDa MWCO. Dialysis was performed at 348 K. Precipitate was removed via centrifugation. Prior to further experiments the appropriate ligands and cofactors were added to the protein solution. Phosphorylated *ckc*ACD1 was obtained by either incubation with 10 mM ATP in ACD buffer for 2 min at 348 K or with 1 mM Ac-CoA in 10 mM Na_2HPO_4 and 5 mM sodium acetate for 1 min at 348 K using magnesium ion-reduced protein solution.

Diffraction experiments. Prior to any diffraction experiment the crystals were swiped through a cryo protection solution composed of 20 % (v/v) MPD and 80 % (v/v) mother liquor with 10 % higher concentration compared to the reservoir solution and eventually flash frozen in

liquid nitrogen. To avoid the washing out of bound ligands during this procedure, the ligands were added as well to the cryo buffer at the same concentration as in the crystallization droplet. X-Ray diffraction was performed at 100 K. Diffraction data were collected at various beamlines: ID23-1 at ESRF (Grenoble, France), beamlines BL14.1, BL14.2 and BL14.3 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron storage ring (Berlin-Adlershof, Germany)(2) and beamline P14 (MX2) at PETRA III (DESY, Hamburg, Germany). Diffraction images were indexed and scaled with XDS and XSCALE (3). Merging of reflections was done using AIMLESS (4).

Phasing and structure refinement. The crystallographic phase problem was solved using molecular replacement. A molecular model for the alpha subunit with the correct amino acid annotation was automatically generated by the SWISS-MODEL server (5) using the crystal structure of PH0766 from *Pyrococcus horikoshii* OT3 (PDB entry 2CSU) as template structure. For the beta subunit a model automatically built within MOLREP (6) based on PDB entry 1WR2 by providing the target amino acid sequence. Placement of the subunits was calculated using MOLREP and software from the PHENIX suite (7). Refinement of the crystal structures was performed with REFMAC5 (8) and PHENIX.REFINE (9). Manual model building was carried out using COOT (10). Validation of the obtained model was performed with tools within COOT as well as PROCHECK (11), SFCHECK (12) and MOLPROBITY (13). Ligands were included for refinement when they were visible in the electron density maps. Ligand restraint files were generated with the Grade Web Server (Smart, O. S., T. O.Womack, A. Shar, C. Flensburg, P. Keller, W. Paciorek, C. Vonrhein and G. Bricogne (2011). Grade, version 1.2.7. http://www.globalphasing.com.). The PHENIX suite was used to create simulated annealing omit maps. Structural representations were designed with PyMOL (Schrödinger).

Enzyme activity. Enzyme activity was monitored spectrophotometrically at 50 °C by HSCoA liberation using Ellman´s thiol reagent (DTNB) (14). The assay contained 100 mM MES/NaOH pH 7.0, 0.1 mM DTNB, 5 mM $MgCl_2$, 1 mM of either ADP or GDP, 0.1 mM acetyl-CoA, 20 mM KH_2PO_4. For determination of K_m values, the concentration of ADP and GDP was varied between 0 and 2 mM.

Generation of kcr_0115 mutants. The *ckc*ACD1-H68A and *ckc*ACD1-H68N mutants of the beta subunit were generated using the QuikChange II XL protocol (Stratagene) according to the manufacturer's manual. The primers used to introduce the mutations were 5´-

TATCCCCTCAAGTGGTCGCTAAAAGCGATGTGGGCG-3′ (forward) together with 5′-CGCCCACATCGCTTTTAGCGACCACTTGAGGGGATA-3′ (reverse) for generation of pET17bkcr_0115H69A (*ckc*ACD1-H68A) and 5′-TATCCCCTCAAGTGGTCAAT-AAAAGCGATGTGGGC-3′ (forward) together with 5′-GCCCACATCGCTTTTATT-GACCACTTGAGGGGATA-3′ (reverse) for generation of pET17bkcr_0115H69N (*ckc*ACD1-H68N). *E. coli* BL21(DE3) RIL cell pellets, transformed with pET17bkcr_0115 or pET17bkcr_0198, respectively, and *E. coli* Rosetta (DE3) pLysS cell pellets transformed with pET17bkcr_0115H69A or pET17bkcr_0115H69N, respectively, were suspended in TRIS/HCl, pH 8.85 containing 5 mM $MgCl_2$ and disrupted by sonication. Supernatant was heat-precipitated at 85 °C for 20 min and precipitated proteins were removed by centrifugation. Subunits were reconstituted and applied to a Superdex 200 16/60 column equilibrated with 50 mM TRIS/HCl, pH 7.5, containing 150 mM NaCl. Protein was eluted at a flow rate of 1 ml min^{-1}, yielding pure protein.

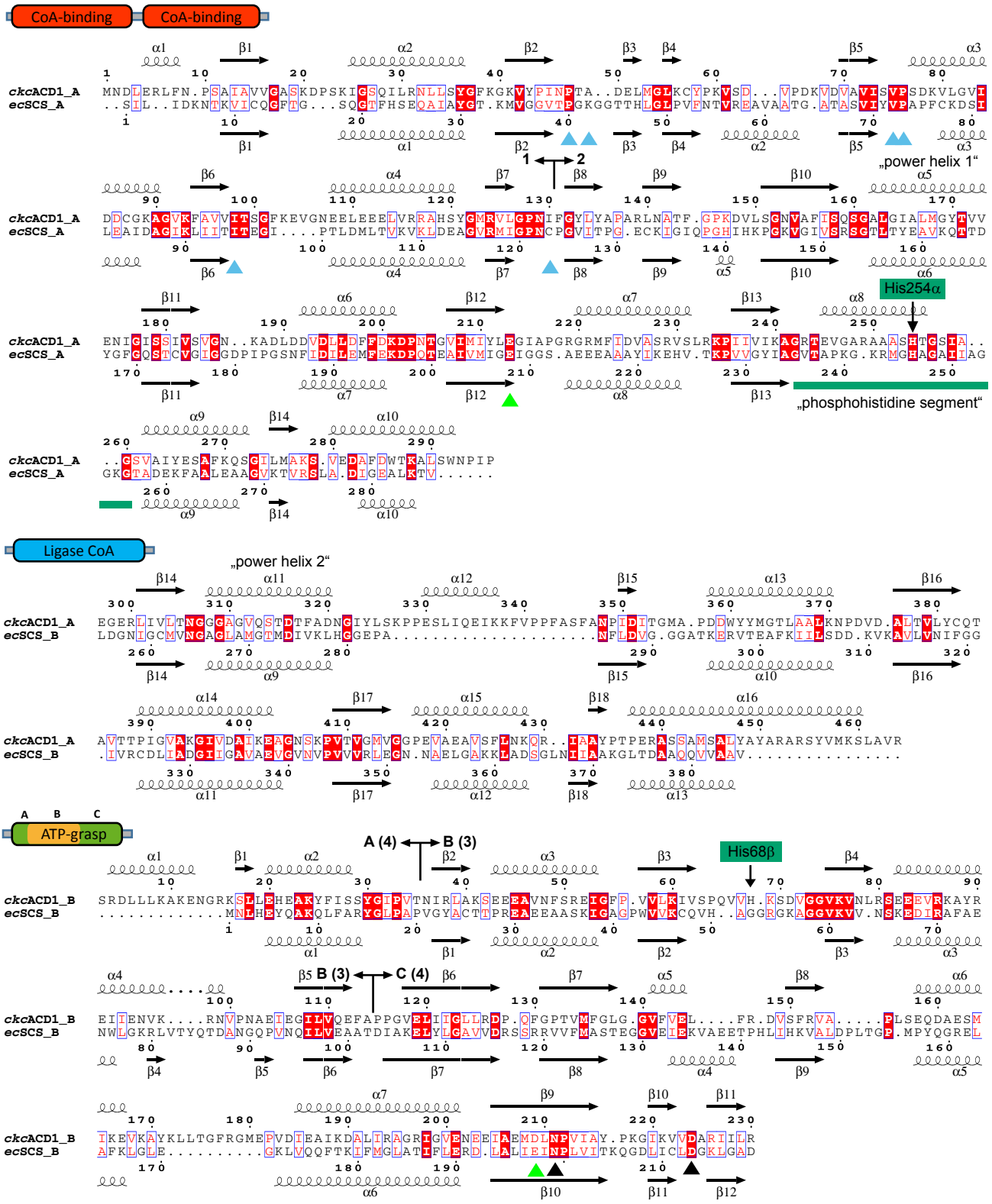
CoA-binding
CoA-binding
ckcACD1_A
ecSCS_A
„power helix 1“
His254α
„phosphohistidine segment“
Ligase CoA
„power helix 2“
ckcACD1_A
ecSCS_B
A B C
ATP-grasp
A (4) B (3)
His68β
B (3) C (4)
ckcACD1_B
ecSCS_B

Fig. S1. Sequence alignment of *ckc*ACD1 with *ec*SCS. The sequence alignment is based on superimposed structures using the program GESAMT (15). Since due to domain shuffling the order of individual domains is rearranged, equivalent domains are individually aligned. The respective domain notation is given. Residues of *ec*SCS which are known to be important for catalysis are marked by a triangle in green (16, 17) (*ec*SCS-residues: His246α, Glu208α – interacting with His246α, Glu197β – mostly Asp in other family members). Residues interacting with coenzyme A (CoA) are marked by a triangle in blue (18)(*ec*SCS residues: Pro40α, Lys42α, Val72α, Pro73α, Ile95α and Cys123α). Residues coordinating the magnesium ion of the ADP-binding site within the ATP-grasp domain are marked by a triangle in black (*ckc*ACD1 residues: Asn211β and Asp224β). The phosphohistidine segment with the phosphorylated histidine residue His254α is underlined in green. The second, highly conserved histidine residue His68β known to be important for catalysis in *ckc*ACD1 is highlighted in green. The "power helices" defining the phosphate binding site I are labeled (*ec*SCS α5(α-subunit) and α9(β-subunit); *ckc*ACD1 α5(α-subunit) and α11(α'-subunit)). The " ' " indicates the symmetry mate within the $\alpha_2\beta_2$-heterotetramer. The subdivision in three domains (A, B, and C) for the ATP-grasp domain is in accordance to Fawaz *et al.* (19). The hinge region for the lid-domain B of the ATP-grasp domain (beta subunit in *ckc*ACD1 and subdomain 3-4 in *ec*SCS) is located at the two changeovers from one domain to the other as indicated (A←→ B and B ←→ C, respectively).

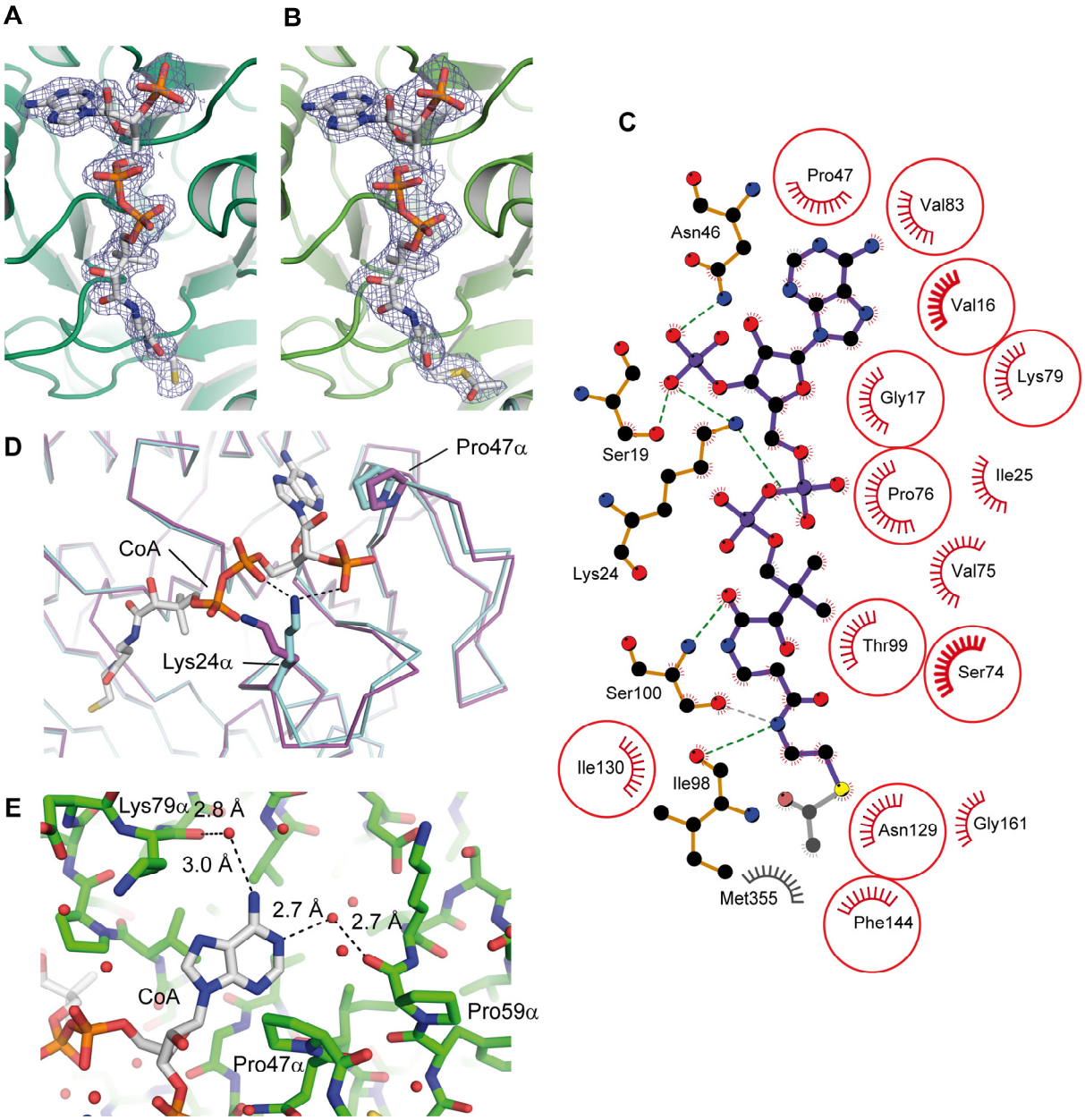

Fig. S2. Binding mode of (acetyl-)CoA within the alpha subunit. Electron density map defining the position of the bound cofactor CoA (**A**) and acetyl-CoA (**B**). Carbon atoms in acetyl-CoA are shown in white, nitrogen in blue, oxygen in red, phosphor in orange, and sulfur in yellow. Density is contoured at 1.0σ (blue mesh). The electron density distribution was calculated without the presence of the cofactor. The surrounding protein is displayed as cartoon representation in green. The binding mode is identical for the main part of the cofactor. (**C**) Schematic representation of the binding environment for acetyl-CoA as calculated with the program LigPlot (20). (**D**) Only small conformational adaptions where observed upon CoA binding in comparison to the cofactor free apoenzyme. The sidechain of Lys24α becomes oriented between the 3‘- and 5‘- phosphate groups of CoA. Due to hydrophobic interactions of

the adenosine moiety with the protein, a small closure of the pocket is observed around residue Pro47α. (**E**) Orientation of the adenosine group of CoA in the hydrophobic pocket is facilitated by interaction with water molecules (red spheres) coordinated to the main chain carbonyl group of Pro59α and Lys79α, respectively.

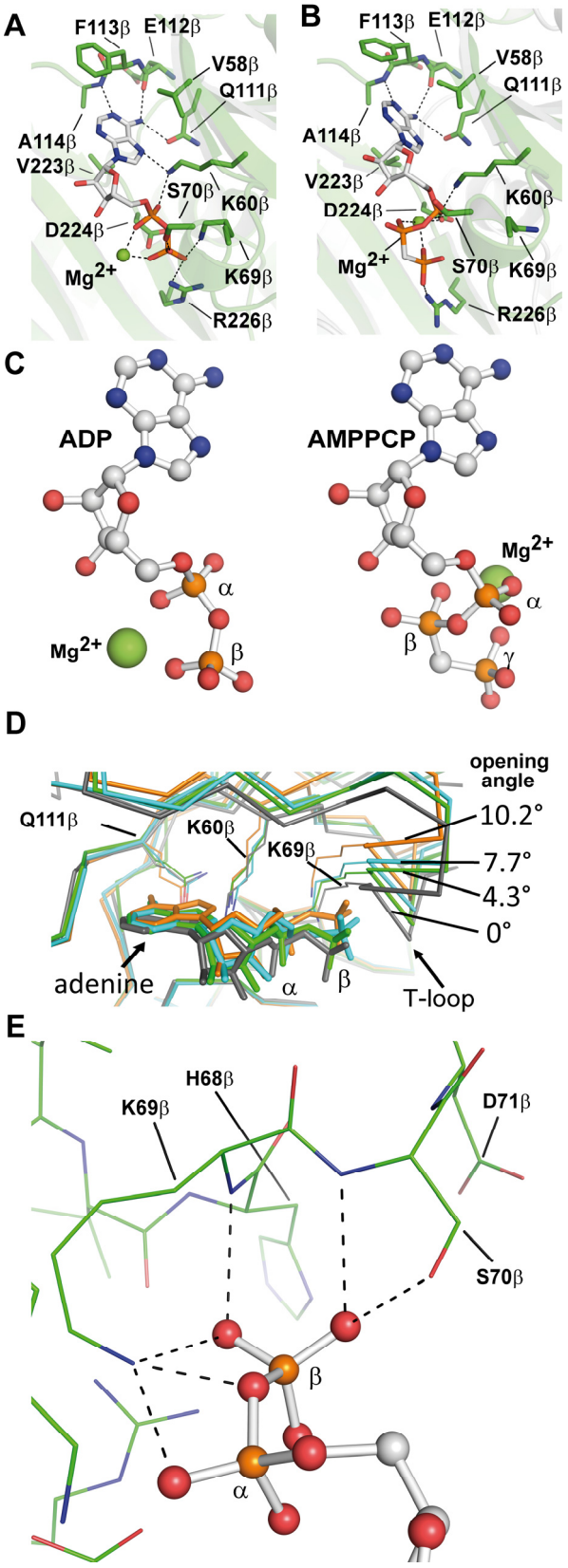

Fig. S3. Binding of nucleotides within the ATP-grasp domain of the beta subunit. Residues interacting with (A) ADP (crystal structure *ckc*ACD1-D) and (B) AMPPCP (crystal structure *ckc*ACD1-C). Carbon atoms in the nucleotides are shown in white, nitrogen in blue, oxygen in

red, and phosphor in orange. (C) Orientation of the phosphate groups of ADP and AMPPCP, respectively, based on the superposition of the adenine and sugar moiety. The green spheres represent magnesium ions, which are complexed by the phosphate groups. (D) Opening of the ATP-grasp domain due to the movement of the lid domain causes changes of the adenosine diphosphate binding. Shown is a superposition of the ATP-grasp domains of the crystal structures *ckc*ACD1-B (grey), *ckc*ACD1-D (green), *ckc*ACD1-H (blue), and *ckc*ACD1-G (orange). (E) Interaction of the beta phosphate group of ADP with the T-loop (Lys60β-Val75β). Depicted is the crystal structure of *ckc*ACD1-D. Several hydrogen bonds (black dashes) are observed, which facilitate strong interactions between the protein and the phosphate moiety.

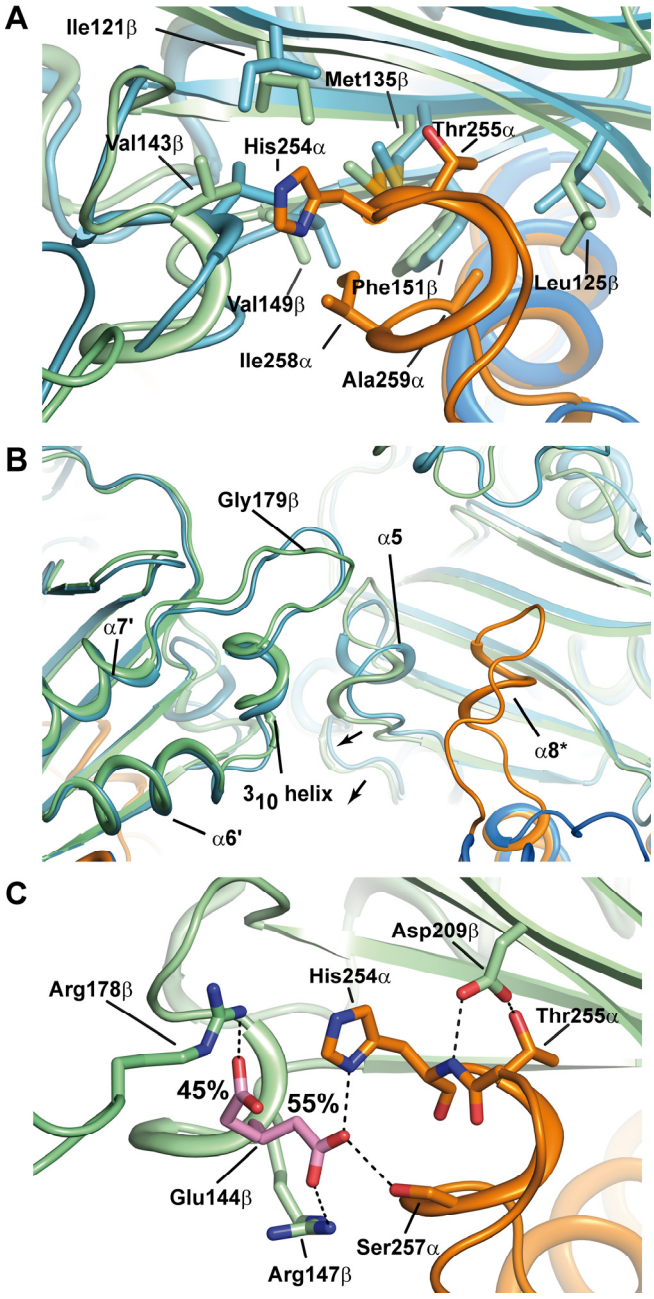

Fig. S4. (**A**) Interaction of the phosphohistidine segment (orange) with the ATP-grasp domain of the beta subunit causes rearrangements. Depicted are the ATP-grasp domain of *ckc*ACD1-B (blue; phosphohistidine segment positioned towards binding site I (see as well **Fig. 3** and **Fig. 4**)) and *ckc*ACD1-C (orange; phosphohistidine segment swung towards binding site II (see as well **Fig. 4** and **Fig. 5**)). Residues from the lid domain are omitted for clarity. (**B**) Formation of a hydrophobic pocket results in reorganization of helix α5 and subsequently in relocation of the residues between the helices α6' and α7'. Arrows indicate the direction of the movement within the ATP-grasp domain. (**C**) Two conformers for the glutamate residue Glu144β (carbon atoms in magenta and oxygen atoms in red) are observed in the crystal structure *ckc*ACD1-C. The value noted next to the conformers are the refined occupancies as observed in the crystal structure. Conformer A (55%) of Glu144β is involved in hydrogen bonds with the phosphohistidine segment side chains His254α and Ser257α (carbon atoms in orange, nitrogen atoms in blue, and oxygen atoms in red). In addition, an ionic interaction to the guanidine group of Arg147β is observed. The conformer B (45%) forms a salt bridge to the guanidine group of Arg178β'. Depicted is also residue Asp209β (carbon atoms in green and oxygen atoms in red). This residue stabilizes the phosphohistidine segment via hydrogen bonds to the side chain of Thr255α and the backbone nitrogen of His254α. Residues from the lid domain are omitted for clarity.

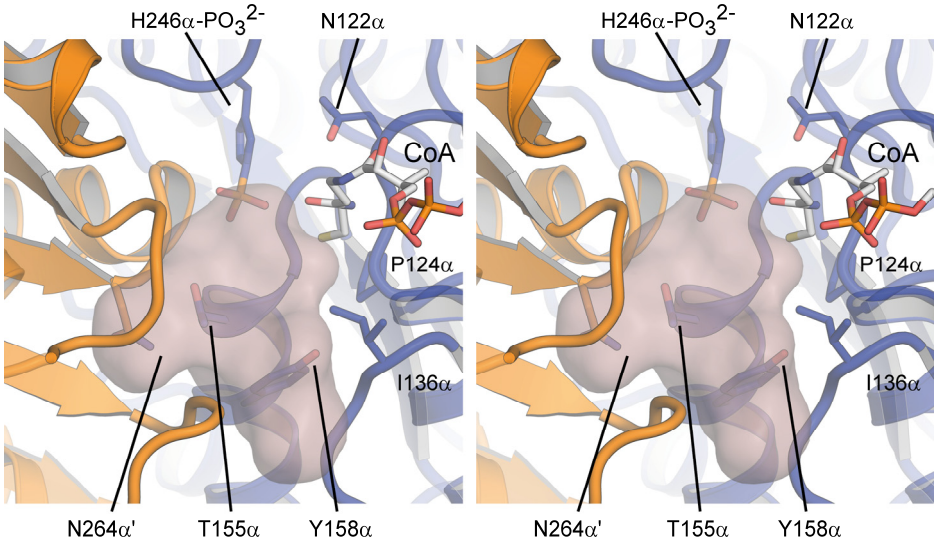

Fig. S5. Binding site for activated acyl-moiety. Stereo representation of the proposed binding pocket for the succinyl group as defined by the *ec*SCS structure (PDB entry 2SCU (18)).

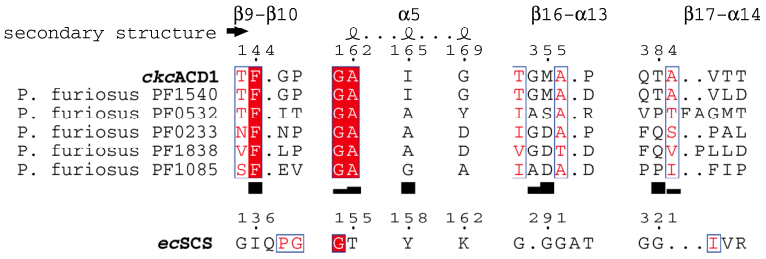

Fig. S6. Sequence variations within the substrate binding region determines substrate selectivity. Multiple sequence alignment of *ckc*ACD1, *ec*SCS and five isoforms of the ACDs from *P. furiosus*. The converted substrates for the different *pf*ACDs are well characterized (21). The potential involvement of individual residues on substrate recognition are based on the interpretation of the obtained structures of *ckc*ACD1 and the importance is highlighted as black bar.

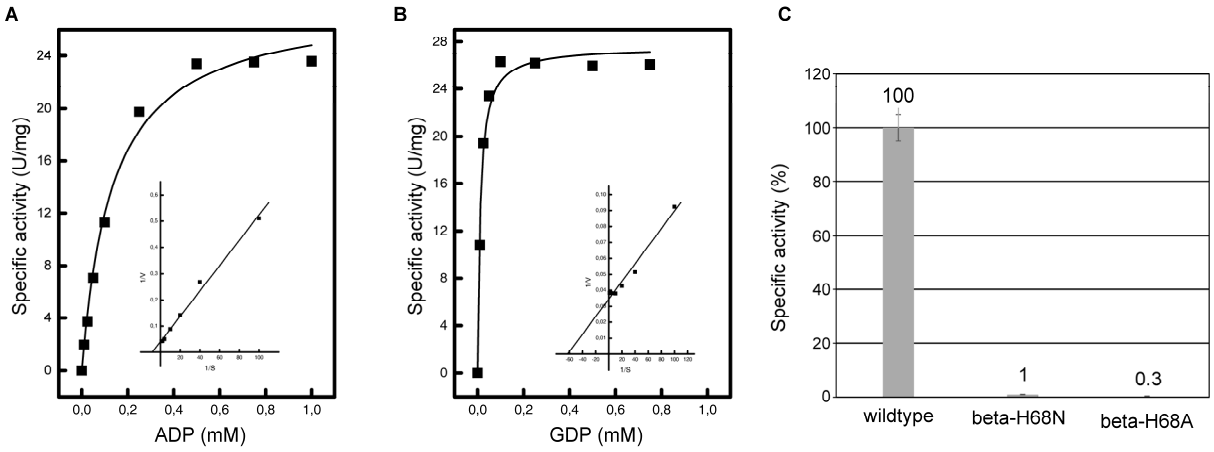

Fig. S7. Enzymatic activity of wild type and mutant *ckc*ACD1.The determination of the rate dependency of recombinant *ckc*ACD1 on ADP (*A*) and GDP (*B*) revealed that *ckc*ACD1 could use ADP as well as GDP as substrate for the activated phosphate. The inserts show the double-reciprocal plots of the rates against the corresponding substrate concentrations. The determined K_m and V_{max} values are 0.1 mM and 24-25 U/mg for ADP (duplicate measurement), respectively, and 0.04 mM and 29-32 U/mg for GDP (triplicate measurement), respectively. (*C*) His68β is crucial for *ckc*ACD1 activity to generate ATP. Levels of acetate-forming activity of wildtype and beta subunit variants of recombinant *ckc*ACD1 are given in percent activity. Specific activity of wildtype (50 U/mg) is set as 100 % activity. The mutant proteins *ckc*ACD1-

H68N (beta-H68N) and *ckc*ACD1-H68A (beta-H68A) have only residual activities of 0.5 U/mg and 0.15 U/mg, respectively. Given error bars indicate standard deviation of 5 to 8 measurements (wildtype = 4.89 U/mg, beta-H68N = 0.18 U/mg, beta-H68A = 0.08 U/mg). Measurements were performed at 85 °C.

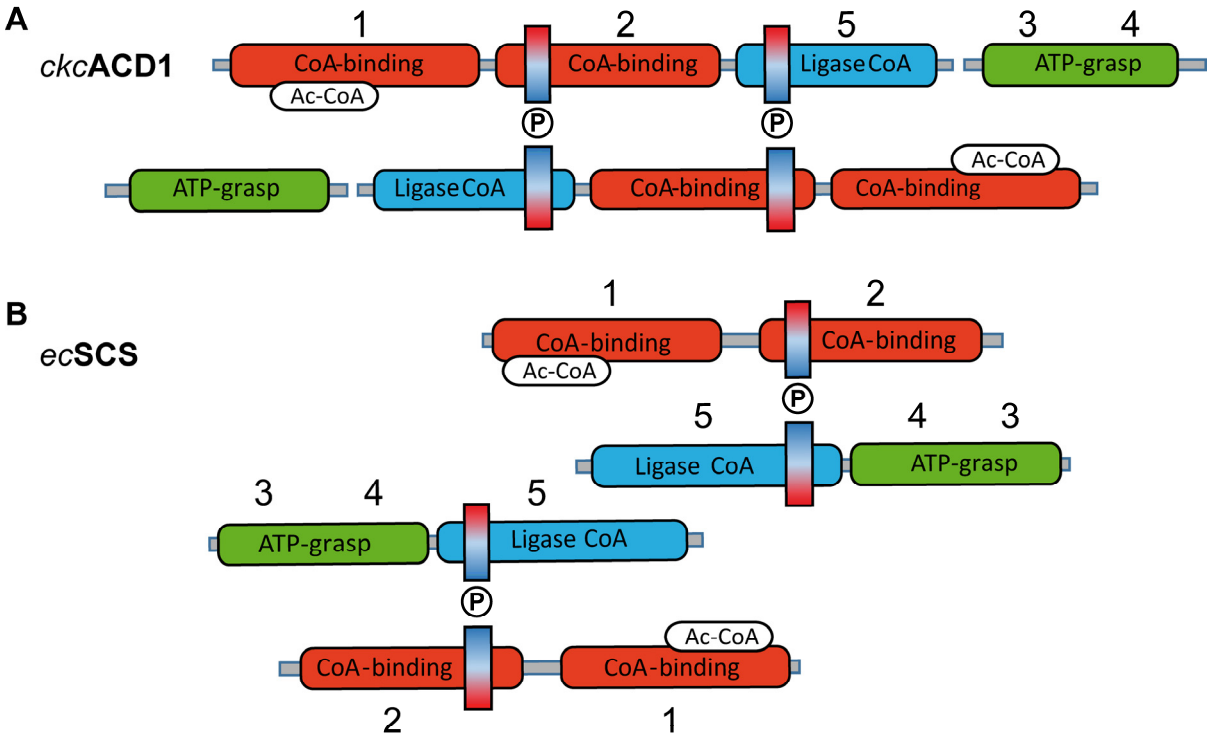

Fig. S8. The domain arrangement of *ckc*ACD1 requires a heterotetrameric $\alpha_2\beta_2$ complex for activity. Scheme of domain arrangement and heterocomplex formation of *ckc*ACD1 (*A*) and *ec*SCS (*B*). The color code for the individual subdomains is the same as used for **Fig. 1 and SI, Fig. S1**. The boxes represent the two power helices, which stabilize the phosphate moiety within the active site I.

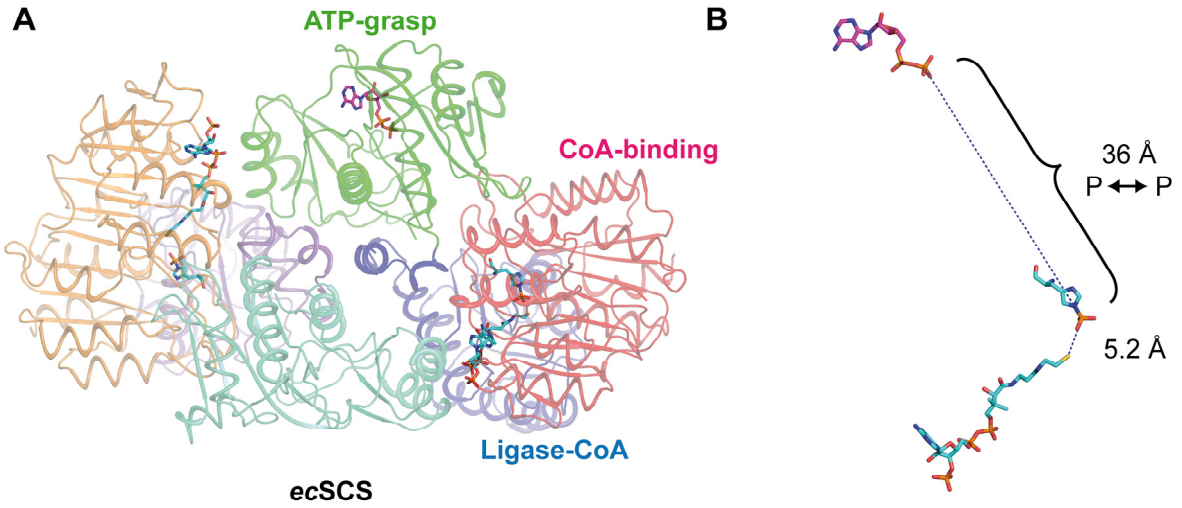

Fig. S9. For the heterotetrameric *ec*SCS a swinging-loop mechanism is proposed. (*A*) Depiction of the arrangement of the five subdomains within *ec*SCS (color scheme in accordance to **Fig. 1)**. The cofactor CoA and the final substrate ADP are shown in stick representation. (*B*) Distance between phosphorylated histidine and CoA (carbon atoms in cyan, nitrogen atoms in blue, phosphor atoms in orange, sulfur atom in yellow, and oxygen atoms in red) oriented towards site I and ADP bound at site II (carbon atoms in magenta, nitrogen atoms in blue, phosphor atoms in orange, and oxygen atoms in red). Representation is based on PDB-entries 2SCU (18) and 1CQJ (22).

The alpha domain features an additional nucleotide binding site. In the electron density of *ckc*ACD1-D and *ckc*ACD1-C, clear evidence for an additionally bound ADP and AMPPCP molecule, respectively, was identified. We observed this only in data collected from crystals grown at 278 K, which indicate an artificial origin. Thus, this observation has most likely no physiological relevance. The terminal phosphate groups of both ADP and AMPPCP facilitate the binding to the alpha domain. In case of ADP, the adenine moiety interacts in addition via π-π-stacking with the imidazole group of His254α. We did not observe electron density for the adenine moiety of AMPPCP. In both cases, we identified an additionally bound metal ion, presumably a magnesium ion, in octahedral coordination through one oxygen atom of the phosphate group as well as five surrounding water molecules (**SI, Fig. S4**). Two of these water molecules are furthermore coordinated to the carboxyl group of Asp351α'. Interestingly, a similar interaction of water molecules with Asp351α' could be observed as well in the crystal structures of *ckc*ACD1 with phosphorylated His254α (*ckc*ACD1-H) and in complex with a

phosphate ion (*ckc*ACD1-G) (**Fig. 3**). This aspartate residue is extraordinarily conserved in all ACDs and SCS (**SI, Fig. S1** and (23)). Hence, we propose that this residue and the observed phosphate binding motif might play an important role in the enzymatic function of ACDs.

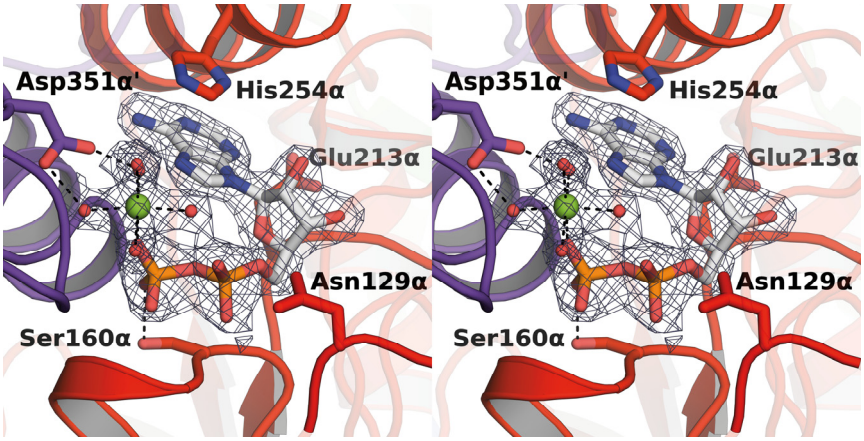

Fig. S10. ADP-binding site within the alpha subunit. Stereo representation of the omit electron density map used to define the position of ADP within the alpha subunit (blue mesh, density is contoured at 1.0σ). ADP as well as interacting residues are represented as sticks (carbon atoms in white, nitrogen atoms in blue, phosphor atoms in orange, and oxygen atoms in red). The bound magnesium ion is shown as green sphere and the coordinating water molecules are shown as red spheres.

Table S1. Data collection and refinement statistics.

***ckc*ACD1-#**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
PDB entry	4XYL	4XYM	4XZ3	4Y8V	4YAK	4YAJ	4YB8	4YBZ	5HBR
Data collection[1]									
Synchrotron X-ray source	PETRAIII P14	PETRAIII P14	BESSY BL14.1	BESSY BL14.1	PETRAIII P14 & BESSY BL14.1	BESSY BL14.1	PETRAIII P14	PETRAIII P14	PETRAIII P14
Wave length (Å)	1.24	1.24	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Resolution range (Å)	83.46-1.95 (1.98-1.95)	127.04-1.90 (1.93-1.90)	48.84-2.40 (2.47-2.40)	49.21-2.10 (2.14-2.10)	83.57-2.46 (2.53-2.46)	125.81-2.20 (2.25-2.20)	110.61-1.90 (1.93-1.90)	83.03-2.10 (2.14-2.10)	126.03-1.99 (2.03-1.99)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell a, b, c (Å)	102.5, 112.2, 124.9	99.6, 114.4, 127.0	100.2, 111.9, 127.6	106.5, 111.0, 126.7	106.4, 111.7, 126.0	105.8, 111.2, 125.8	106.1, 110.6, 125.7	105.3, 109.9, 126.8	106.1, 110.7, 126.0
Total reflections	684725 (33884)	706193 (8766)	711978 (56845)	548838 (15840)	591856 (23566)	570073 (7989)	1515778 (55680)	573438 (30751)	672786 (27311)
Unique reflections	105037 (5118)	112092 (3230)	56770 (4557)	86108 (3595)	54078 (3442)	73935 (2617)	116787 (5673)	85835 (4477)	101504 (4632)
Multiplicity	6.5 (6.6)	6.3 (2.7)	12.5 (12.5)	6.4 (4.4)	10.9 (6.8)	7.7 (3.1)	13.0 (9.8)	6.7 (6.9)	6.6 (5.9)
Completeness (%)	99.8 (99.7)	97.8 (58.1)	99.9 (98.8)	97.6 (77.9)	98.1 (77.3)	97.3 (57.0)	100.0 (99.5)	99.5 (99.1)	99.6 (93.1)
Mean I/σ(I)	6.7 (0.6)	8.1 (0.3)	9.9 (0.8)	10.1 (1.1)	9.7 (0.6)	6.4 (0.4)	9.4 (0.5)	7.5 (0.8)	8.4 (0.5)
CC(1/2)	0.992 (0.169)	0.996 (0.103)	0.995 (0.229)	0.995 (0.422)	0.996 (0.231)	0.990 (0.105)	0.998 (0.159)	0.993 (0.312)	0.998 (0.124)
Wilson *B*-factor ($Å^2$)	25.5	24.9	40.2	20.4	45.3	30.8	37.1	26.2	37.2
Refinement									
R-work (%)	20.64	19.53	19.65	18.44	19.36	21.45	18.16	21.58	17.84
R-free (%)	24.64	23.71	24.44	22.55	23.82	24.13	22.05	25.34	22.22
CC* (24)	0.998	0.999	0.999	0.999	0.999	0.998	0.999	0.998	0.999
CC-work (24)	0.955	0.961	0.960	0.967	0.973	0.961	0.975	0.939	0.975
CC-free (24)	0.938	0.952	0.944	0.945	0.948	0.939	0.966	0.909	0.956
B-factors ($Å^2$) (nonhydrogen atoms) (25)									
All	36.7 (10886)	39.5 (11074)	60.70 (10550)	33.6 (11131)	64.2 (10559)	43.9 (10359)	47.9 (10782)	41.8 (10972)	52.2 (10944)
Main chain	35.5 (5508)	37.7 (5532)	56.9 (5461)	31.7 (5544)	62.5 (5516)	42.7 (5432)	46.0 (5544)	40.2 (5532)	50.1 (5544)
Side chain	38.3 (4916)	41.6 (4939)	62.9 (4826)	35.4 (4956)	66.2 (4926)	45.7 (4754)	50.1 (4876)	43.6 (4950)	55.0 (4951)
solvent	32.5 (364)	38.4 (436)	45.0 (56)	31.1 (519)	50.5 (18)	35.1 (171)	42.2 (293)	40.5 (461)	46.2 (340)

ADP				55.5 (54)			78.3 (54)	68.8 (27)	
ADP*				53.8 (54)					
AMPCP		40.7 (54)							
AMPPCP			97.0 (62)						
AMPPCP*			82.8 (44)						
CoA	33.6 (96)	40.2 (96)	62.1 (96)		58.8 (48)				49.9 (96)
AcCoA					63.0 (51)				
Phosphate							52.6 (10)		54.3 (10)
Estimated coordinate error (Å)	0.29	0.32	0.38	0.25	0.44	0.27	0.30	0.31	0.30
rmsd (bonds) (Å) (26)	0.005	0.007	0.004	0.004	0.006	0.003	0.012	0.004	0.008
rmsd (angle) (°) (26)	0.940	1.056	0.763	0.921	0.880	0.771	1.201	0.970	1.047
Molprobity all-atom clashscore (27)	2.22	2.20	1.81	1.64	2.78	0.83	2.61	2.79	2.15
Rotamer outliers (%)	2.16	0.71	1.04	1.78	2.61	1.12	0.54	2.42	2.50
Ramachandran plot statistics (%)									
Favoured	97.95	97.82	96.60	98.26	97.30	97.02	97.75	97.75	97.68
Allowed	1.90	2.04	3.25	1.60	2.56	2.83	2.10	2.11	2.18
Outliers	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15

Parameters for the outermost shell are shown in parentheses. CC(1/2), percentage of correlation between intensities from random half datasets. Correlation significant at the 0.1 % level (24). CC*, the CC of the full dataset against the true intensities (24). R-work, $R=100 \times \sum_{hkl} \big| |F_{obs}| - |F_{calc}| \big| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively; R-free is equivalent to R-work but is calculated from reflections (5%) that were omitted from the refinement process (28, 29). Description of *B*-factors: values in parentheses corresponds to the number of atoms. Nucleotides marked with an asterisk are bound to the alpha subunit.

Table S2. Opening angle within the ATP-grasp domain of the beta subunit of *ckc*ACD1.

Crystal structure	α-β	α‘-β‘
ACD1-A	0.0 ° §	1.5 °
ACD1-B	23.8 °	13.7 °
ACD1-C	16.6 °	9.3 °
ACD1-D	9.3 °	4.3 °
ACD1-E	18.2 °	12.2 °
ACD1-F	10.8 °	9.3 °
ACD1-G	10.2 °	4.6 °
ACD1-H	12.7 °	7.7 °
ACD1-I	24.6 °	10.2 °

§ The crystal structure *ckc*ACD1-B was set as the reference for the ‚closed‘ conformation. Its opening angle was set to 0.0 °. For the determination of the opening angle the relative orientation of the lid domain (subdomain 3, residues Thr35β – Phe113β) was compared with the subdomain 4 (residues Ser3β – Pro33β / Gly117β – Arg230β) **(**see **Fig. 1**).

Reference List

1. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., ppel, R. D., & airoch, A. (2005) in *The Proteomics Protocols Handbook*, ed. Joh M.Walker (Humana Press, pp. 571-607.

2. Mueller U, Darowski N, Fuchs MR, Forster R, Hellmig M, Paithankar KS, Puhringer S, Steffien M, Zocher G, Weiss MS (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J. Synchrotron. Radiat.* 19(Pt 3): 442-449.

3. Kabsch W (2010) XDS. *Acta Crystallogr. D. Biol. Crystallogr.* 66(Pt 2): 125-132.

4. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* 69(Pt 7): 1204-1214.

5. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* 31(13): 3381-3385.

6. Vagin A, Teplyakov A (1997) MOLREP: An automated programm for molecular replacement. *J. Appl. Cryst.* 30 1022-1025.

7. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* 66(Pt 2): 213-221.

8. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* 67(Pt 4): 355-367.

9. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* 68(Pt 4): 352-367.

10. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* 66(Pt 4): 486-501.

11. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26 283-291.

12. Vaguine AA, Richelle J, Wodak SJ (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D. Biol. Crystallogr.* 55(Pt 1): 191-205.

13. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* 66(Pt 1): 12-21.

14. Ellman GL (1958) A colorimetric method for determining low concentrations of mercaptans. *Arch. Biochem. Biophys.* 74(2): 443-450.

15. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372(3): 774-797.

16. Majumdar R, Guest JR, Bridger WA (1991) Functional consequences of substitution of the active site (phospho)histidine residue of Escherichia coli succinyl-CoA synthetase. *Biochim. Biophys. Acta* 1076(1): 86-90.

17. Fraser ME, Joyce MA, Ryan DG, Wolodko WT (2002) Two glutamate residues, Glu 208 alpha and Glu 197 beta, are crucial for phosphorylation and dephosphorylation of the active-site histidine residue in succinyl-CoA synthetase. *Biochemistry (Mosc).* 41(2): 537-546.

18. Fraser ME, James MN, Bridger WA, Wolodko WT (1999) A detailed structural description of *Escherichia coli* succinyl-CoA synthetase. *J. Mol. Biol.* 285(4): 1633-1653.

19. Fawaz MV, Topper ME, Firestine SM (2011) The ATP-grasp enzymes. *Bioorg. Chem.* 39(5-6): 185-191.

20. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8(2): 127-134.

21. Scott JW, Poole FL, Adams MW (2014) Characterization of ten heterotetrameric NDP-dependent acyl-CoA synthetases of the hyperthermophilic archaeon *Pyrococcus furiosus*. *Archaea.* 2014 176863.

22. Joyce MA, Fraser ME, James MN, Bridger WA, Wolodko WT (2000) ADP-binding site of *Escherichia coli* succinyl-CoA synthetase revealed by X-ray crystallography. *Biochemistry (Mosc).* 39(1): 17-25.

23. Sanchez LB, Galperin MY, Muller M (2000) Acetyl-CoA synthetase from the amitochondriate eukaryote *Giardia lamblia* belongs to the newly recognized superfamily of acyl-CoA synthetases (Nucleoside diphosphate-forming). *J. Biol. Chem.* 275(8): 5794-5803.

24. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084): 1030-1033.

25. Popov AN, Bourenkov GP (2003) Choice of data-collection parameters based on statistic modelling. *Acta Crystallogr. D. Biol. Crystallogr.* 59(Pt 7): 1145-1153.

26. Engh RA, Huber R (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.* A47 392-400.

27. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, III, Snoeyink J, Richardson JS *et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35(Web Server issue): W375-W383.

28. Tickle IJ, Laskowski RA, Moss DS (2000) R_{free} and the R_{free} ratio. II. Calculation of the expected values and variances of cross-validation statistics in macromolecular least- squares refinement. *Acta Crystallogr. D.* 56 (Pt 4) 442-450.

29. Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355 472-475.