Oligomeric structures of metabolic protein assemblies

Structure-function analysis of acetyl-CoA carboxylase and urease

Inauguraldissertation

zur Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Leonie Anton

aus Basel, Basel-Stadt

und Deutschland

2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Maier, Timm
Prof. Dr. Basler, Marek

Basel, den 23. Juni 2020

Prof. Dr. Martin Spiess
Dekan

I Abstract

Oligomeric protein assemblies play a pivotal role in metabolism. The advantages of such complexes are increased stability, protection from degradation and additional options for regulation through allostery. The two enzymes discussed in this thesis utilize oligomerization as a central mechanism, to shield from damaging conditions and control their activity.

Urease is a nickel-metalloenzyme with varying assembly structures expressed in all branches of life except animals and is an essential part of the nitrogen cycle. It catalyzes the breakdown of urea into ammonia, which is used as a nitrogen source. Oligomerization of urease leads to the local increase of ammonia in a concentrated area and contributes to urease stability in extreme environments. Ureases have a huge impact on agricultural practices, but their function as virulence factors in pathogens also makes them important drug targets.

Acetyl-CoA Carboxylase (ACC) catalyzes the first and rate limiting step in the production of fatty acids. Acetyl-CoA is carboxylated by ACC in two consecutive and distinct reactions, forming malonyl-CoA. Fatty acid synthase (FAS) uses malonyl-CoA as a precursor for the formation of fatty acids. Eukaryotic ACCs are multienzymes, which contain all the catalytic domains in a single polypeptide chain. Human ACC is a dimer and oligomerizes into a filament in its most active state. Upregulation of ACC activity has been linked to cancer, metabolic syndrome and viral infections.

The aim of this thesis is the investigation of the mechanisms leading to oligomerization of these essential enzymes by using cryo-electron microscopy (cryo-EM).

In chapter two, the nickel –metalloenzyme urease of *Yersinia enterocolitica* is revealed to form a tetramer-of-trimers. A similar assembly has only been observed for the pathogen *Helicobacter pylori*, where it is described as a crucial factor in survival of acidic environments. Including higher order aberration correction in data processing of cryoEM movies, a reconstruction of *Y. enterocolitica* urease at 1.98 Å resolution was obtained. The structural data revealed an oligomerization loop and a specific alpha-helix as central elements involved in the oligomerization. At better than 2.0 Å resolution, so far only reached by single-particle analysis for four others proteins, salt bridges, alternate conformations and protonation are visualized for the entire protein, and in particular the active site. The two nickel ions in the active site

are at a distance of only 3.2 Å, much closer than in other urease active sites described so far by X-ray crystallography. Frame-wise analysis of cryoEM movies indicates the onset of radiation-damage in the environment of the Ni^{2+} ions during electron-beam exposure as the cause of a distorted overall bimetal center structure.
The mechanism of allosteric regulation of ACC is discussed in chapter three, uncovering the mechanism behind citrate-induced filament formation. Prior and newly collected cryoEM data of ACC filaments were processed using advanced algorithms. Corrections for higher order aberrations resulted in a 3.84 Å resolution reconstruction, which supported identification of a potential citrate interaction region. Site specific point mutants of residues in the suspected citrate binding pocket showed decreased activity in *in vitro* activity assays. Citrate acts on a pocket of positively charged residues in the central domain of the multienzyme. This central domain region has no catalytic properties, but was observed to function as a hinge where large conformational changes occur. Binding of citrate influences the conformational ensemble of the central domain region, creating an interface for other dimers of ACC to dock and form a filament. This filament locks ACC in an active conformation and is the most active state of the protein.
The findings of this thesis show how different proteins employ oligomerization to preserve or augment their function. Ureases grant protection from acidic environments through their assembly while citrate-induced filament formation increases ACC activity.

II Table of Contents

III List of Figures

IV List of Tables

V Abbreviations

AC	ACC central domain
ACC	Acetyl-CoA carboxylase
ACP	Acyl carrier protein
ACYL	ATP-citrate lyase
ADP	Adenosine diphosphate
AH	allophanate hydrolase
AMPK	AMP-activated protein kinase
ATP	Adenosine triphosphate
BC	Biotin carboxylase
BCCP	Biotin carboxyl carrier protein
β-ME	β-Mercaptoethanol
BT	BC-CT interaction domain
BRCA1	Breast cancer type 1 susceptibility protein
BRCT	Breast cancer type 1 susceptibility protein C Terminus
CD	Central domain
CT	Carboxyl transferase
CTF	Contrast transfer function
CPT 1	Carnitine palmitoyltransferase 1
CTP	Cytidine triphosphate
CoA	Coenzyme A
ChREBP	Carbohydrate response element binding protein
DED	Direct electron detector
EM	Electron microscopy
EMDB	Electron microscopy data bank
GCC	Geranyl-CoA carboxylase
FAS	Fatty acid synthase
FSC	Fourier shell correlation
kDa	Kilo Dalton
KO	Knock out
MCC	3-methylcrotonyl-CoA carboxylase
mTORC	Mechanistic target of rapamycin complex
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NBPT	N-(n-Butyl)thiophosphoric Triamid

NCS	Non-crystallographic symmetry
NMR	Nuclear magnetic resonance
o/n	Over night
PAGE	Polyacrylamide gel electrophoresis
PC	Pyruvate carboxylase
PCC	Propionyl-CoA carboxylase
PCP	Peptidyl carrier protein
PDB	Protein data bank
PDC	Pyruvate dehydrogenase complex
PHD3	Proyl hydrolase domain protein 3
PIN1	Proyl isomerase 1
PKA	cAMP-dependent protein kinase
PKS	Polyketide synthase
RMSD	Root mean square deviations
SDS	Sodium dodecyl sulfate
SEC	Size exclusion
SGD	Stochastic gradient descent
SNR	Signal-to-noise ratio
SREBP	Sterol regulatory element binding proteins
TCA cycle	Tricarboxylic acid cycle
TEM	Transmission electron microscopy
UA	urea amidolase
UC	Urea carboxylase

1 Introduction

1.1 Oligomeric enzymes in metabolism

The metabolism of an organism is the total sum of the synthesis of cellular components from energy and chemicals (anabolism) and the breakdown of organic material (catabolism). These chemical transformations are catalyzed by specific proteins called enzymes, which lower the activation energy needed for a chemical reaction and permit chemical turnover under the conditions of life. Enzymes, and all other proteins, are extended linear chains of 20 different types of amino acids. They are involved in metabolism, but also in many other cellular process, such as signal transduction. A subset of all proteins functions as intrinsically disordered proteins without a permanent three-dimensional structure. Most proteins, however, adopt a folded three-dimensional structure linked to their biological function and determined by the sequence of amino acids.

A series of successive chemical reaction steps in metabolism leading to one product or a group of products is known as a metabolic pathway. Controlling the flux through different metabolic pathways is essential for cell survival and proliferation. Oligomerization in the biological context, is the ordered and specific non-covalent binding of multiple proteins in one complex. Homo-oligomers consist of copies of the same protein [1-5] and assemblies of different proteins are hetero-oligomers [5, 6]. Both types of complexes can be either permanent or transient. 80% of *Escherichia Coli* proteins and two thirds of all identified human enzymes are oligomeric, demonstrating their distinctively favorable properties for an organism [7, 8]. The functional advantages for such assemblies are potential for allosteric regulation and stabilization of proteins [9]. An increase in stability of the protein protects against degradation and denaturation [7, 10].

1.2 Urease is a crucial enzyme in ecology and health

Ureases are ubiquitous enzymes responsible for the breakdown of urea into ammonia, which is used as a nitrogen source by many organisms. The oligomeric assemblies of ureases have been suggested to protect the enzyme from degradation and increase its stability. Structural analysis has revealed distinct urease complexes across different species. Ureases were the first enzymes to be

crystallized [11] and the first enzymes to be identified as a nickel-metalloenzymes [12].

1.2.1 Nickel ions as cofactors in redox- and non-redox metalloenzymes

A third of all reported enzymes so far are metalloenzymes, but only a small portion of them contain nickel ions [13]. There are 15 entries for enzymes containing nickel ions on the ExPASy database when searching for enzymes with cofactor Ni^{2+} [14]. Nickel can be present as a mono-, bi- or multimetal centers or as part of a mix of metal ions in an enzyme. Nickel containing enzymes are present in all branches of life except mammals and impact the environment around them [15].
Nickel-metalloenzymes are divided into redox and non-redox enzymes [16]. Ni^{2+} ions have only limited redox chemistry. The redox properties of nickel-enzymes are determined by the ligand and/or surrounding coordinating side chains. Redox nickel-enzymes usually have a sulfide as a ligand or a sulfur containing sidechain (Cys) coordinating the metallocenter [16]. The ligands of non-redox nickel-enzymes contain O and N atoms and use Ni^{2+} as a Lewis acid [15].

1.2.2 Urease is highly efficient and increases the pH of the environment

Ureases are non-redox nickel-enzymes and were the first enzymes to be shown to contain a nickel in their active site [12]. They are ubiquitous in plants, fungi, some bacteria and archaea [13] and are crucial players in environmental (or biological) nitrogen fixation [17]. The substrate of urease is urea, which is a waste product necessary to excrete accumulated surplus of toxic ammonia from an organism [18]. The breakdown of urea occurs in a two-step reaction where the first one is catalyzed by urease [19]. Urea is split into one molecule of ammonia and carbamate at a rate estimated to be 10^{14} to 10^{15} times higher than the non-catalyzed reaction [20]. The second step is the spontaneous hydrolyzation of carbamate into another molecule of ammonia and bicarbonate (Figure 1.2.1) [21].
A consequence of this reaction is a local increase of pH, which has emerged as a problem in agriculture [22]. Ureases are abundant in the soil within bacteria but also as extracellular ureases. Cell-free ureases alkalinize the soil inducing calcium

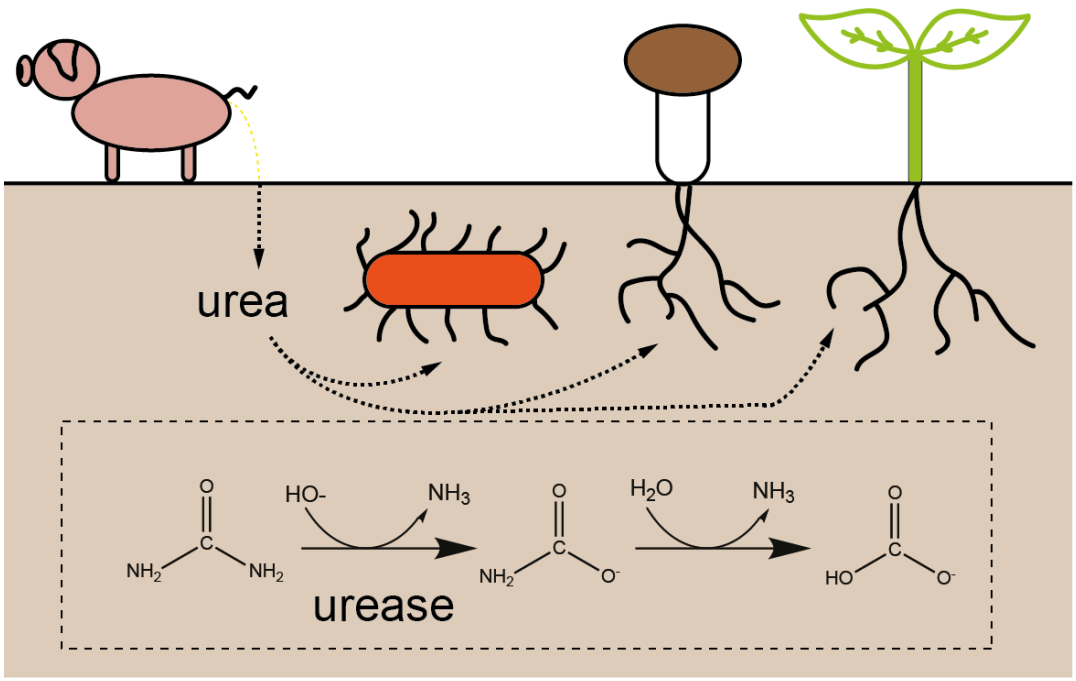

Figure 1.2.1 Schematic overview of urea cycle and urease catalytic reaction

carbonate precipitation, which affects availability of minerals. This is detrimental to agriculture and makes added fertilizers less effective. Up to 70 % of nitrogen from urea fertilizers is used up by the ureases in the soil rather than by the plants. Therefore, research to identify effective urease inhibitors is of high agricultural interest [23]. The increase in local pH caused by the urease reaction is also used by many pathogens to survive in acidic environments [24].

1.2.3 Urease is a virulence factor for different types of pathogens

Bacterial and fungal ureases have emerged as a virulence factors in pathogens and therefore, as important drug targets [25]. The fungal infections cryptococcosis [26] and coccidiomycosis have been connected to the activity of ureases [27]. The bacterium *Proteus mirabilis* causes urinary stones through urine alkalization by the activity of its urease, which can lead to pyelonephritis and catheter encrustation [28].

The impact of bacterial ureases on human pathology is best studied for *Helicobacter pylori*, the causal agent of many gastric ulcers leading to stomach cancer. In *H. pylori,* urease constitutes 10% of all its proteins leading to the neutralization of the acidic environment of the stomach, thereby promoting survival of *H. pylori* [29]. However, there are also other effects on the host which may occur independent of

enzymatic activity, like the delay of apoptosis in neutrophils and gastric epithelial cells [30]. These non-enzymatic influences might cause extra-gastric disease phenotypes like cardiovascular diseases[31]. The ureases of *E. Coli* [31], *Klebsiella pneumonia* [32], *Brucella abortus* [33], *Haemophilus influenza* [34] and *Yersinia enterocolitica* [35] have all been described as virulence factors[31].

1.2.4 *Yersinia enterocolitica* uses urease to survive in acidic conditions

Y. enterocolitica is a small rod-shaped, Gram-negative coccobacillus and the causative agent of yersiniosis. This pathogen causes the majority of all zoonotic illnesses contracted through consumption of contaminated food. It causes mesenteric adenitis, terminal ileitis and gastroenteritis. *Y. enterocolitica* has a high potential to develop multi-drug resistance and several multi-resistant strains have been isolated [36]. Ureases enable *Y. enterocolitica* to passage through the gastrointestinal tract and to survive in the phagosome by protecting the bacterium against the acidic environment [37, 38]. The catalysis of ammonia by urease is notably faster under low pH conditions than in neutral environments. This provides the pathogen with a mechanism by which urease activity is directly stimulated by a drop in pH [35]. *Y. enterocolitica* urease prevents the cytosol from acidifying, but does not affect the pH outside the bacterium, in contrast to observations of acidification for *H. pylori* urease [35]. For *H. pylori,* the ability of urease to preserve activity in acidic environment was attributed to its dodecameric enzyme assembly, demonstrating an impact of the quaternary structure of ureases on physiological activity [39]. The oligomeric structure of *Y. enterocolitica* urease is currently unknown.

1.2.5 Ureases assemble distinct types of oligomers

Plant and fungal ureases present as single polypeptide chains while bacterial and archaeal ureases can contain two (ureA, ureB) or three polypeptide chains (ureA, ureB, ureC) (Figure 1.2.2 a). The chains can also be denoted as alpha, beta and gamma, but do not always refer to the same orthologue in different organisms. Therefore, for the purpose of this thesis we will not use this naming scheme. Conserved domains of ureases from different species share around 50% sequence similarity irrespective of the number of polypeptide chains (see Table 1). The three

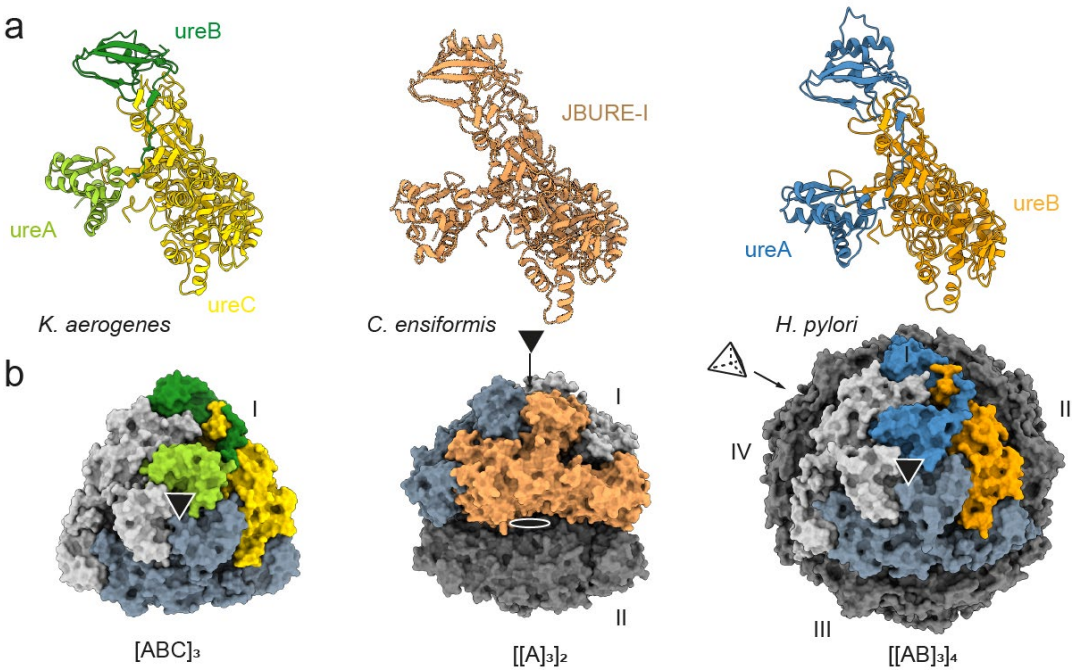

Figure 1.2.2 Architecture of ureases of different domains of life.

a) shows cartoon representation of different ureases with single two and three polypeptide chains encoding the protein. Protein names indicated in color code below cartoon structure. *Klebsiella aerogenes* (PDB ID: 1EJW); *Canavalia ensiformis* (PDB ID: 3LA4); *Helicobacter pylori* (PDB ID:1E9Z) b) surface representation of the same three ureases in their full oligomeric assembly. Same color scheme as in a). Steel and light grey represent the other protomers in one trimer. Dark grey shows the other trimers indicated with roman numerals. Ellipse two-fold symmetry, triangles indicate threefold symmetry axis, tetrahedron indicates tetrahedral symmetry. Assembly type indicated below structure.

dimensional fold adopted by ureases is homologous (see section below) across all species and protein architectures. The two hypothesis on the evolution of ureases are that either a fusion or a splitting event occurred in a common ancestral protein [40]. Phylogenetic inference concluded that it likely consisted in a single 3-to-1 fusion event. This suggests that bacterial three-chained ureases existed before single chained plant and fungal ureases. Fusion could occur by a read-through of the genes, bypassing the stop codons. Notably, two-chained ureases likely emerged from a separate event from the 3-1 fusion and are not evolutionary intermediates [41].

All ureases assemble into oligomeric structures with similar protomer folds, but different species form different higher-order structures [15]. The assembly most commonly found in bacteria is a homo-trimer of the urease with threefold rotational

(C3) symmetry, containing three active sites [42]. A hexamer with threefold dihedral (D3) symmetry is mostly found in plants and fungi [43] (Figure 1.2.2 b). *H. pylori* urease was the first to be described as a dodecamer formed by a tetramer of trimers with tetrahedral (T) symmetry [39]. This oligomeric organization has been hypothesized to increase survival in acidic environments. Low pH can cause the dissociation of Ni^{2+} ions from the active site resulting in subsequent loss of activity [44]. A higher oligomeric state increases the number of active sites in one urease complex to 12. The products accumulate around the complex effectively increasing the local ammonia concentration. The subsequently higher pH microenvironment around the complex ensures the ions are kept in the active site even in acidic conditions [39].

Species compared	**ureA**	**ureB**	**ureC**
C. ensiformis/ Y. enterocolitica	51%	33%	55%
Y. enterocolitica/ H. pylori	52%	52%	58%
H. pylori /C. ensiformis	57%	49%	59%

Table 1.2.1 Global sequence alignment of ureases

Urease sequences were aligned based on their domain boundaries and individually compared. Sequence identity is given in percent. *Canavalia ensiformis* (Jack bean); *Yersinia enterocolitica*; *Helicobacter pylori*.

1.2.6 Nickel bimetal center coordinated by modified lysine is central to urease activity

All ureases have the active site located in their largest domain (ureC or ureB in bacteria), which needs to be accessible for urea. A helix-turn-helix motif called “mobile flap” covers the active site during catalysis and opens to mediate substrate and product diffusion. Urease has a nickel bimetal center coordinated by six side chains: one carbamylated lysine, four histidines and one aspartate. Ni1 is ordered by three of the histidines and Ni2 by the two other histidines and the aspartate. The carbamylated lysine coordinates both Ni ions, which in turn are bridged by a hydroxide ion[31].

When urea enters the active site, it replaces three coordinated waters. First the carbonyl oxygen of urea binds Ni1, which makes it more susceptible to nucleophilic attack. One of the amino nitrogen atoms binds Ni2 forming a bidentate bond, which

is stabilized by an alanine and two histidines brought into proximity through the closing of the mobile flap. A nucleophilic attack by the hydroxide on the carbonyl carbon results in a tetrahedral intermediate and the proton is transferred to one of the amino nitrogens. This yields one molecule of ammonia and, through the binding of the oxygen from the hydroxide, a molecule of carbamate. Finally, the mobile flap opens and releases ammonia and carbamate and the active site fills again with water [31, 45-47]. For this reaction to occur urease need two post-translational modifications: carbamylation of the lysine and addition of Ni^{2+}, which are mediated by accessory proteins of urease.

1.2.7 Post-translation modification and co-factor integration are mediated by accessory proteins

In bacteria, four accessory proteins are tasked with the maturation of apo-urease to fully active enzyme, ure D, E, F and G. The expression of these genes is regulated together with ureA, B, C [21]. The activity of the GTPase ureG has been linked to carbamylation of urease [48] and ureE is a metallochaperone delivering the Ni^{2+} ions [31]. A structure of the *H. pylori* ureDFG (ureD is ureH in *H. pylori*) complex was a breakthrough for the understanding of how hydrogen carbonate for carbamoylation and Ni^{2+} ions get delivered to apo-urease [49]. Molecular dynamics simulations of ureDFG suggest the presence of a set of tunnels within the complex, which would allow for the diffusion of the ions to the active site [50].

The oligomeric assembly of urease allows it to function under a variety of conditions. Stability is protecting it from denaturation while multiple active sites ensure high throughput. Urease catalyzes a single reaction extremely efficiently, but other metabolic pathways require more complex sequences of reactions and tight regulation of the enzyme activity.

1.3 Human ACC is regulated by oligomerization

ACC is an enzyme present in all domains of life and catalyzes the first committed step in fatty acid synthesis. Carboxylation of acetyl-CoA forms malonyl-CoA, which is the substrate for FAS. Activation of ACC from higher eukaryotes is correlated with formation of large filament-like oligomers [51]. ACC is a central and tightly regulated component of the lipogenic metabolic pathway.

1.3.1 Carrier proteins are crucial for the function of multienzymes

The multiple reactions required for a metabolic pathway can be catalyzed by complexes of enzymes or single enzymes. An assembly of enzymes can be transient or permanent and is called a multi-subunit enzyme. Multienzymes are single polypeptides containing several catalytic domains, which execute multiple reactions. Multienzyme complexes consist of multiple enzymes where at least one is a multienzyme [52, 53]. Multienzymes are more frequently found in eukaryotes than in prokaryotes. A multi-subunit enzyme of prokaryotes and a multienzyme of eukaryotes can catalyze the exact same reaction (Figure 1.3.1, adapted from [54]). The tryptophan synthase in *E. coli* is a heterotetramer made of two protein chains, whereas in fungi the proteins are fused and form a homodimer. The advantage of multienzymes over multi-subunit enzymes could be that it ensures a fixed stoichiometry of proteins. In prokaryotes the parallel transcription and translation of proteins regulates their ratio [52].

Analogous to an assembly line, each enzyme catalyzes one step in a sequence of reactions. Prime examples of such enzymes are polyketide synthases (PKSs), which catalyze the formation of secondary metabolites including important antibiotics by condensation of carboxylic acid building blocks [55]. Catalysis of a sequence of reactions within one enzyme complex leads to concentration of metabolites in one location. Intermediates benefit from short diffusion distances or are covalently linked to a protein of the complex[56].

These carrier proteins shuttle intermediates between catalytic sites and are present in a subsection of multienzyme complexes and multienzymes. Carrier proteins can be separate proteins as part of a multi-subunit enzyme or a domain in a multienzyme. Within multienzyme, they are connected via long flexible linkers that allow them to reach the different active sites. Carrier proteins can confer substrate specificity, enhance substrate solubility and protect them from sequestration by other enzymes [57].

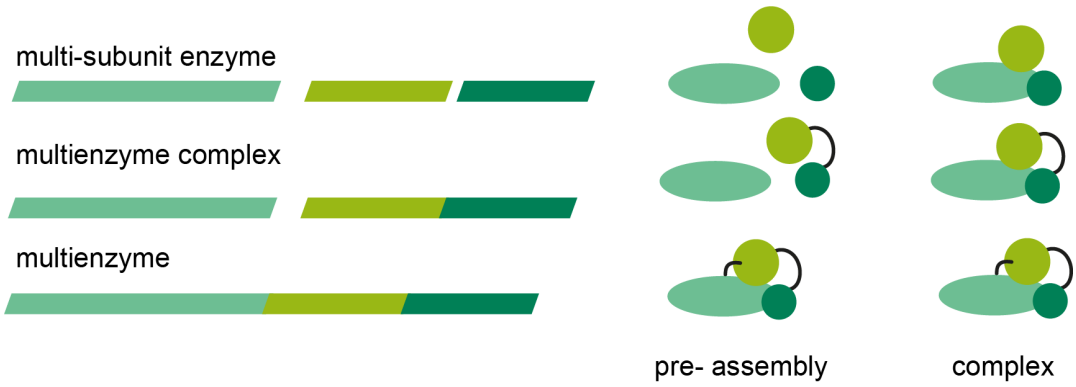

Figure 1.3.1 Multi-subunit enzyme, multienzyme complex and multienzymes
Multi-subunit enzymes (top) consist of multiple proteins functioning as together as one complex. Multienzyme complexes (middle) are composed of single proteins and multienzymes. Multienzymes (bottom) encode all catalytic domains on a single polypeptide chain. Adapted from [53].

Carrier proteins typically need post-translational modifications with a co-factor to bind their respective target intermediate. These co-factors are covalently linked to residues on the opposite side of the linkers to ensure full range of motion. Three co-factors, pantetheine acid, lipoic acids and biotin, define three classes of enzymes [57]. Pantetheine-dependent enzymes have a pantetheine moiety on their carrier protein, called an acyl-carrier protein (ACP). FAS is a pantetheine-dependent enzyme with a phosphopantetheine co-factor attached at the ACP [58]. The pyruvate dehydrogenase complex (PDC) has a lipoic acid moiety on its lipoyl domain. It enables binding of the substrate pyruvate and is in the class of lipoic acid dependent enzymes [59]. The biotin-dependent carboxylases have a biotin moiety attached to their carrier protein, the biotin carboxyl carrier protein (BCCP), at a highly conserved Met-Lys-Met motive [60]. The moiety allows the BCCP to transport carboxyl groups between the different reaction sites [61].

1.3.2 Biotin-dependent carboxylases have a conserved reaction mechanism

Biotin-dependent carboxylases are ubiquitous enzymes which are involved in fatty acid, amino acid and carbohydrate metabolism, as well as urea utilization and polyketide biosynthesis [62-66]. All biotin-dependent carboxylases have the same three components: biotin carboxylase (BC), BCCP and carboxyl transferase (CT). They can be organized as multienzymes or multi-subunit enzymes. Biotin-dependent carboxylases catalyze a conserved two-step reaction: in the first step, a

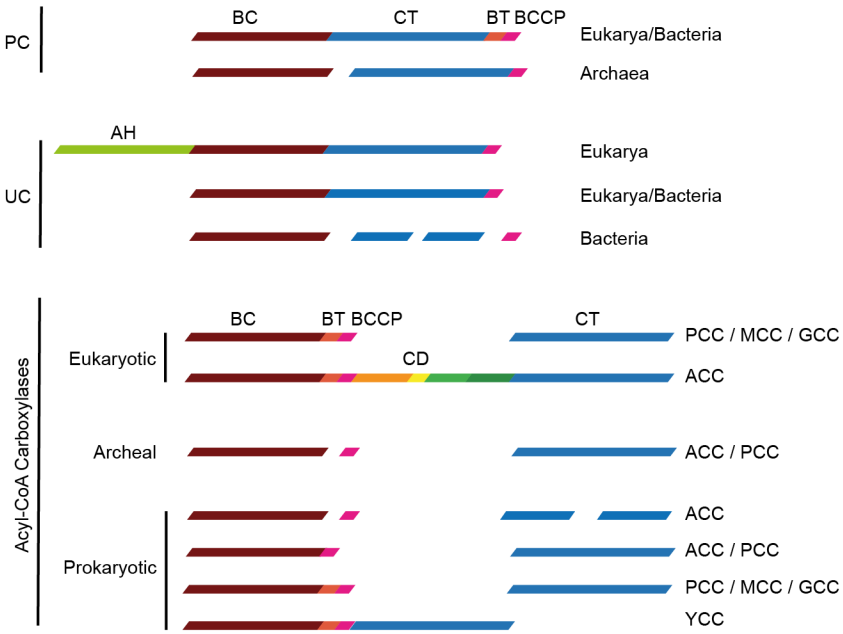

Figure 1.3.2 Domain arrangement of biotin-dependent carboxylases

Linear representation of domain arrangements in biotin-dependent carboxylases. The carboxylases are divided by their families and labeled according to enzyme and class. The domain color scheme will be used throughout the introduction. Adapted from [68].

carboxyl group of a bicarbonate molecule is transferred by the BC to the biotin moiety of the BCCP in a ATP-dependent manner. The BCCP then carries the carboxyl group to the CT, which transfers the carboxyl group onto the substrate as the second step [67, 68].

Biotin-dependent carboxylases are classified by their substrate, which leads to three main classes: Urea carboxylases (UC), pyruvate carboxylases (PC) and acyl-CoAs carboxylases [61] (Figure 1.3.2, adapted from [69]). The PC carboxylates pyruvate to form oxaloacetate to resupply the tricarboxylic acid cycle (TCA). In mammals, PC also catalyzes the first committed reaction in gluconeogenesis. PC is therefore crucial in intermediary metabolism and is localized in the mitochondria in mammals and in the cytosol in fungi [62, 70]. There is some evidence that PC has a role in virulence during infection with *Listeria monocytogenes* [71]. As an alternative to the urease reaction described in the previous chapter, some bacteria, fungi and algae use urea amidolyase (UA) to access nitrogen. UA contains UC to carboxylate urea to allophanate and allophanate hydrolase (AH) to convert it into ammonium. They can organize as multienzymes or multi-subunit enzymes [72, 73]. Acyl-CoA carboxylases are further classified according to the specific acyl-CoA they carboxylate. Propionyl-CoA carboxylase (PCC), 3-methylcrotonyl-CoA carboxylase (MCC), geranyl-CoA carboxylase (GCC) and acetyl-CoA carboxylase (ACC,

covered in the next chapter) [61]. Propionyl-CoA is carboxylated to D-methylmalonyl-CoA and is found across many organisms from bacteria to humans. PCC is central for the catabolism of β-branched amino acids, cholesterol side chains and fatty acids with odd number of carbon atoms [66, 74]. MCC converts 3-methylcrontonyl-CoA to 3-methylglutaconyl-CoA, which is essential for the catabolism of leucine and isovalerate [75]. GCC like MCC also carboxylates the γ carbon of an unsaturated acid and are therefore often grouped together. Carboxylation of geranyl-CoA yields γ-carboxygeranyl-CoA, which is involved in bacterial metabolism of geranyl group and other acyclic terpenes. GCC can also accept the MCC substrate 3-methylcrontonyl-CoA, but not vice versa [76].

1.3.3 Human ACC has two isoforms with distinct functions in the cell

Bacterial ACC is organized as a multi-subunit enzyme or a multienzyme complex, where the BC and BCCP are fused. There is genomic evidence that some prokaryotes have multienzyme ACC plus an additional BC-CT interaction domain (BT). Eukaryotic ACC is a multienzyme and contains a BT domain as well as an additional non-catalytic domain called central domain (CD) not found in other biotin-dependent carboxylases [61].

ACC1 and ACC2 are the two isoforms of higher eukaryotic ACC. They are encoded by the genes ACACA and ACACB, which share 74% sequence identity [77]. While both isoforms function as 500 kDa dimers, ACC1 can assemble higher order oligomers [60]. Both produce malonyl-CoA from acetyl-CoA, but in different tissues and with different downstream uses.

ACC2 is found in heart and skeletal muscle and is anchored to the mitochondrial outer membrane with a unique N-terminal region [78]. Production of malonyl-CoA inhibits carnitine palmitoyltransferase 1 (CPT1) and therefore fatty acid oxidation [79]. ACC1 is mostly found in adipose tissue, mammary glands and lipogenic tissue where *de novo* fatty acid production is required. The produced malonyl-CoA is the substrate for FAS, therefore ACC1 catalyzes the first committed step in fatty acid production [65]. Fatty acids can be used as precursors for lipids and membrane anchors, energy storage and as secondary messengers [58]. The oxidation and production of fatty acids are essential functions for all domains of life, making the regulation of ACC is crucial to the metabolism.

1.3.4 ACC is regulated by an interplay of transcription, phosphorylation and metabolites

ACC of higher eukaryotes is regulated at transcriptional and post-translational level to tie production of fatty acids to metabolic demand. Transcriptional regulation involves sterol regulatory element binding proteins (two isoforms, SREBP1a and SREBP1c) and carbohydrate response element binding protein (ChREBP) as transcription factors [80-82]. The master regulators of metabolism, mechanistic target of rapamycin complex 1 and 2 (mTORC1 and 2), regulate expression of SREBP based on nutrient availability [83, 84]. ACC is a 265 kDa monomer and functions as a dimer of over 500 kDa. Given the large size of the active ACC dimer, regulation through transcription only would be extremely costly for the cell.

Citrate functions as a feedforward activator increasing ACC activity [85]. As observed more than five decades ago, citrate activation leads to the formation of filaments of ACC in higher eukaryotes of up to one micron in length [51]. Citrate is a product of the TCA cycle and is converted into acetyl-CoA by the ATP-citrate lyase (ACYL) in the cytosol. Acetyl-CoA is the substrate of ACC and citrate functions as an allosteric activator of catalysis. However, how citrate binds ACC and how it exerts its function is still unclear.

The product of ACC, malonyl-CoA and the downstream intermediate of FAS, palmitoyl-CoA are feedback inhibitors, signaling the surplus of fatty acids [63, 77]. Palmitoyl-CoA has detergent like properties, which leads to unspecific binding of proteins and could be part of a more general inhibition of metabolic pathways [86]. However, blocking of unspecific binding sites by the addition of BSA does not prevent palmitoyl-CoA binding and ACC inhibition. Addition of citrate, palmitoylcarnitine and decanoylcarnitine can also reverses inhibition by palmitoyl-CoA [87] (Figure 1.3.3).

Inhibition of ACC is also induced by phosphorylation of specific residues. In ACC1 AMP-activated kinase (AMPK) phosphorylates Ser80, Ser1201 and Ser1206, while cAMP-dependent kinase (PKA) phosphorylates Ser78 and Ser1201. Phosphorylation of the conserved residues Ser80 and Ser1201 cause an especially strong reduction of activity in ACC1 [88, 89]. Ser222 in ACC2, the Ser80 equivalent, is phosphorylated by AMPK in order to decrease catalysis [90]. Ser80 and Ser222 are located in the BC domain of ACC1 and ACC2, respectively. Phosphorylation of these residues localizes them to a binding pocket, which is also the target of inhibitors (see next section). Yeast ACC is regulated through phosphorylation of Ser1157 by Snf1, a homologue of AMPK [91-93].

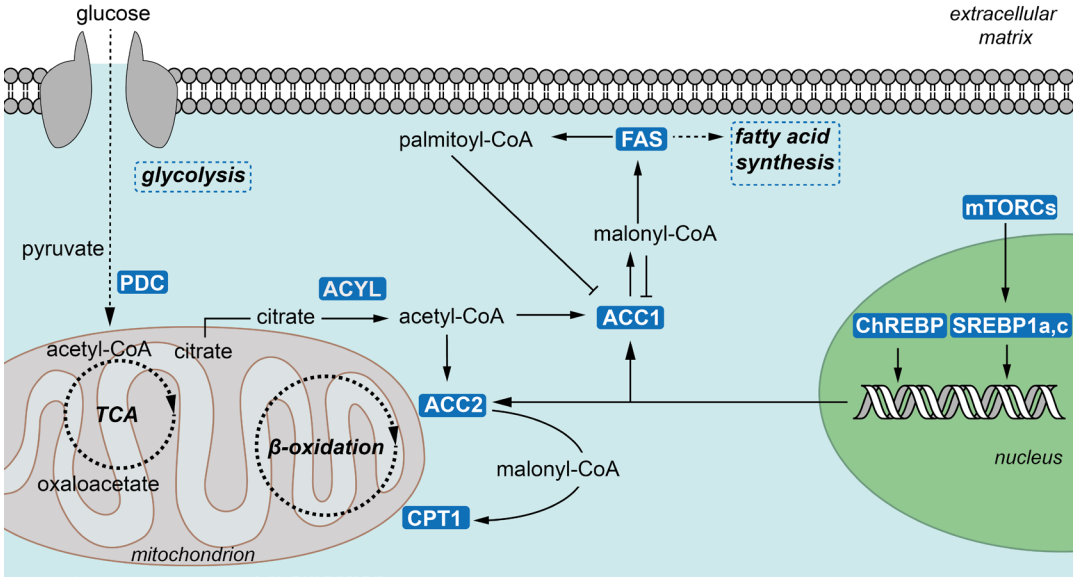

Figure 1.3.3 Regulation of acetyl-CoA carboxylase

Scheme of regulation of human acetyl-CoA carboxylase. PDC: pyruvate dehydrogenase complex; TCA: tricarboxylic acid cycle; ACYL: ATP-citrate lyase; ACC: acetyl-CoA carboxylase; CPT1: carnitine palmitoyltransferase 1; FAS: fatty acid synthase; mTORC: mechanistic target of rapamycin; ChREBP: carbohydrate response element binding protein; SREBP: sterol regulatory element binding protein.

ACC is regulated by several protein interaction partners. Mig12 is a cytosolic protein of 22 kDa size and can lower the citrate threshold for activation and induce filament formation by binding to ACC1 and ACC2 [94]. Mig12 and the paralog Spot14, can form a complex which attenuates filament formation of both ACC1 [95] and ACC2 [96]. If ACC2 forms filaments in physiological conditions in presence of the N-terminal membrane anchor is unclear. Prolyl hydrolase domain protein 3 (PHD3) is part of the metabolic response to changing cellular conditions. It hydroxylates a proline of ACC2 increasing its activity and inhibiting fatty acid oxidation [97]. Increased levels of proyl isomerase 1 (PIN1) are associated with malignant tumors and upregulation of ACC1. Pin1 can bind and stabilize the C-terminal domain of ACC1 and, by preventing protein degradation, lead to an increase in protein levels of ACC1. Pin1 also inhibits AMPK and prevents phosphorylation and further activation of ACC [98].

Breast Cancer Susceptibility gene 1 (BRCA1) is one of the most important tumor-suppressor genes in breast cancer. BRCA1 functions in DNA repair, transcription and ubiquitination. The tandem domain (BRCT) in the C-terminal region of BRCA1 binds ACC1 and the interaction is abrogated by mutations associated with breast cancer [99]. The binding of monomeric BRCT to a specific phosphopeptide (1258-

DSPPQ-pS-PTFPEAGH-1271) of ACC1 in a 1:1 ratio was revealed in a crystal structure [100]. The tight control of ACC activity is necessary for the whole metabolism and any dysregulation leads to detrimental outcomes.

1.3.5 ACC is a drug target for treatment of different metabolic diseases

The effect of ACC on the organism as a whole has been demonstrated in knock out (KO) studies. ACC1 KO in mice is embryonically lethal, but KO of ACC1 in liver showed no obvious phenotype. *De novo* fatty acids synthesis was reduced, while fatty acid oxidation was unaffected [101]. KO in adipose tissue caused prenatal growth impairment and postnatal issues in skeletal growth in mice [102]. The induction of fatty acid synthesis by ACC1 is essential to proliferation of cells in growth and development. ACC2 KO mice were described to have a normal life span and fertility, but depending on KO strategy and genetic background additional phenotypes were observed. Some mice displayed reduced fat mass, increased energy expenditure and improved insulin sensitivity and were protected against obesity on a high-fat/high-carbohydrate diet [103-106]. The inhibition of fatty acid oxidation by ACC2 is a crucial factor in maintaining fatty acid homeostasis in the cell.

ACC is implicated in two major health disorder: metabolic syndrome and cancer. Metabolic syndrome includes type 2 diabetes, obesity, cardiovascular diseases and arteriosclerosis [107-109]. ACC1 and ACC2 are both attractive drug targets for metabolic syndrome, however no drug has made it to the market yet. There are promising studies of inhibitors targeting both isoforms, which decreased *de novo* fatty acid synthesis and upregulation of fatty acid oxidation. A new inhibitor improved insulin sensitivity, reduced weight gain and reduced hepatic steatosis [110]. ACC2 selective inhibitors mimicking the ACC2 KO, showed serious side effects and significantly affected metabolic pathways outside of ACC2. Reduction of off-target effects improves performance of ACC2 selective inhibitors [111]. Soraphen A was first discovered as a potent antifungal drug. It has also inhibitory effects on other eukaryotic ACCs [112]. Soraphen A was able to inhibit *de novo* lipogenesis and improve insulin sensitivity in rats, but poor water solubility and bioavailability limited its use as a drug[113].

Lipid metabolic pathways have been heavily implicated in the progression of different cancers for a long time. This includes ACC overexpression found in types of lung, liver, breast and prostate cancer [83, 114-116]. The lipogenic pathway is a

very attractive drug target for tumors. Cancer cells depend on lipid anabolism for proliferation to a greater extent than healthy cells [115, 117]. RNA interference mediated knock down of ACC1 and FAS resulted in apoptosis of breast cancer cells [116]. Downregulation of ACC2 by tumors increases fatty acid oxidation, which provides acetyl-CoA for the TCA cycle and energy in acidic conditions. Most cancer cells are in a low pH environment and can reach a state where both fatty acid synthesis and fatty acid oxidation are concomitantly active [97, 118]. Soraphen A was able to inhibit proliferation of a type of prostate cancer cells [113] and an ACC1 selective inhibitor was recently reported as a potential cancer treatment [119]. An ACC inhibitor developed initially for cancer treatment had an unforeseen effect: it reduced *de novo* lipogenesis in sebaceaous glands and reduced acne [120]. The ACC inhibitor GS0976 is in clinical trials as part of a combination treatment for non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) [113].

In recent years more evidence has emerged that ACC plays a crucial role in infectious diseases. Zika virus, Porcine reproductive and respiratory syndrome virus, West Nile virus and dengue virus all showed potential for treatment with ACC inhibitors [121-123]. *Plasmodium falciparum*, the causative agent of malaria, upregulates host ACC to acquire essential lipids it cannot produce itself [124].

Two Alzheimer drug candidates were shown to increase the acetyl-CoA concentration in cells by inhibiting ACC, which was neuroprotective and increased acetylation of critical histones [125]. Two ACC inhibitors in plants have been used as herbicides for a long time [126, 127]. There are also several drugs being studied, indirectly affecting ACC through AMPK or PKA activation [128-130].

Balancing the trade-offs between efficient inhibition and off-target effects on the metabolism will be crucial for the future development of ACC inhibitors. Essential to the understanding of potential inhibitors are detailed structural descriptions of ACC.

1.3.6 Structural characterization of ACC highlights the relevance of dynamics for function

Even though the research field of eukaryotic ACC has garnered much attention for its role in disease, solving a complete structure of the full multienzyme has been challenging. ACC is too large to investigate via NMR and highly dynamic, which is challenging for crystallization. Structures of single domains have furthered our understanding of how ACC and its inhibitors work. Only recently, full length

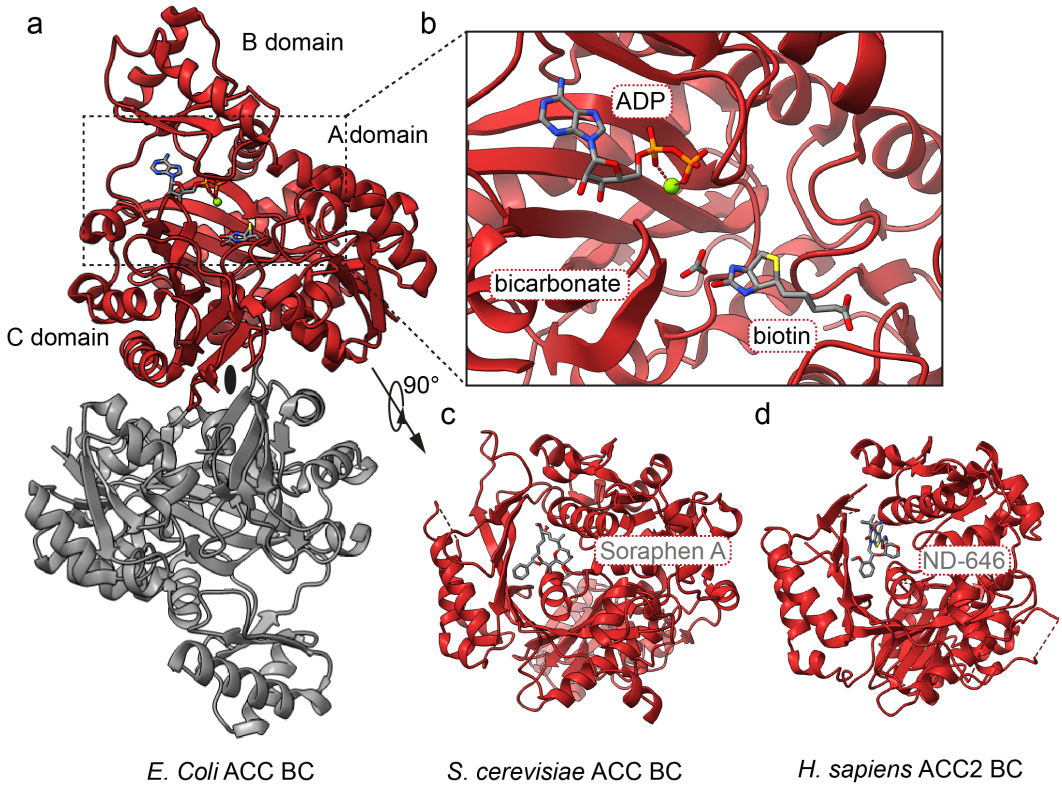

Figure 1.3.4 BC domains of ACC

a) Dimeric BC domains with two-fold symmetry indicated by an oval. BC domain has A, B and C subdomains indicated for red protomer. b) Inset shows enlarged active site with biotin, ADP and bicarbonate (PDB ID:3G8C). c) 90° rotation from view of top protomer in a) for BC monomer from S. cerevisiae bound with inhibitor soraphen A (PDB ID:1W96). d) Inhibitor ND-646 inhibiting human BC domain (PDB ID:5KKN). View same as in c).

structures of ACCs were solved to high enough resolution to unambiguously identify all the domains.

The BC domain of *E. Coli* was the first solved x-ray structure of an ACC domain and crystallized as a dimer [131] (Figure 1.3.4 a). A structure of BC with ADP, biotin and bicarbonate gave insight into the mechanism for biotin carboxylation [132] (Figure 1.3.4 b). Structural analysis of BC bound by Soraphen A and ND-630 inhibitor compounds revealed their binding sites in the dimer interface, incompatible with activity [110, 133, 134] (Figure 1.3.4 c, d). Eukaryotic and prokaryotic BC domains are structurally very similar and have three subdomains A, B and C, with B being the most flexible. BC domains of *S. cerevisiae* and human ACC have only been isolated as monomers and are catalytically inactive as such. This is due to the interface in the monomer being incompatible with dimerization, as was seen in the structure of full length *S. cerevisiae* ACC [135]. This demonstrates conformational flexibility as a necessary part of ACC activity.

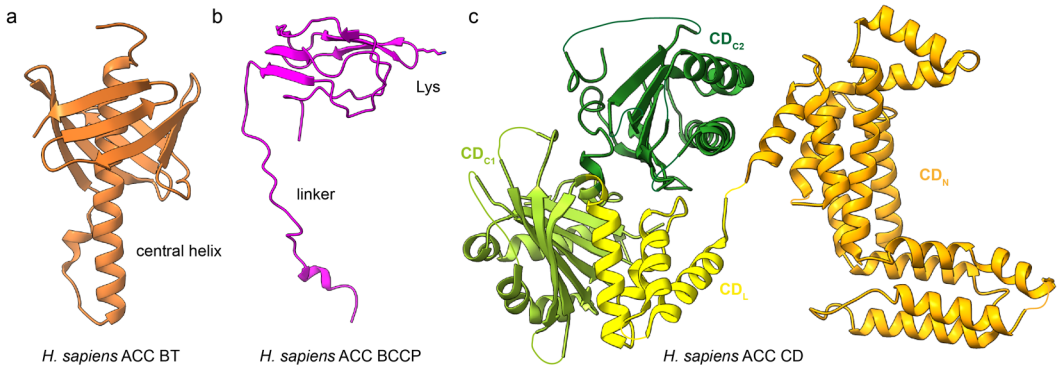

Figure 1.3.5 BT, BCCP and CD domains of ACC

a) BT domain of ACC with central helix indicated (PDB ID:6G2D). b) BCCP domain with flexible linker and conserved lysine indicated (PDB ID:6G2D). c) CD domain with subdomains indicated (PDB ID: 5I87).

The BT domain has a rarely observed backbone fold and was first described in the structure of a PCC. The central helix of the BT is surrounded by an antiparallel eight stranded β-sheet [74] (Figure 1.3.5 a). The BCCP core structure is a β-hairpin, which consists of four antiparallel β-strands. The β-strands are folded around a hydrophobic core. The conserved biotinylation motif Met-Lys-Met is located at the tip of the hairpin. At the opposite end are the linkers connecting BCCP to the rest of the multienzyme [136-138] (Figure 1.3.5 b).

The CD is only present in eukaryotic ACCs and there are interesting differences between regions of human CD and fungal ACC central domain (AC). The domain region consists of subdomains of similar fold: the CD_N (N-terminal)/AC1 and AC2, CD_L (linker)/ AC3, CD_{C1} (C-terminal 1)/AC4 and CD_{C2} (C-terminal 2)/ AC5, respectively. CD_N is composed of a four-helix bundle connected by six helices to a helical hairpin, all ordered in a horseshoe shape. CD_L is a small irregular four-helix bundle bridging CD_N to CD_{C1}. CD_{C1} and CD_{C2} adopt the same fold of a six-stranded β-sheet, flanked by two bent helices. Comparison of human CDs and fungal ACs show that, even if the folds of the subdomains are highly similar, the overall architecture of the central domain region is very different. The subdomains rotate and move relative to each other, influencing the conformational ensemble, identifying the CD as a highly dynamic region of ACC [135, 139] (Figure 1.3.5 c).

There is low sequence identity of CT among different acyl-CoA carboxylases, because it specifies the substrate. However, the overall fold of CT domains is very similar across all types of acyl-CoAs, but can present as different higher oligomeric

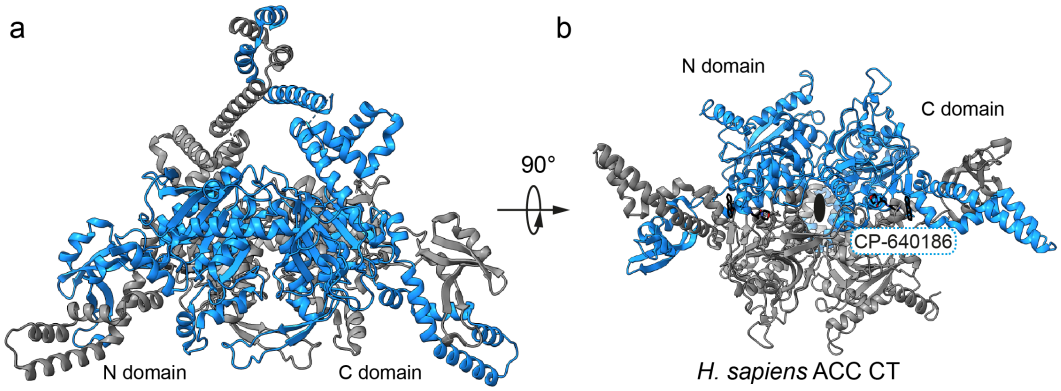

Figure 1.3.6 CT domain of ACC

a) Dimer of human CT domain of ACC one protomer in blue and one in grey. N and C domain indicated for blue protomer. (PDB ID:4ASI). b) 90° rotated dimer with inhibitor CP-640186 shown in black. N and C domains are indicated for blue protomer (PDB ID:3FF6).

assemblies [74, 75]. The CT forms a dimer where each protomer has a N and C domain (Figure 1.3.6 a). Dimerization is necessary for activity, because the active site is located in the highly hydrophobic interface [140]. Many of the small molecule inhibitors bind the active site of CT domain [113] (Figure 1.3.6 b).

X-ray structure of full length yeast ACC revealed that the dimer adopts a very different structure from the other biotin-dependent carboxylases. The CD region is crucial in positioning the CT and BC domains relative to each other so both catalytic sites can be reached by the BCCP (Figure 1.3.7 a). Overlay of full length ACC crystal structure with CD structures, shows that there are substantial rotations of the CD subdomains, allowing for dimerization of the BC. There are also extensive conformational changes in the BC domain itself, making dimerization and therefore activity, possible. This dimerization is not compatible with binding of Soraphen A and similar compounds [135].

Human ACC has not been crystallized and negative stain EM, showed a highly heterogeneous mixture of conformational states of inactive ACC1 [141]. The revolution in cryo-electron microscopy (cryo-EM, see next section) has allowed for the structure determination of many large oligomeric proteins, including ACC filaments.

Citrate-induced filaments of human ACC (ACC-Cit) assemble a helix of ACC dimers, which have a highly similar structure to full length yeast ACC. The helix has a 120° twist and 154 Å rise and is formed by the dimers stacking together via interaction through their CD domains. The conformation of the CD_N in the filament

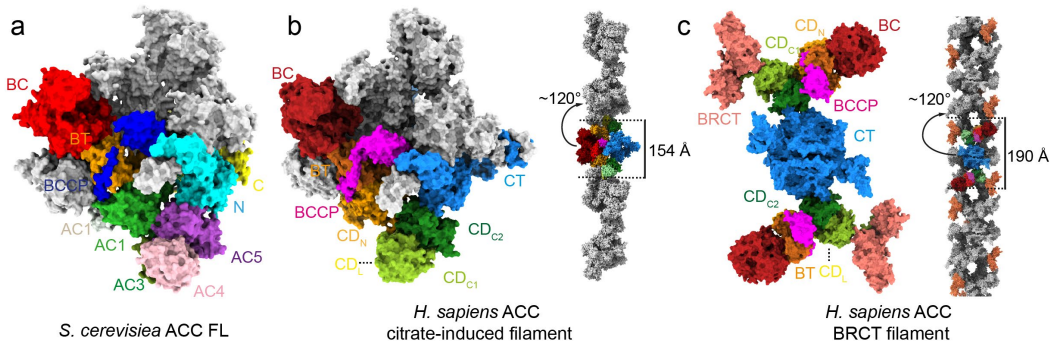

Figure 1.3.7 Structure of full length fungal ACC and human ACC filaments

a) Crystal structure of full length S. cerevisiae ACC. Domains are labeled and colored according to original publication (PDB ID: 5CSL). b) ACC dimer shows domains colored and labeled with color code. ACC filament with one dimer in domain colors and rest in grey. Rise (154 Å) and twist (120°) are indicated. c) ACC dimer bound by BRCT domain of BRCA1. Domains are colored and labeled with color code. ACC-BRCT filament with one dimer in domain colors and rest in grey. Rise (154 Å) and twist (120°) are indicated.

interface is different from fungal yeast ACC, explaining why fungal ACC cannot form filaments. The BC domains are dimerized and the BCCP is able to reach both catalytic sites (Figure 1.3.7 b). Phosphorylation of Ser80 in the BC domains would not be compatible with this dimerization. The filament state locks ACC in a perpetually active and closed conformation.

The binding of the BRCT tandem domain to full length ACC induces a filament (ACC-BRCT) distinct from ACC-Cit. The BRCT binds a phosphorylated loop of the CD domain and prevents its movement. This locks ACC in an inactive and open conformation within a filament. The BC domains are monomeric and BCCP is not able to reach both BC and CT domains (Figure 1.3.7 c). These two filaments highlight once again the dynamic nature of ACC and how crucial large conformational changes are to its regulation and function [141].

Citrate-induced filament formation is tightly linked to increasing activity. However, the mechanism by which citrate induces the necessary conformational rearrangements in ACC is still not resolved.

1.4 CryoEM as a tool to study macromolecular assemblies

Electron microscopy has long been used to image and study large protein complexes. Since the 'resolution revolution', cryo-EM has become essential for structure determination of macromolecular assemblies [142]. The advantages of cryo-EM are low amount of sample required, the physiological state in which proteins are imaged and the ability to separate heterogeneous samples into their components. The impact cryo-EM has made on science was recognized in 2017 with the Nobel prize in chemistry being awarded to Jaques Dubochet, Joachim Frank and Richard Henderson *'for developing cryo-electron microscopy for the high resolution structure determination of biomolecules in solution'*.

The number of structures solved by cryo-EM started exponentially increasing around 2010. At the time of writing this thesis, there are over 10'000 deposited cryoEM maps according to the electron microscopy data bank. The average resolution has significantly decreased from 15 Å in 2010 to 5.6 Å in 2019 [143]. While writing this thesis several <2Å structures came out and as of May 2020 the highest resolved reconstruction is apoferritin at 1.22 Å [144]. There are also now structures available of proteins and oligomers, which have been historically difficult to resolve using x-ray crystallography [142]. This includes membrane proteins, large assemblies and filaments.

A large driver in the surge of structures of large oligomeric complexes is the constant development in the optimization of sample preparation, image acquisition and computing. The first giant leap occurred with commercially available direct electron detectors (DEDs), that enabled fast data collection while preserving high resolution information [142, 145]. DEDs enables counting of single electrons by using large pixels to ensure only one electron excites one pixel. A very thin radiation-hard camera chip prevents scattering of electrons within the detector, which prevents image blurring.

New cryo-EM algorithms have improved calculation and refinement of 3D reconstructions. Beam-induced motion of particles during recording blurs the final image. Motion correction software realigns the particles within a series of images of the same particles to account for their movement [146-148]. Observation of additional anisotropic image motion and better modelling of particle trajectories aided in the development of more robust motion correction algorithms [149, 150].

The contrast transfer function (CTF) describes which information of the sample is in the recorded image. It also reflects optical aberrations of the microscope system. The more accurately the CTF can be estimated, the more information can be

extracted from the data [151]. To improve calculation of the CTF, there are several new and complementary algorithms. Ice thickness can vary across one micrograph and affects the local CTF values. A new approach calculates a variable CTF landscape instead of a uniform CTF correction across one micrograph. Particles distributed in ice at different heights have different defocus values and the accurate approximation of defocus per particle directly influences the accuracy of the CTF [152].

The increase in achievable resolution through cryo-EM has led to the limitations of the optical systems coming into focus. There are three proposed optical effects: symmetrical aberration, anti-symmetrical aberration and magnification anisotropy. In the newest releases of some of the cryo-EM programs these higher-order aberrations can be taken into consideration in the calculation of the CTF [153]. The cryo-EM refinement algorithms are still under constant development to improve quality of reconstructions and get more information out of collected data [150, 154]. Software capable of separating different conformational states of a protein complex are further pushing the limits of cryo-EM [155-157].

The development in cryo-EM remains ongoing with new DEDs having better signal-to-noise ratios (SNR), larger fields of view and faster data collection. Together with more efficient computing and advanced algorithms, cryo-EM will further deepen our insight into structure and function of previously unobserved macromolecular assemblies. Major challenges for cryo-EM are still macromolecular complexes with relatively small heterogeneity and flexibility in single regions or domains.

1.5 Aims of the thesis

The aim of this thesis is to investigate oligomerization as an evolutionary asset of enzymes for enhanced stability and advanced regulation. Oligomerization of the two systems studied here, urease and ACC plays an important role for human health. Their giant protein assemblies are attractive drug targets for infectious diseases and cancer, respectively.

The oligomeric assembly of urease in other bacteria has been linked to enhanced stability and virulence, but the structure of the three-chain urease of the pathogen *Y. enterocolitica* is unknown.

Citrate-induced formation of filamentous oligomers was recognized as a critical component of the regulation of the essential ACC enzyme over 50 years ago, but the mechanistic link between allosteric activation, filament formation and enzymatic turnover remained enigmatic.

In my thesis, I address these questions using state-of-the-art tools of structural biology in combination with functional analysis. Specific aims of these thesis work are:

Aim 1.1: Reveal how *Y. enterocolitica* urease assembles oligomers at atomistic level.

Aim 1.2: Understand the sequence and structural determinants for oligomer formation and stabilization in *Y. enterocolitica* urease.

Aim 1.3: Explore potential for structure-based drug discovery for bacterial pathogens by single particle cryoEM.

Aim 2.1: Reveal the binding mechanism of citrate that induce activated acetyl-CoA carboxylase filaments.

Aim 2.2: Validate and analyze the process of allosterically-induced filament formation by mutational functional analysis.

The central approach that proved successful for addressing the questions raised above is cryo-electron microscopy single particle analysis, in combination with biochemical analysis.

2 High resolution cryo-EM structure of urease from the pathogen *Yersinia enterocolitica*

This research was published in Nature Communications

Ricardo D. Righetto*, Leonie Anton*, Ricardo Adaixo*, Roman Jakob, Jasenko Zivanov, Mohamed-Ali Mahi, Philippe Ringler, Torsten Schwede, Timm Maier and Henning Stahlberg

Nat Commun. 2020; 11: 5101

*Authors contributed equally

High resolution cryo-EM structure of urease from the pathogen Yersinia enterocolitica

2.1 Contribution to the manuscript

I build the model based on available crystallographic models and refined them against the cryo-EM map. I did sequence alignments and analyzed the conservation. I analyzed the interfaces, quality of the model and the active site residues and metal ions. I wrote the parts of the manuscript pertaining to model, interfaces and active site. I made the corresponding figures and the movie showing radiation damage on the active site. Contributions of the other authors is indicated in author contribution section.

2.2 Abstract

Urease converts urea into ammonia and carbon dioxide and makes urea available as a nitrogen source for all forms of life except animals. In human bacterial pathogens, ureases also aid in the invasion of acidic environments such as the stomach by raising the surrounding pH. Here, we report the structure of urease from the pathogen *Yersinia enterocolitica* at 2 Å resolution from cryo-electron microscopy. *Y. enterocolitica* urease is a dodecameric assembly of a trimer of three protein chains, ureA, ureB and ureC. The high data quality enables detailed visualization of the urease bimetal active site and of the impact of radiation damage. The obtained structure is of sufficient quality to support drug development efforts.

2.3 Introduction

Ureases are nickel-metalloenzymes produced in plants, fungi, and bacteria, but not in animals. They facilitate nitrogen fixation by metabolizing urea, but in some pathogenic bacteria they serve a dual function and consume ammonia to promote survival in acidic environments. This function is vital during host infection where the bacteria have to survive the low pH of the stomach1, and of phagosomes in host cells[15, 35, 158]. The relevance of ureases in early stages of infection renders them attractive targets for novel anti-microbials.

Ureases catalyze the breakdown of urea into ammonia and carbamate at a rate 10^{14}–10^{15} times faster than the non-catalyzed reaction (Figure 2.3.1 a), and are arguably the most efficient hydrolases[15, 159]. In a second non-catalyzed step, carbamate is spontaneously hydrolyzed to yield another molecule of ammonia as well as one molecule of bicarbonate[159]. Jack bean urease was the first enzyme to be crystallized, offering evidence that enzymes are proteins[11] and was the first metalloenzyme to be shown to use nickel in its active site[12, 160]. All ureases characterized to date share the architecture of their active site[31, 159]. Two Ni^{2+} ions in the active site are coordinated by a carbamylated lysine, four histidines, and one aspartate. The two metal ions are bridged by a hydroxide ion, serving as a nucleophile[161, 162]. Urea first interacts with Ni(1) through its carbonyl oxygen, and following a conformational change of a mobile flap covering the active site, one of the amino nitrogens then binds to Ni(2), making the resulting metal-chelate urea molecule more available for nucleophilic attacks. Subsequently, the bridging hydroxide acts as the nucleophile attacking the urea C atom, and a proton is then transferred from the hydroxide ion to the distal NH_2 group, yielding a nascent

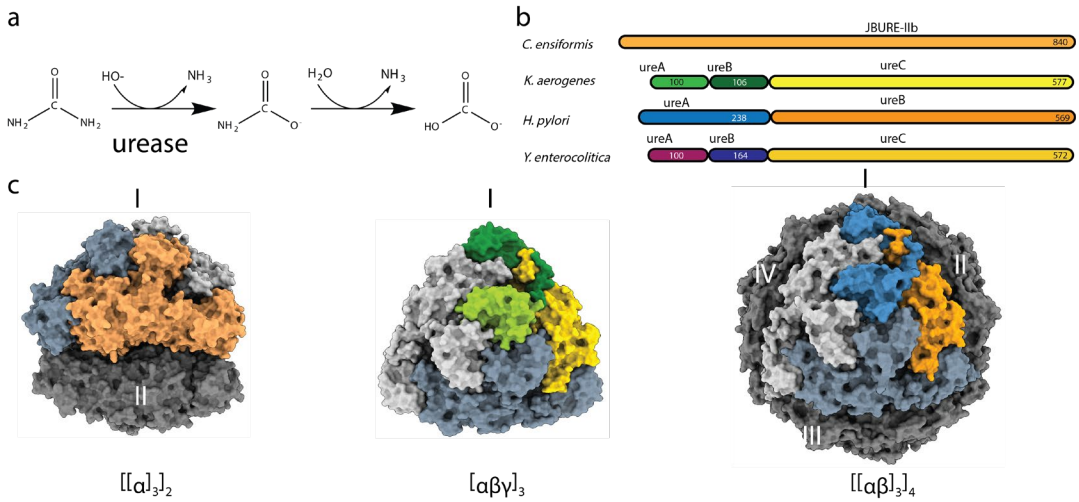

Figure 2.3.1 Protein architectures and oligomeric assemblies of ureases.
a) Schematic of biochemical reaction catalyzed by urease. **b** Protein architecture of urease functional unit from full length *C. ensiformis* (Jack bean urease)[160], *K. aerogenes, H. pylori*, and *Y. enterocolitica*. **c** Surface representation of oligomeric assembly of urease in *C. ensiformis* (PDB: 3LA4), *K. aerogenes* (PDB: 1EJW), *H. pylori* (PDB: 1E9Z), respectively. Proteins are color-coded as in **b**. Oligomeric state of urease assembly is indicated at the bottom and the trimeric assemblies are indicated in roman numerals.

ammonia molecule. The latter is released upon breaking of the C–NH_3+ bond, yielding carbamate as the byproduct, the latter spontaneously undergoing hydrolysis. The active site is closed off during the reaction by a conformationally variable helix-turn-helix motif, referred to as the mobile flap. Despite their common catalytic mechanism, ureases display different chain topologies and higher order oligomeric assemblies (Figure 2.3.1 b, c). Jack bean urease assembles into an oligomer of a single type of polypeptide chain with D3 symmetry ($[[\alpha]_3]_2$ D3) (Figure 2.3.1 b, c)[15, 159]. The urease of the human pathogen *Helicobacter pylor*i is composed of two types of polypeptide chains (ureA, ureB) and assembles into a dodecamer (tetramer-of-trimers) with tetrahedral symmetry ($[[\alpha\beta]_3]_4$ T) (Figure 2.3.1 b, c). The ability of ureases to raise the pH of their environment benefits *H. pylor*i in colonizing the stomach and downstream gut, causing gastric ulcers[31]. The urease of the opportunistic pathogen *Klebsiella aerogenes*[163] assembles into a heterotrimer of three proteins ureA, ureB, and ureC, which in turn oligomerizes into a trimer ($[\alpha\beta\gamma]_3$ C3)11 (Figure 2.3.1 b, c). This oligomeric assembly is common also to most other structurally characterized bacterial ureases[15, 159].

Yersinia enterocolitica is the causative agent of yersiniosis, a gastrointestinal infection, reactive arthritis, and erythema nodosum. The infection spreads to

humans through consumption of contaminated food with pigs being one of the largest reservoirs of *Y. enterocolitica*. The symptoms of yersiniosis are fever, abdominal pain, diarrhea, and/or vomiting and is one of the most reported enteritis in some countries, although outbreaks are rare[36]. *Y. enterocolitica* is a facultative intracellular bacterium and can survive in very different environments. In the presence of urea *Y. enterocolitica* can tolerate extremely acidic conditions of pH 1.52. *Y. enterocolitica* urease comprises three polypeptide chains (ureA, ureB, and ureC), an architecture similar to that of *K. aerogene*s urease (Figure 2.3.1 b) and has been biochemically characterized previously[164].

Here we present the structure of *Y. enterocolitica* urease at an overall resolution of 2 Å, which was achieved using recent advances in cryo-EM data collection and processing. The structure shows that *Y. enterocolitica* urease assembles into a dodecameric hollow sphere with a $[[\alpha\beta\gamma]_3]_4$ oligomeric assembly structure of tetrahedral symmetry. A tightly embedded kinked loop is interacting with neighboring domains and is potentially responsible for the assembly of the oligomer. The data allows model building of the active site carbamylated lysine, and visualization of radiation damage to the nickel-metal center as well as of hydration networks throughout the protein.

2.4 Results

2.4.1 Structure determination of *Y. enterocolitica* urease by cryo-EM

We have used single particle cryo-EM to determine the structure of the fully assembled *Y. enterocolitica* urease. We acquired 4,494 movies of urease particles using a Titan Krios transmission electron microscope (TEM) equipped with a K2 direct electron detector and an energy filter (see Methods for details). Approximately half of the movies (2,243) were acquired by illuminating three locations (shots) per grid hole using beam-image shift in order to speed up the data collection [165], whereas the remaining movies were recorded without this feature i.e. just a single shot at the center of the hole. This allowed us to measure and assess the extent of beam tilt and other optical aberrations, as well as the behavior of sample drift between each condition and beam-image shift position. Typical micrographs from the imaged grids are shown in Supplementary Figure 2.10.1 a and a summary of data collection information is given in Table 2.4.1.

Each dataset was processed separately for 3D reconstruction following the strategy depicted in Supplementary Figure 2.10.2. The first obtained 3D map, at an overall resolution of 2.6 Å, revealed that this urease assembly is a dodecamer of tetrahedral (T) symmetry with a diameter of approximately 170 Å. The separate processing of each dataset yielded refined 3D maps at nominal resolutions of 2.10 Å and 2.20 Å for the multi-shot and single-shot cases, respectively (see Methods).

For comparison, we also processed the merged set of particles from both datasets altogether. We observed on the 2D class averages a preferential orientation for the three-fold symmetric view of urease and also the presence of isolated monomers and broken assemblies (Figure 2.10.1). The presence of such incomplete assemblies was further confirmed by performing 3D classification without imposing symmetry, as shown in Supplementary Figure 2.10.1. The 3D class corresponding to the complete dodecameric assembly of urease contained 119,020 particles, of which 69,512 (58.4%) came from the multi-shot and 49,518 (41.6%) from the single-shot dataset. With respect to the number of particles picked from each dataset, 64.7% of the particles from the multi-shot and 56.8% from the single-shot datasets were retained at this stage and throughout the final reconstruction. While coma-free alignment was performed and active beam-tilt compensation in SerialEM [166] was used on our data collections, after performing beam tilt refinement in RELION-3 [150] we observed that the single-shot case has a residual beam-tilt higher than the smallest residual observed in the multi-shot case (Supplementary Table 2.12.2). These two values are however very close to zero and are possibly within the error margin of the *post hoc* beam tilt refinement procedure.

The reduced need to move the specimen stage in beam-image shift mode not only speeds up data collection but also minimizes stage drift. The second and third shots from the multi-shot dataset have comparatively less drift than both the first multi-shot and the single shot, as suggested by the parameter values obtained from the Bayesian polishing training [167] on each beam-tilt class separately (Table 2.12.3). As all the three multi-shots are taken in nearby areas within the same foil hole, this observation is consistent with the annealing of the vitreous ice layer and its carbon support after pre-irradiating the specimen as reported previously [146].

At this point, the nominal resolution of the map after 3D refinement was 2.05 Å. Finally, correcting for residual higher-order aberrations in CTF refinement [153] (Supplementary Figure 2.10.3) yielded a map at a global resolution of 1.98 Å (Figure 2.10.4 a). Local resolution estimation reveals that the core of the map is indeed at this resolution level or better (Figure 2.4.1 2a and Supplementary Figure 2.10.4b), and the local resolution-filtered map was then used for model building as explained

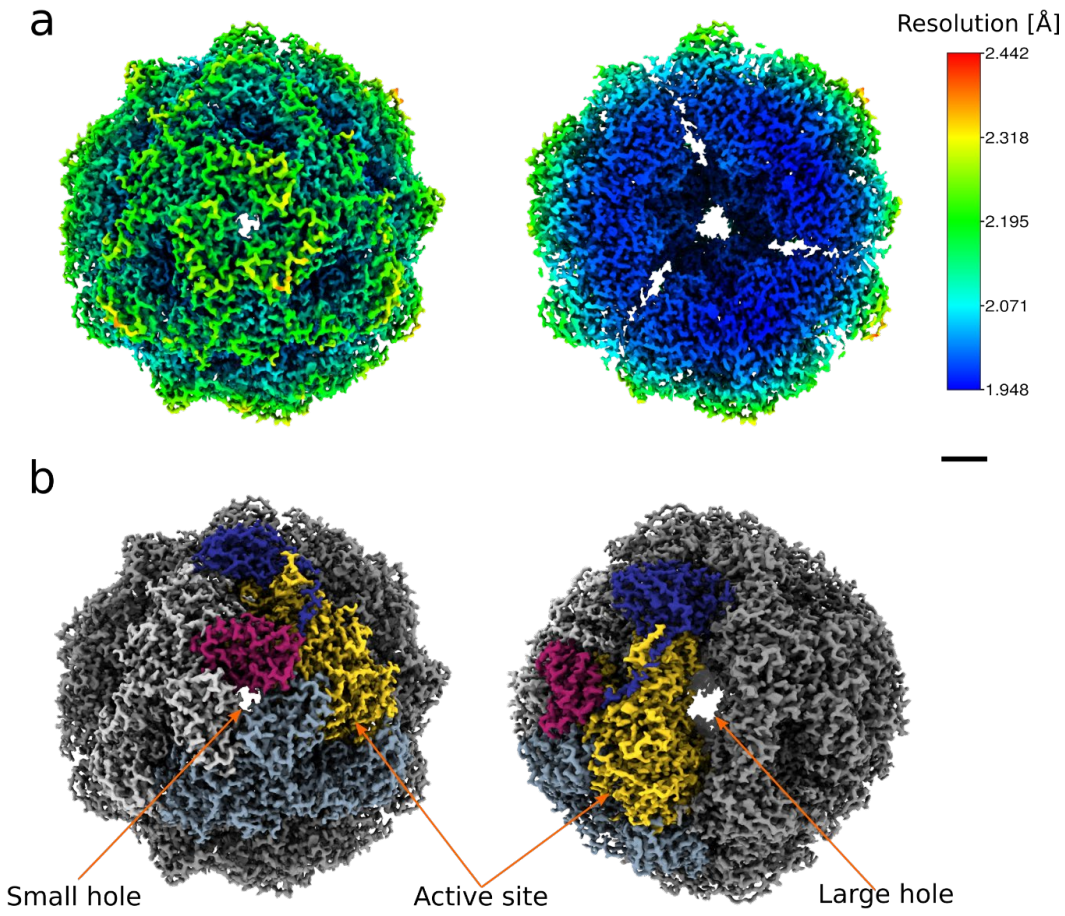

Figure 2.4.1 Cryo-EM analysis of the *Y. enterocolitica* dodecameric urease assembly a) The cryo-EM map (left) filtered and colored by local resolution and a slice cut through the map (right) to show the internal details. b) The assembly architecture highlighted on the map. The three chains that form the basic hetero-trimer are shown in different colors, with the other hetero-trimers shown in shades of gray. Two different views are shown to indicate the location of the small and larger holes at the interfaces, as well as the active site. Scale bars: 20 Å.

in the next section. Despite the twelve-fold symmetry of the urease assembly, a limiting factor in the resolution of the map is the strong presence of preferential orientation, as confirmed by the plot of the final orientation assignments (Supplementary Figure 2.10.4 c). The estimated angular distribution efficiency is 0.78[168]. An overview of the cryo-EM map and its main features are depicted in Supplementary Movie 1.

2.4.2 *Y. enterocolitica* urease assembles as a tetramer of trimers

Model building was initiated from available crystallographic models with subsequent fitting and refinement against the cryo-EM map. The model was built and refined for one asymmetric unit containing one copy of the ureA, ureB and ureC protein each. The model was then expanded using NCS (see Methods). The complete model

Data Collection and processing	
Magnification	78,247 (215,000 nominal)
Voltage (kV)	300
Electron exposure (e-/Å)	42
Defocus range (μm)	−0.2 to −1.5
Pixel size	0.639
Symmetry imposed	T
Initial particle images (no.)	194,603
Final particle images (no)	19,020
Map resolution (Å)	1.98
FSC threshold	0.143
Map resolution range (Å)	1.94–2.40
Refinement	
Initial model used (PDB Code)	4Z42
Model resolution	2.02
FSC threshold	0.5
Model resolution range (Å)	-
Map sharpening B factor (Å^2)	-38
Model composition	
Non-hydrogen atoms	77076
Protein residues	9552
Ligands (Ni^{2+})	24
B factors (Å^2)	
Protein	14.5
Ligand	22.9
R.M.S. deviations	
Bonds (Å)	0.066
Angles (°)	3.665
Validation	
MolProbity score	1.93
Clashscore	4.37
Poor rotamers (%)	2.74

Ramachandran plot	
Favored (%)	94.35
Allowed (%)	5.27
Disallowed (%)	0.38

Table 2.4.1 Data Acquisition, model building and refinement.

Statistics shown for full assembly calculated from the asymmetric unit using NCS.

covering the whole oligomeric assembly contains 9,552 residues, 3,672 waters and 24 nickel ions (two per active site, twelve active sites) (Table 2.4.1). The quality of the model was assessed with the cryo-EM validation tools in the PHENIX package [169]. The map allowed for the building of all residues of ureA (1-100), and residues 31-162 of ureB and 2-327/335-572 of ureC (Figure 2.10.5). The hetero-trimer formed by the three protein chains (ureA, ureB, ureC) (Figure 2.3.1 b) oligomerizes into a homo-trimer. The homo-trimer is arranged in a tetramer-of-trimers making the full complex a dodecamer of the hetero-trimer (Figure 2.4.1 b). There are four large oval shaped holes between the trimers (64 Å long, 12 Å wide, high electrostatic potential) and four smaller holes at the center of the trimer with a diameter of 6 Å (low electrostatic potential), as shown in Figure 2.4.1 b, and the center of the enzyme assembly is hollow (Figure 2.4.1). The holes provide ample opportunity for diffusion of the uncharged substrate and product, and the hollow inside potentially leads to local increase of reaction product. The assembly has the same symmetry as the urease homologue in *H. pylori*, which was postulated to increase stability and/or resistance to acidic environments [39].

For analysis of the protein sequences, the ConSurf Server [170, 171] was used with the sample list of homologs option to get a diverse set of 150 sequences. The protein chains of *Y. enterocolitica* urease are highly conserved across different organisms. The ureA chain is split after a LVTXXXP motif and is 99-100 amino acids long in most cases, with a sequence identity of 55.7%. The ureB chain of *Y. enterocolitica* has between 20 and 30 N-terminal amino acids more compared to the other sequences (except *Kaistia sp.* SCN 65-12), which share an identity of 51.5%. This N-terminal extension is located on the outside of the holoenzyme where ureA and ureB chain split occurs and are too disordered to be modeled in the structure (Figure 2.10.5). The charges and properties of this stretch of amino acids vary and if they still serve a function remains unclear. The last 20 amino acids of the C-terminus of ureB are only represented in half of the compared sequences and accurate sequence conservation could not be determined in this part. This stretch

contains a loop and a C-terminal helix (Supplementary Figure 2.10.5). The ureC protein of the compared sequences has a shared sequence identity of 60.3%. All amino acids involved in catalysis are highly conserved (Figure 2.10.5). The ureA and ureB chains show lower conservation compared to ureC. They are not involved in catalysis but in scaffolding, so the differences could stem from their role in different types of oligomeric assembly (Figure 2.4.2).

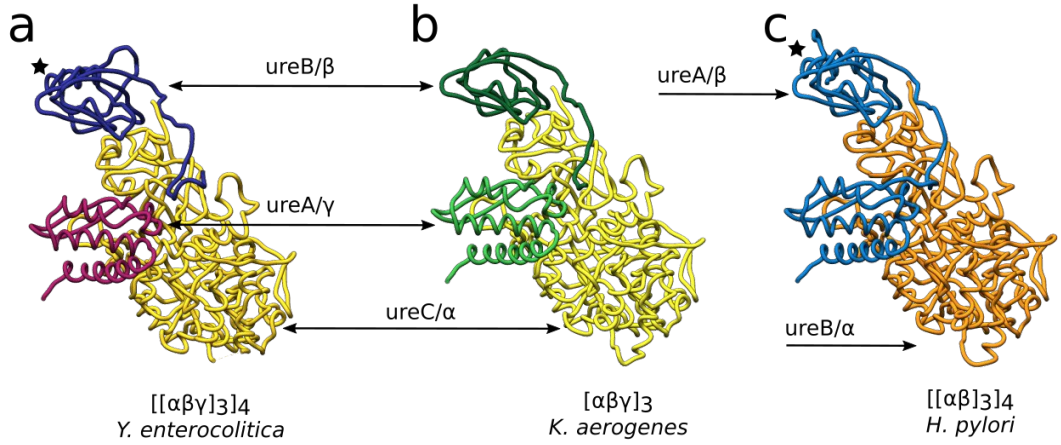

Figure 2.4.2 Comparison of *Y. enterocolitica* urease chain architecture with ureases with different modes of assembly from other pathogens.

Hetero-trimers are shown in tube representation with each chain in the same colors of the sequences in Figure 2.3.1 b. The black star indicates the central helix of the β subunit. The *Y. enterocolitica* and *H. pylori* ureases form the same dodecameric assembly despite having different types of chain splitting, while *K. aerogenes* urease has the same type of chain splitting as in *Y. enterocolitica* but forms only a trimeric assembly.

To investigate this aspect further, we compared the presented structure to the ureases of *H. pylori*, *S. pasteurii* and *K. aerogenes*. Sequence identity scores among these ureases are provided in Supplementary Table 2.12.4. The *H. pylori* urease is made up of two protein chains ureA (that contains the equivalent of ureA and ureB in *Y. enterocolitica*) and ureB (that is the equivalent of ureC in *Y. enterocolitica*) (Figure 2.3.1 c and Figure 2.4.2 c). It assembles into a T-symmetric oligomer like in *Y. enterocolitica* and the crystal structure was solved to 3 Å. *S. pasteurii* and *K. aerogenes* ureases both assemble into a trimer from the hetero-trimeric unit (Figure 2.3.1 c and Figure 2.4.2 b).

S. pasteurii urease has been solved by X-ray crystallography in different conditions (for example, PDB IDs: 2UBP, 3UBP, 4CEU)[45, 172]. Here we use the highest resolution urease structure, that was solved in the presence of the inhibitor N-(n-Butyl)thiophosphoric Triamid (NBPT) to 1.28 Å, for comparison (PDB ID: 5OL4)[158]. The crystal structure of *K. aerogenes* used for comparison in this paper

has a similar resolution range and was solved to 1.9 Å in absence of substrate or inhibitors (PDB ID: 1EJW)[42].

There are two main regions with high root mean square deviations (RMSDs) when comparing these three ureases to the *Y. enterocolitica* model (Supplementary Figure 2.10. 6 and Supplementary Table 2.12.5). The first region with high deviation is the mobile flap, which opens and closes over the active site (residues 312 to 355 of ureC). Both the open and the closed conformations of the mobile flap have been observed in crystal structures, stabilized at pH values lower and higher than the pKa of the conserved His323, respectively[161]. The residues of its connecting loop could not be built with confidence in the cryo-EM model (residues 326 to 333 of ureC). The other region with large differences is on the edges of ureA and ureB where the interactions with the next protomer occur. The *H. pylori* assembly contains an additional C-terminal loop (residues 224-238 of ureA) after the top alpha helix (residues 206-223 of ureA). This helix (central helix) forms the three-fold axis of three neighboring trimers and the loop binds in a head-to-tail fashion to the next trimer forming the tetramer (Figure 2.4.2 a, c) [39]. The core of the assembly is identical in its structure. For whole-chain superposition scores and RMSD values between the compared models please see Supplementary Table 2.12.5.

In the dodecameric assembly seven different interfaces are formed between the hetero-trimers (Figure 2.4.3 a, b and Supplementary Figure 2.10.7). Intra-trimer interactions occur between the three basic hetero-trimers in one assembled trimer, forming a three-fold symmetry axis (Figure 2.4.3 a). The interactions between these trimers to form the tetramer then make up a different three-fold symmetry axis (Figure 2.4.3). The three largest interfaces (interfaces 1-3) are formed intra-trimeric between ureC of one hetero-trimer and ureC, ureA and ureB of the next trimer (Figure 2.4.3 a). The three ureA proteins make up the intra-trimer-core (first three-fold axis) with interface 4 (Figure 2.4.3 c and Supplementary Figure 2.10.7). Comparison of the interface areas formed in the trimer assembly shows no

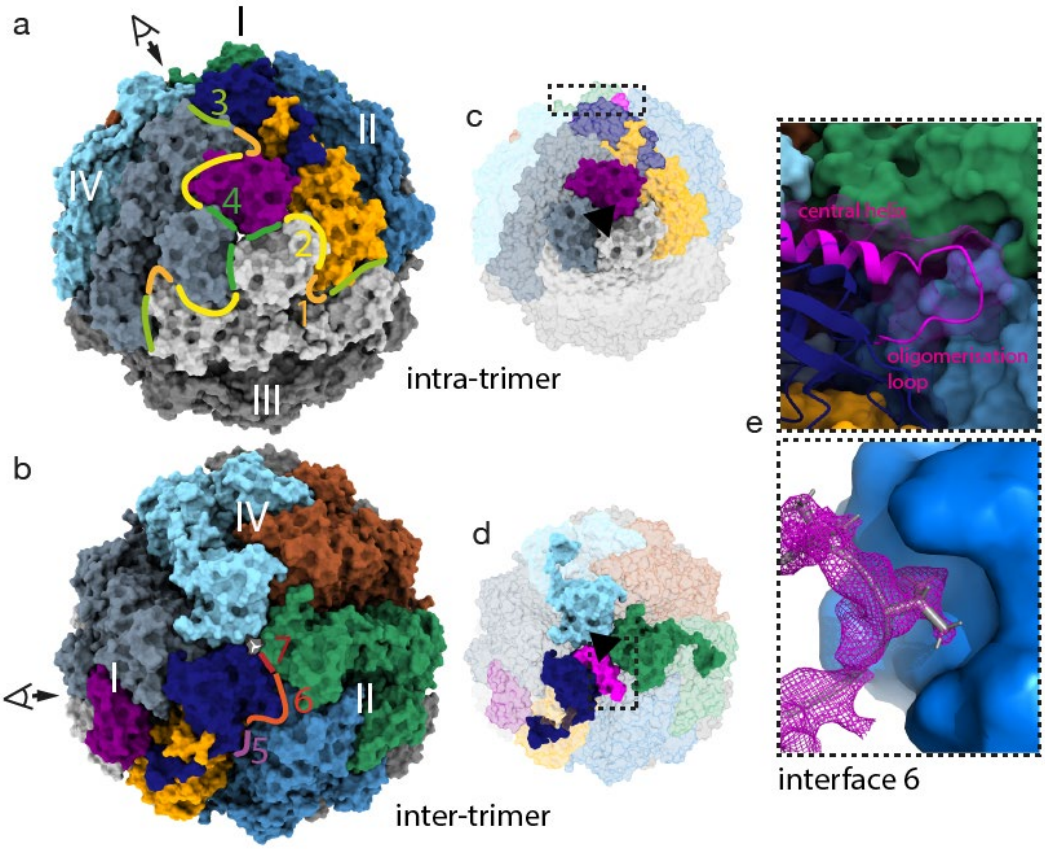

Figure 2.4.3 Interfaces in dodecameric assembly of *Y. enterocolitica*.

a) surface model in front view of trimer of *Y. enterocolitica* urease with intra-trimeric interfaces 1-4 indicated with color-coded lines and numbers. b) Same model shown from the top (view indicated with eye) and the inter-trimeric interfaces. c) Front view with intra-trimeric-core highlighted and three-fold axis indicated with black triangle. Inset for e) in dashed box. d) same as c) but from the top view. e) Interface 6 with loop from ureB (magenta) binding into pocket of ureC of neighboring trimer. Upper inset shows ureB in cartoon and transparent surface and ureC in surface representation. Lower panel shows ureC as surface and ureB loop as cartoon with density.

substantial differences between the four organisms (Supplementary Figure 2.10.7). Inter-trimer interfaces (interfaces 5, 6, 7) formed in the dodecameric *Y. enterocolitica* and *H. pylori* ureases have similar areas (Figure 2.4.3 b and Supplementary Figure 2.10.7). Part of the interactions occur between ureB and ureC forming interfaces with each other (interface 4, 6). The other interaction is between the three ureB proteins and forms interface 7 and the inter-trimer-core (second three-fold axis) with their central helices (Figure 2.4.3 b). *Y. enterocolitica* does not have the same oligomerization loop after the central helix proposed for *H. pylori*. However, there is a short loop before the central helix, which is extended in *Y. enterocolitica*. It binds into a pocket of ureC of the neighboring trimer in interface 6 (Figure 2.4.3 e). These types of loops or extensions are missing from *S. pasteurii* and *K. aerogenes* ureB proteins. *S. pasteurii* has the central helix, but there is no extended loop before or after it (Supplementary Figure 2.10.6 b). *K. aerogenes* urease does not have a helix nor a loop in this region (Figure 2.4.2 b). This suggests that the presence of oligomerization loops in ureB is crucial for determining the oligomeric state of the enzyme.

The dodecameric holoenzyme structure of ureases might aid in stabilizing the protein at acidic pH, and in combination with 12 active sites producing ammonia enables the formation of a pH-neutralizing microenvironment around the assembly [39]. This ensures the continued function of the enzyme and makes this type of oligomeric assembly essential to survival of *Y. enterocolitica* in the host. It is remarkable that *Y. enterocolitica* is the first organism outside the *Helicobacteraceae* family to have a known dodecameric urease. Considering the different subunit organization between these ureases, it raises the question of what particular events in the evolutionary history of *Y. enterocolitica* could have led to this type of assembly[41].

2.4.3 The empty active site is filled with water

At the global resolution of 1.98 Å, detailed structural features can be observed. All throughout the highly resolved areas of the protein, salt bridges, backbone and side chain hydration and alternative side chain conformations can be visualized (Supplementary Figure 2.10.8 a-c). Furthermore, the high resolution allows for a detailed description of the nickel-metallo-center and the active site. The active site is located on the ureC protein at the edge of the hetero-trimer and is wedged in between the ureA and ureB proteins of the next hetero-trimer in the homo-trimeric assembly (Figure 2.4.4 a).

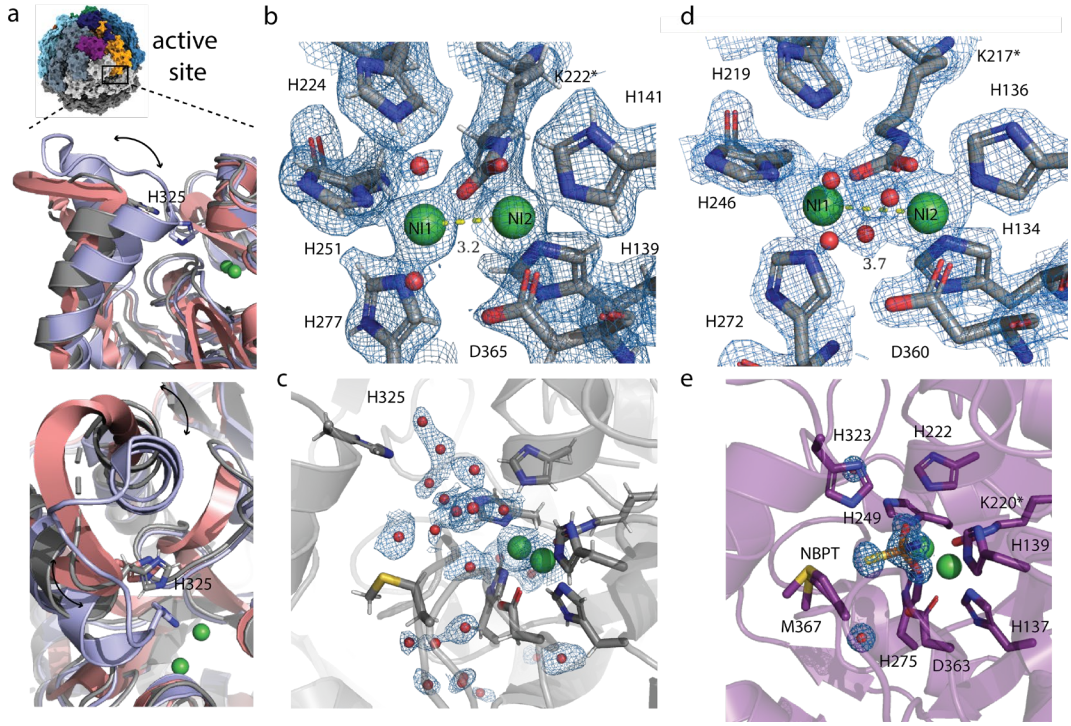

Figure 2.4.4 Active site of *Y. enterocolitica* urease.

a) Overview of urease assembly with the active site location indicated. Inset shows in top panel side view of urease crystal structures from *S. pasteurii* mobile flap shown in open conformation in salmon (PDB: 2UBP) and in closed position as light purple (PDB: 3UBP). In gray the cryo-EM structure of *Y. enterocolitica* is overlaid and the green spheres represent the Ni^{2+} ions of the active site. Bottom panel shows top view of the three structures. Arrows indicate movement of helix and catalytic HIS325 is shown as stick. b) Model of active site residues and Ni^{2+} ions with the cryo-EM map of *Y. enterocolitica* at 1.98 Å nominal resolution. Yellow line indicates distance between Ni^{2+} ions in Å; c) shows the water molecules in the active site. d) Crystal structure of *K. aerogenes* urease at 1.9 Å resolution (PDB: 1EJW). Yellow line indicates distance between Ni^{2+} ions in Å. e) Crystal structure of *S. pasteurii* at 1.28 Å with inhibitor NBPT (PDB: 5OL4).

The catalysis of ammonia and carbamate from urea occurs in two steps (Figure 2.3.1 a). Urea first interacts with the nickel ions through its carbonyl oxygen and amino nitrogens. The active site contains two Ni^{2+} ions which are coordinated by six different amino acids (Figure 2.4.4 b). Both Ni^{2+} ions are coordinated by the carbamylated LYS222*. Ni(1) is additionally coordinated by HIS224, HIS251 and HIS277 and Ni(2) by HIS139, HIS141 and ASP365. Close to the active site is a methionine (MET369), which can be modelled in different alternative conformations. One conformation could potentially reach the active site. There is no described function for this amino acid (Figure 2.4.4 c and Supplementary Figure 2.10.9).

The active site is protected by a helix-turn-helix motif, called the mobile-flap. Its function is to coordinate the access of substrate to the catalytic site and the release of the product from it [15, 159]. The protonation state of a conserved histidine on the mobile flap (His325) is essential for catalysis by determining opening and closing of the mobile flap and strongly depends on the solution pH[159, 161](Figure 2.4.4 a). By closing of the mobile flap His325 moves closer to the active site, stabilizing the distal amine of urea in the active site pocket[15, 159, 161, 162]. After closing of the mobile flap, the urea molecule chelates the two Ni ions in the active site, and following the nucleophilic attack by the bridging hydroxide onto the urea C atom, a proton is transferred to the distal amine group from the metal-bridging C–OH group, yielding an ammonia molecule after breakage of the resulting C-NH_3+ bond. Flap opening then releases ammonia and carbamate, where the latter spontaneously hydrolyzes into another molecule of ammonia and bicarbonate. The mobile flap of the cryo-EM structure presented here is modeled in an open position, however the local resolution is lower than in the surrounding areas, indicating flexibility. The sample was frozen in a buffer of pH 7.0, where the mobile flap of urease can adopt both open and closed conformations[161]. The twelve active sites of each particle adopting different conformations are averaged by single particle reconstruction with symmetry imposition into a mainly open conformation. In the absence of substrate or inhibitors in the sample, the mobile flap cannot be stabilized in one conformation (Figure 2.4.4 a). Coordinated water molecules can be seen in the empty pocket of the active site, which do not only form hydrogen bonds with side chains or the protein backbone, but also with each other constituting a hydration network (Supplementary Figure 2.10.8 d).

The resolution in the active site is sufficient for complete atomic description of the coordinated Ni^{2+} ions, including the carbamylated lysine. The protonation states of the active site residues are also represented in the map (Figure 2.4.4 b). One of the hydroxide molecules in the active site is essential as it performs the nucleophilic attack on urea while other molecules are displaced by urea and the closing of the mobile flap[15, 159, 161, 162].

Comparison to the crystal structure of *K. aerogenes* of similar nominal resolution (1.9 Å) shows differences in the visualization of these features. This crystal structure was solved in absence of inhibitors or substrate such that the active site is also empty and the mobile flap in an open conformation (Figure 2.4.4 a). The details of the map provide finer details around the Ni^{2+} ions in the cryo-EM map than in the crystallographic data. The protonation of the histidines is clearly visible in the cryo-EM density (Figure 2.4.4 b). The positions of the side chains and the Ni^{2+} ions in

the active site are very similar to the *Y. enterocolitica* urease structure with a RMSD of 0.270 Å (Table 2.12.5). The highest resolution *S. pasteurii* crystal structure was solved in presence of the inhibitor NBPT, which displaces the essential water molecules needed for the reaction from the active site. The closing of the mobile flap displaces the rest of the waters and brings the catalytic HIS323 closer to the active site. The tight packing of side chains prevents urea from entering the active site, efficiently blocking it (Figure 2.4.4 b, e). The active site residues and Ni^{2+} ions have a RMSD of 0.293 Å between *S. pasteurii* and *Y. enterocolitica*.

2.4.4 Nickel atoms come closer together

The distance between the Ni^{2+} ions is 3.7 Å in X-ray structures of *K. aerogenes* and *S. pasteurii*, but only 3.2 Å in the *Y. enterocolitica* cryo-EM model (Figure 2.4.4 b, d). Short distances of 3.1-3.3 Å were described for *S. pasteurii* and *K. aerogenes* at high resolutions for structures in presence of β-Mercaptoethanol (β-ME) [173]. Knowing that metallic cores are particularly sensitive to radiation [174], we tried to determine the extent to which radiation damage can explain the shorter distance between the Ni^{2+} ions. For this purpose, we generated per-frame reconstructions for the first 25 frames of our data collection and refined the model on each of them (see Methods) and measured the distances between the residues involved in ion coordination, shown in Figure 2.4.5. Bayesian particle polishing[167] was run again before calculating each per-frame reconstruction. At the beginning of the exposure, in which the frames contribute more to the full reconstruction due to dose-weighting [167, 175], there is a trend of the ions coming closer together (Figure 2.4.5 a-i). While we cannot determine exactly how this arises from radiation damage, it is likely a result of several interactions in the active site changing simultaneously along the exposure. For example, both Ni(1) and Ni(2) tend to come closer to the carbamylated LYS222 (Figure 2.4.5 a-iv, v) as ASP365 vanishes (Figure 2.4.5 a-vi), which can be seen in the Supplementary Movie 2. Aspartic acid is known to have its side chain damaged very early on [176]. The dynamic interplay between residues along the exposure (Figure 2.4.5 b) is likely due to the different rates at which specific types of bonds and residues are damaged [177]: first negatively charged residues, then positively charged ones followed by aromatic side chains, as also observed in Supplementary Movie 2. A possible explanation of how these events may account for the shorter distance between the Ni2+ ions is then that their bridging hydroxide molecule becomes deprotonated into the oxide form, either by radiation damage directly or by local pH changes arising from it.

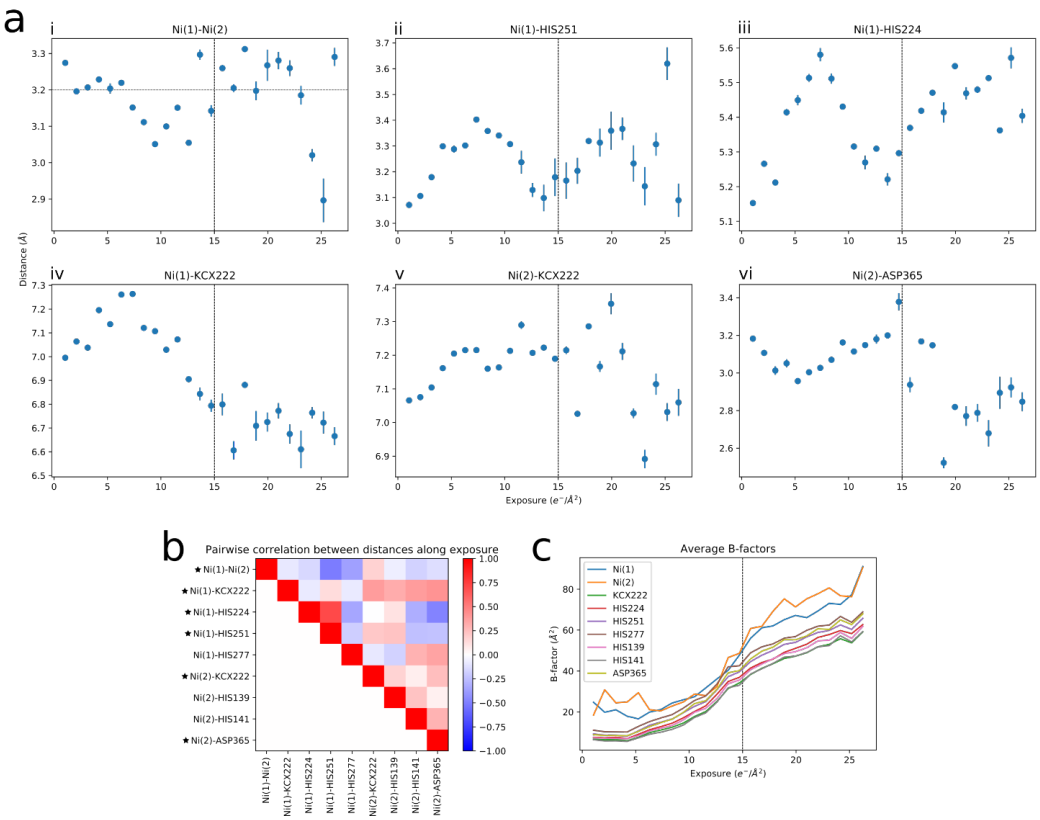

Figure 2.4.5 Radiation damage affects the distance between residues in the active site

a) Distances between the Ni^{2+} ions and selected residues involved in their coordination are plotted against the accumulated exposure. For each reconstruction calculated along the exposure, the model was refined, and distances measured. Dots indicate the average and error bars show +/- one standard deviation across five refinement runs with different random seeds. Horizontal dashed line in panel a-i) shows the distance in the model obtained from the full reconstruction with all frames. Vertical dashed lines show approximately the exposure at which the density for charged residues completely vanishes (see Supp. Mov. 2). b) Correlation coefficients between distance changes along the exposure for selected residues involved in ion coordination. Distance plots shown in a) are indicated with a star. c) Average B-factors of selected residues plotted against the accumulated exposure.

The oxide form is known to have a more favorable ferromagnetic interaction with the two nickel ions[45], although not found in ureases under native conditions[178]. Furthermore, the B-factors suggests that some residues in the active site, and in particular the Ni2+ ions, are indeed damaged more strongly than the rest of the protein, right from the beginning of the irradiation as shown in Figure 2.4.5 c. We note however that, in the present analysis, radiation damage cannot be completely disentangled from other effects such as residual sample movement, which is

especially difficult to correct in the initial frames of the exposure. The later part of the exposure must also be interpreted with caution, as atomic coordinates become less reliable, which is verified by the overall increase in B-factors in Figure 2.4.5 c and the error bars in Figure 2.4.5 a.

2.5 Discussion

Large urease assemblies have been historically difficult to study by X-ray crystallography [39]. We have determined the structure of a dodecameric urease assembly, a metalloenzyme from the pathogen *Y. enterocolitica* at an overall resolution of 1.98 Å using cryo-EM. The collection of datasets with and without beam-image shift demonstrates the advantages of using this feature of modern TEMs and invites further investigations on the behavior of optical aberrations and specimen drift.

Our results demonstrate the feasibility of cryo-EM as a technique for obtaining structures of clinically relevant enzymes with sufficient quality for *de novo* model building and drug design. The cryo-EM map has allowed a detailed description of the active site and the oligomeric assembly. More specifically, we could observe the putative oligomerization loop that enables the dodecameric assembly, which was hypothesized to be responsible for the enhanced survival of *Y. enterocolitica* in highly acidic environments [35]. This urease is the first outside the Helicobacteraceae family, and therefore without an αβ subunit organization, to have a dodecameric assembly reported. What evolutionary events have led to this intriguing combination of subunit organization and quaternary structure are unknown.

Furthermore, in comparison to the *K. aerogenes* structure, which is at approximately the same nominal resolution, the cryo-EM map offers an improved representation of protons and Ni^{2+} ions. A possible explanation is that the error in the phases derived in the X-ray structure determination grows faster towards the limit of observed diffraction. Another aspect to be considered is that X-rays and electrons probe different properties of matter, respectively the electron density and the integrated Coulomb potential. Our results prompt a more detailed investigation of these effects and how they affect the representation of features at high resolution.

Finally, we noticed that radiation damage can partially explain the shorter distance observed between the nickel atoms in the active site. Given that ions and charged residues are damaged very early on in the exposure [174, 176], this effect cannot

be neglected in structures derived from cryo-EM reconstructions. Novel direct electron detectors with higher frame rates may allow time-resolved experiments to investigate these effects in more detail.

2.6 Data availability

The model has been deposited at the PDB under accession code 6YL3. The map has been deposited at the EMDB under accession code EMD-10835. Raw electron microscopy data is deposited in EMPIAR, accession code EMPIAR-10389. All other data supporting the findings of this study are available from the authors upon request.

2.7 Acknowledgments

The authors would like to thank L. Kovacik and K. Goldie for assistance in data collection. R.D.R. acknowledges S. Klumpe and A. Nunes-Alves for discussions on tracking atomic coordinates. Cryo-EM data processing calculations were performed at sciCORE (http://scicore.unibas.ch/) scientific computing center at the University of Basel. R.D.R. and L.A. acknowledge funding from the Fellowships for Excellence program sponsored by the Werner-Siemens Foundation and the University of Basel. This work was in part supported by the Swiss National Science Foundation (grants 177195 and 185544, NCCR TransCure).

2.8 Author Contributions

R.D.R. and R.A. performed the cryo-EM experiments and data analysis. L.A. built and analyzed the atomic model. R.P.J. performed X-ray crystallography experiments. M.A.M. expressed and purified the protein. P.R. prepared and screened EM samples. J.Z. performed the higher-order aberration corrections and analysis. T.S., T.M. and H.S. initiated and supervised the project. R.D.R, L.A., R.A., T.M. and H.S. wrote the manuscript with assistance from all authors.

2.9 Methods

2.9.1 Protein expression and purification

The *Y. enterocolitica* urease was purified for cryo-electron microscopy according to the protocol of [179]. The strain was precultured overnight at 37°C for 18 hours in a

medium containing 37 g/l of brain/heart infusion (Oxoid, CM0225), 50 µg/ml streptomycin sulfate (Applichem, A1852.0100), 35 µg/ml nalidixic acid (Applichem, A1894.0025), 50 µg/ml meso-diaminopimelic acid (Sigma, D1377) and 100 µM nickel(II) chloride hexahydrate (Sigma, N6136). 6 x 600ml of expression cultures were inoculated at OD of 1 at 28°C for 23 hours. Cells were harvested by centrifugation and the cell pellet resuspended in 0.15M NaCl, 50mM Tris pH 8.0. The cell lysate was applied directly to a Sephacryl S-300 HR 26/60 column equilibrated with 150 mM NaCl, 50mM TrisHCl pH 8.0. The active fractions as identified by a phenol-hypochlorite assay [180] were buffer-exchanged to 50mM Tris pH 7.0 within a centrifugal filter unit (Sartorius, Vivaspin MWCO 50kDa) and applied to a Mono Q HR 5/5 column pre-equilibrated with 50mM TrisHCl, pH 7.0. The protein was eluted in 50mM Tris pH 7.0 by a gradient to 1M NaCl, concentrated on a centrifugal filter unit (Sartorius, Vivaspin MWCO 50 kDa) and purified by SEC as before. The purity of the urease sample of the two preparations was verified on a 4%/12% SDS-PAGE and by mass spectroscopy.

2.9.2 Sample preparation

Approximately 3 µl of the 0.39 mg/ml urease solution were applied to glow-discharged Quantifoil holey carbon grids. After 3-second blotting, the grids were flash-frozen in liquid ethane, using a FEI Vitrobot IV (Thermo Fisher Scientific) with the environmental chamber set at 90% humidity and 20 °C temperature.

2.9.3 Data acquisition

Cryo-EM data were collected on a FEI Titan Krios (Thermo Fisher Scientific) transmission electron microscope, operated at 300 kV and equipped with a Quantum-LS imaging energy filter (GIF, 20 eV zero loss energy window; Gatan Inc.) and a K2 Summit direct electron detector (Gatan Inc.) operated in dose fractionation mode. Data acquisition was controlled by the SerialEM [166] software, performed in counting mode, with a 42 $e^-/Å^2$ total exposure fractioned into 40 frames over 8 seconds. The physical pixel size was 0.639 Å at the sample level. The data was pre-processed via the FOCUS package [181], including drift-correction and dose-weighting using MotionCor2 [149] (grouping every 5 frames and using 3x3 tiles) and CTF estimation using CTFFIND4 [182] (using information between 30 Å and 5 Å from the movie stacks). With these settings, we collected two datasets: one using beam-image shift [165], with three shots per grid hole, comprising 2,243 movies, and a second one taking a single shot per hole, with 2,252 movies. A summary of

data collection information is given in the Table 2.12.1.

2.9.4 Image processing

The two datasets were initially processed separately as shown in the flowchart of **Supplementary Figure 2.10.2**. We excluded all movies whose resolution of CTF fitting was worse than 4 Å according to CTFFIND4, leaving 2,197 movies in the multi-shot dataset or 2,115 in the single-shot dataset for further processing. Using the template-free LoG-picker algorithm [150] we picked an initial set of 157,699 particle coordinates on the multi-shot dataset. These particles were extracted and subjected to one round of 2D classification with the aim of removing “bad” or false-positive particles. Best results in 2D classification were observed when enabling the RELION option “Ignore CTFs until first peak?”. Selecting only the classes displaying views of urease with high resolution features, a new subset containing 60,271 particles was obtained. Using this subset, a first 3D map was obtained by the *ab initio* stochastic gradient descent (SGD) algorithm [150, 152] with and without tetrahedral symmetry imposed. The symmetric map was consistent with previously determined structures of ureases [39, 183]. The particles were then subjected to 3D refinement using the map from the *ab initio* procedure as starting reference, resulting in a map at 2.6 Å resolution. We then generated new templates for particle picking by low-pass filtering the unsharpened map from this first 3D refinement to 20 Å and calculating evenly oriented 2D projections from it. These templates were then used for picking with Gautomatch [184], detecting 107,399 particle coordinates on the multi-shot dataset or 87,204 particle coordinates on the single-shot dataset. Visual inspection of randomly selected micrographs indicated this set of coordinates was better than that previously found by the LoG-picker, in the sense that it contained fewer false positives and more true particles. The newly extracted particles were then subjected to 2D and 3D classification procedures to get rid of false positive, damaged or broken particles, which yielded cleaner subsets with 62,884 (multi-shot) or 51,173 (single-shot) particles. Using the current best map from the multi-shot dataset as a starting reference, we then performed masked 3D refinements on the two datasets separately, interleaved with rounds of CTF refinement and Bayesian particle polishing [167]. More specifically, we refined defocus *per particle*, astigmatism *per micrograph* and beam tilt *globally* in CTF refinement. In the multi-shot dataset, each of the three relative “shooting targets” were assigned a different class for separate beam tilt refinement. The parameters for Bayesian particle polishing were trained separately on ~5,000 particles from

each dataset at this stage. Each dataset yielded refined maps at 2.10 Å (multi-shot) and 2.20 Å (single-shot) resolution. Best results, however, were obtained when merging the particles picked by template-matching on each dataset (194,603 particles in total) and processing them altogether. After 2D classification, 141,069 particles remained (Figure 2.10.1 b), and after 3D classification, there were 119,020 particles (Figure 2.10.1 c). CTF refinement was then performed using four beam tilt classes, with the particles from the single-shot dataset belonging to a new, fourth class (Table 2.12.2). Defocus and astigmatism were both refined *per particle* this time, resulting in a map resolution of 2.05 Å. We compared polishing the full merged dataset at once and each beam tilt class separately, to verify if there were differences in the patterns of particle motion. For training the polishing parameters ~10,000 particles were used in each case this time. Although we did observe different statistics of particle motion (Table 2.12.3), resolution and overall quality of the map did not improve further by performing either merged or separate polishing of the different shots. Finally, correction of third-order aberrations in RELION-3.1 [153] (Figure 2.10.3) followed by local 3D refinement resulted in a global map resolution of 1.98 Å.

All resolution estimates reported were obtained by considering the 0.143 threshold [185] on the Fourier shell correlation (FSC) curve [186] between independently refined half-maps [187]. A solvent-excluding mask was generated by low-pass filtering the maps to 12 Å, binarizing the filtered map and adding a soft edge consisting of a cosine-shaped falloff to zero. The FSC curve was corrected for artificial correlations introduced by the mask [188]. Local resolution was estimated using the approach implemented in RELION [189].

2.9.5 Model building, refinement and analysis

After refinement of the map to high resolution it had to be flipped in UCSF Chimera [190] to match the correct handedness. The biological assembly from the crystal structure of *Y. enterocolitica* urease (Note 2.13) was rigid-body fitted to the map in Chimera. The non-crystallographic symmetry (NCS) was calculated with PHENIX v1.17 [191] from the crystal structure. Only using the hetero-trimer of the three proteins, backbone and side chains were built, corrected or confirmed in Coot [192]. After several rounds of manual refinement of the model in Coot, applying NCS and real-space refinement in PHENIX, the model comprised side chains of residues 1-100 of ureA, 31-162 of ureB and 2-327, 335-572 of ureC. Residues 328-334 are disordered and could not be modeled with confidence. NCS constraints were not

used during final refinements as to include alternative side chain conformations. Waters were built manually and refined in PHENIX and were added to the closest chain by the program phenix.sort_hetatoms. The quality of the refinement was assessed by cryo-EM Validation tool (Table 2.10.1) [169].

2.9.6 Structure analysis

The electrostatic potential was calculated using the APBS plugin in PyMOL [193]. The ConSurf server [170, 171] was used to find 150 sequences per urease protein for alignment with ClustalW [194] and calculate conservation per residue. The "sample the list of homologs" option was used to get a diverse representation across all species. The sequence identity of the 150 sequences was determined using BLSM62 in Geneious [195]. The PDBePISA server (https://www.ebi.ac.uk/pdbe/pisa/) was used to find and calculate interface areas [196].

2.9.7 Radiation damage analysis

Bayesian particle polishing in RELION [150, 167] was carried out on a sliding-window basis along the exposure, including 5 frames at a time, starting from frame 1 up to frame 25. Half-set reconstructions were then created from each polished particle stack and post-processed using the same mask as that applied to the reconstruction from all frames. On each post-processed reconstruction, real space refinement of chain C (α-subunit containing the active site) from the full reconstruction was carried out in PHENIX [197] for 5 macro-cycles. This procedure was repeated 5 times with different random seeds. Distance between residues in the resulting refined models were calculated using BioPython [198] and plotted using the NumPy (https://www.numpy.org) and Matplotlib (https://www.matplotlib.org) Python modules.

2.9.8 Structural representations and figure generation

Protein structural representations were generated using the following software: UCSF Chimera [190], ChimeraX [199], and PyMol (Schrödinger, LLC), with the aid of the Inkscape, Adobe Illustrator and Adobe Photoshop (Adobe Inc.) programs for creating figures.

2.10 Supplementary Data

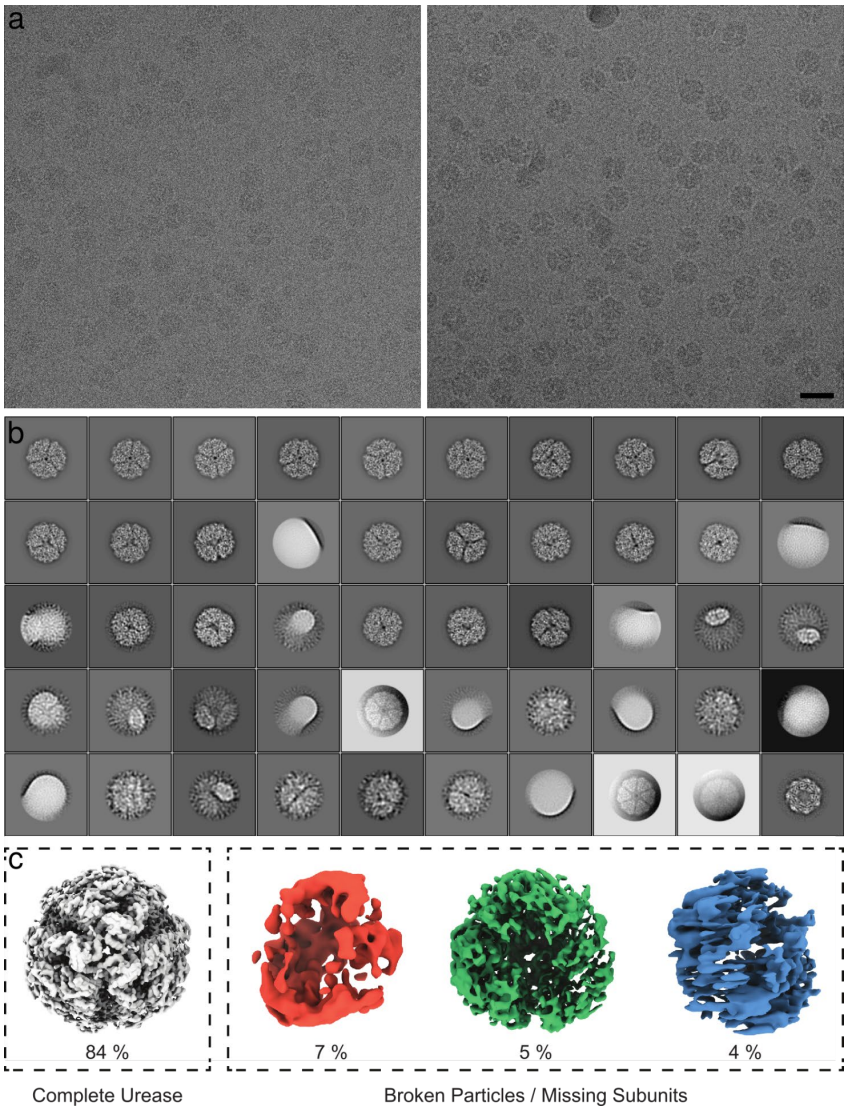

Figure 2.10.1 Cryo-EM data of *Y. enterocolitica* urease.

a) Two representative micrographs from the dataset, acquired at -0.64 and -1.36 µm defocus, respectively. Scale bar: 200 Å. b) 2D class averages obtained from 194'603 particles in the merged dataset. These averages were obtained with RELION's "Ignore CTF until first peak" option enabled and are sorted by decreasing order of number of particles in each class. Views of urease with missing subunits are observed. The bottom right average shows a contamination by GroEL. c) 3D class averages obtained from 141'069 particles in the merged dataset. These averages were obtained without symmetry imposition in RELION. The 3D classes are sorted by the indicated fraction of particles assigned to it. The first class is a complete dodecameric urease assembly, while the other classes represent urease structures with at least one trimer missing from the tetrahedron.

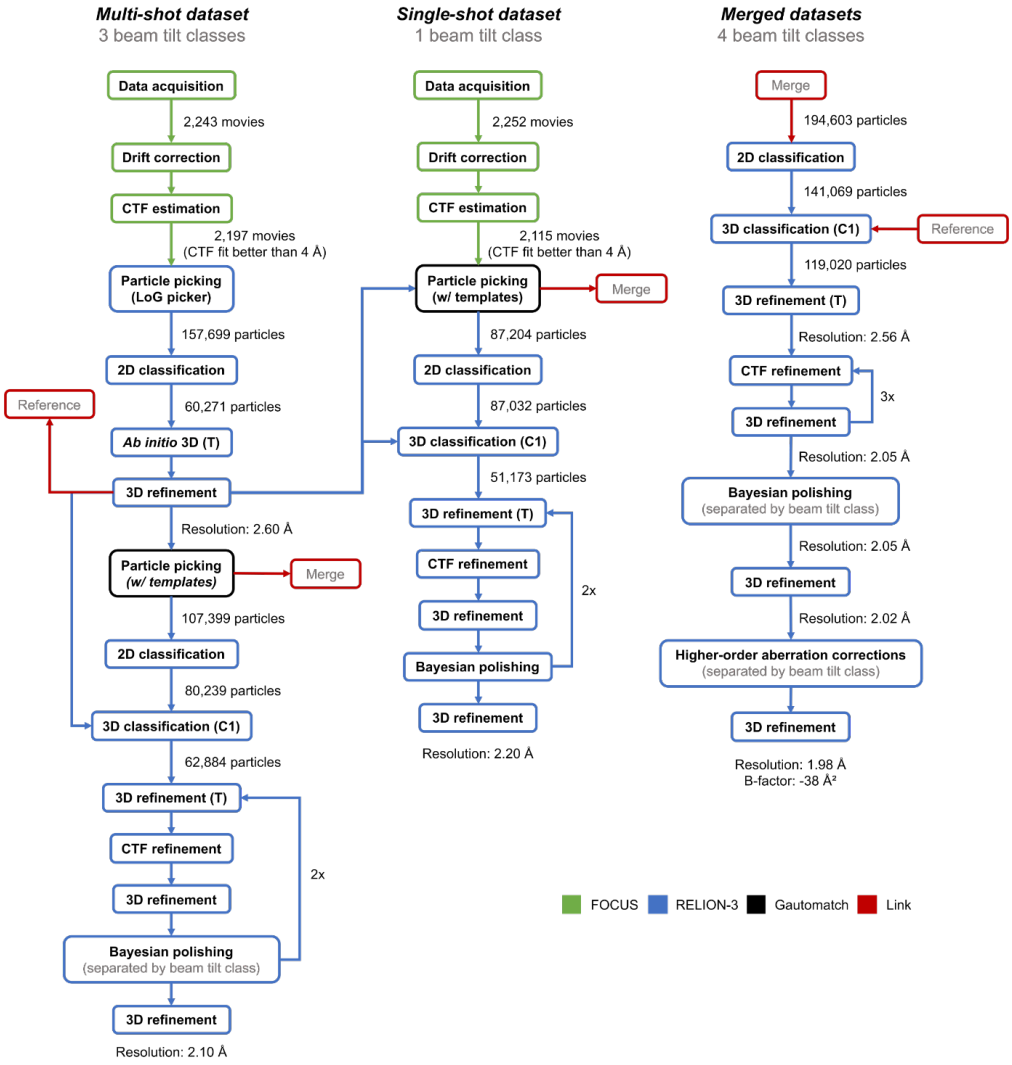

Figure 2.10.2 Data processing flowchart for the urease cryo-EM map.

Masking and postprocessing jobs have been omitted for clarity. All CTF refinement jobs included beam tilt and per-particle defocus refinement (see Methods for details). All resolution estimates given correspond to the corrected FSC curves between masked half-maps obtained from postprocessing jobs in RELION.

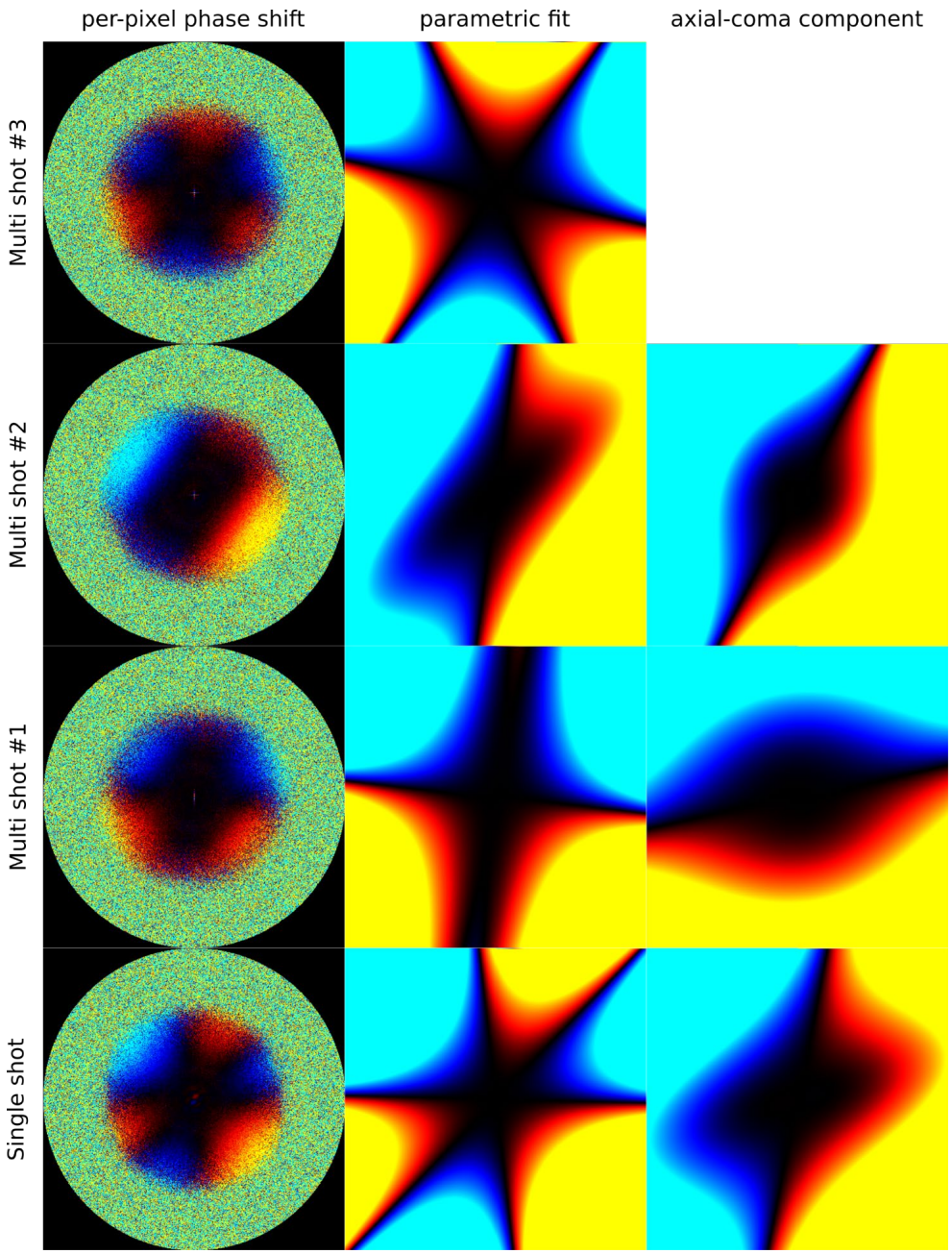

Figure 2.10.3 Fits of the anti-symmetrical aberrations arising at the four different beam-shift positions.

The left column shows phase shifts measured independently for each Fourier pixel, while the center column shows their parametric fits using third-order Zernike polynomials. The first position (top row, the third beam-shifted multi-shot) corresponds to an essentially untilted beam (Supp. Tab. 2), while the other two multishots and the single-shot dataset exhibit tilts to different extents. Note that even the untilted dataset shows a significant trefoil aberration. In the right column, the parametric fit of the first position has been subtracted, yielding residuals roughly consistent with the axial coma produced by a tilted beam.

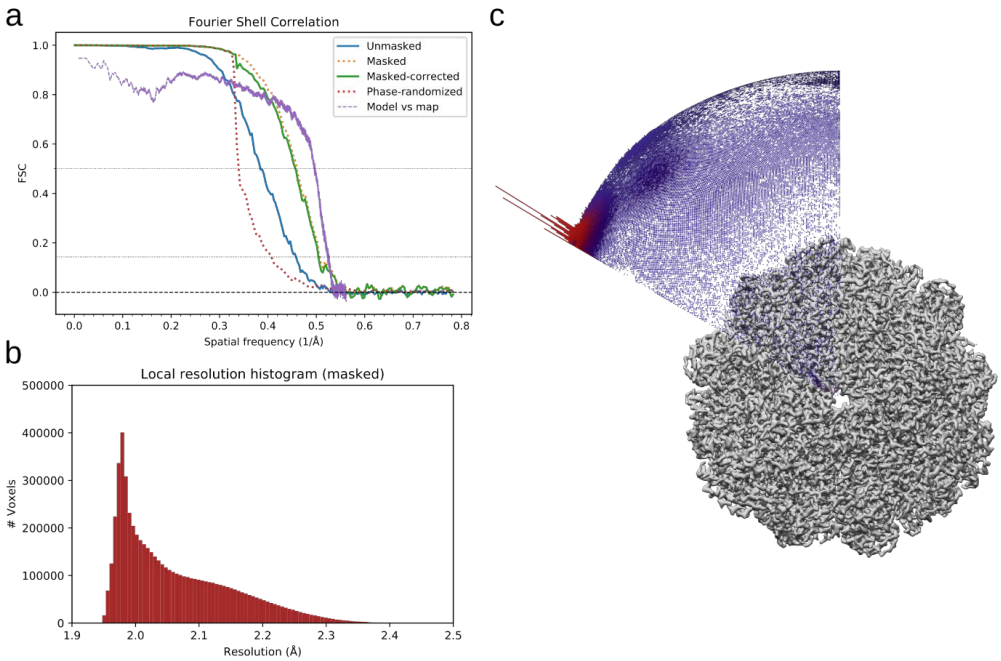

Figure 2.10.4 Resolution estimates of the urease cryo-EM map.

a) FSC curves between the half-maps when unmasked (solid blue line), masked (dashed orange), corrected by high-resolution noise substitution after masking (solid green), phase-randomized (dashed red) and between the atomic model and the full experimental map (dashed violet). b) Histogram of local resolution assigned to each voxel. c) Angular distribution of particles in the urease cryo-EM reconstruction overlaid on the unsharpened map.

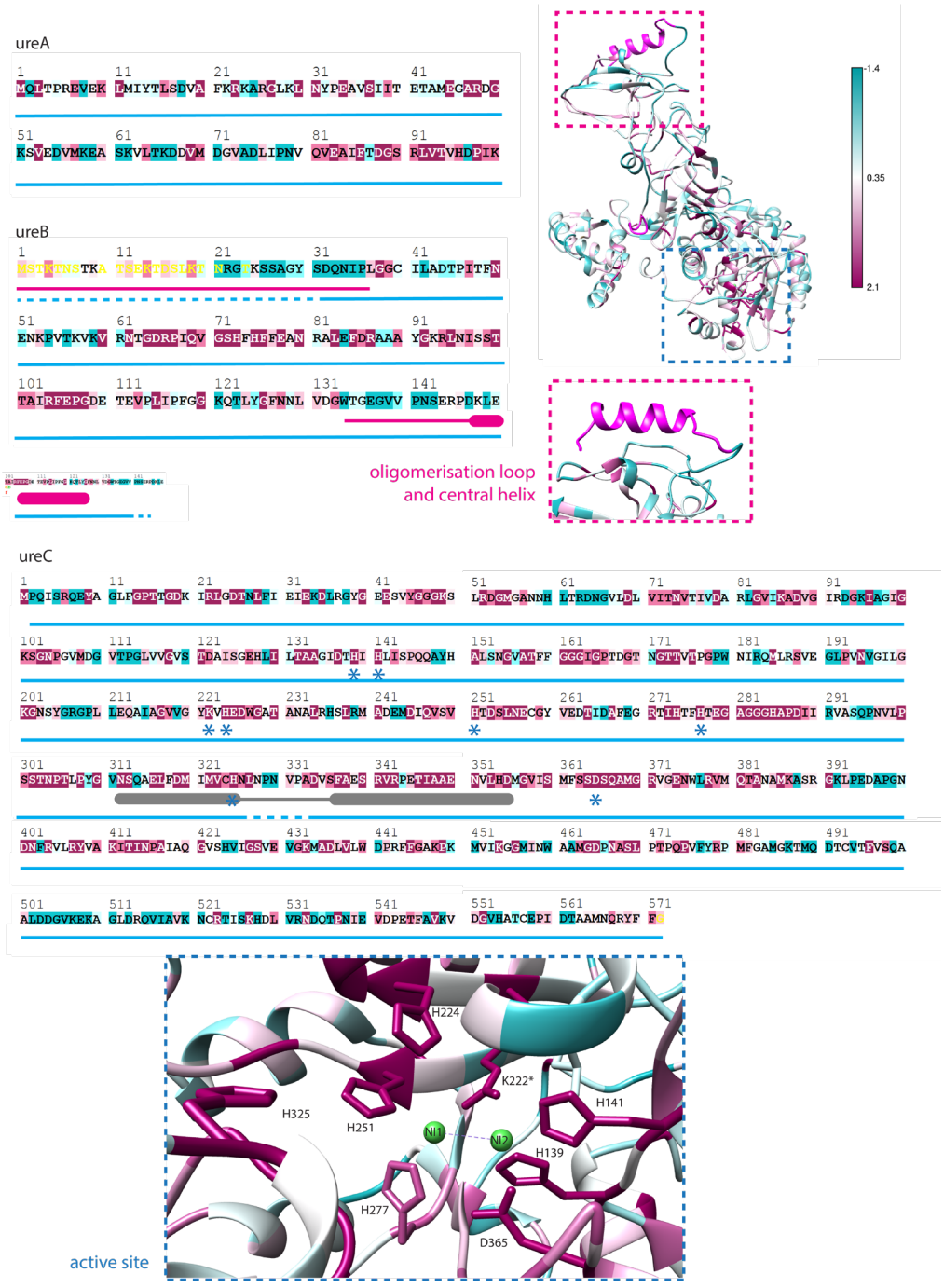

Figure 2.10.5 Alignment of 150 ureA, ureB and ureC sequences

Alignment of 150 ureA, ureB and ureC sequences, chosen by sampling from all homologs found for each protein. ClustalW was used for sequence alignment and the ConSurf server for conservation analysis. Conservation is shown on a gradient from dark purple to white to turquoise (arbitrary units). The light blue bar indicates model completeness. Regions of interest are highlighted: The N-terminal extension in dark blue, the oligomerization loop and its following helix are indicated in magenta, the mobile flap is indicated in grey and the blue asterix indicates residues belonging to the active site.

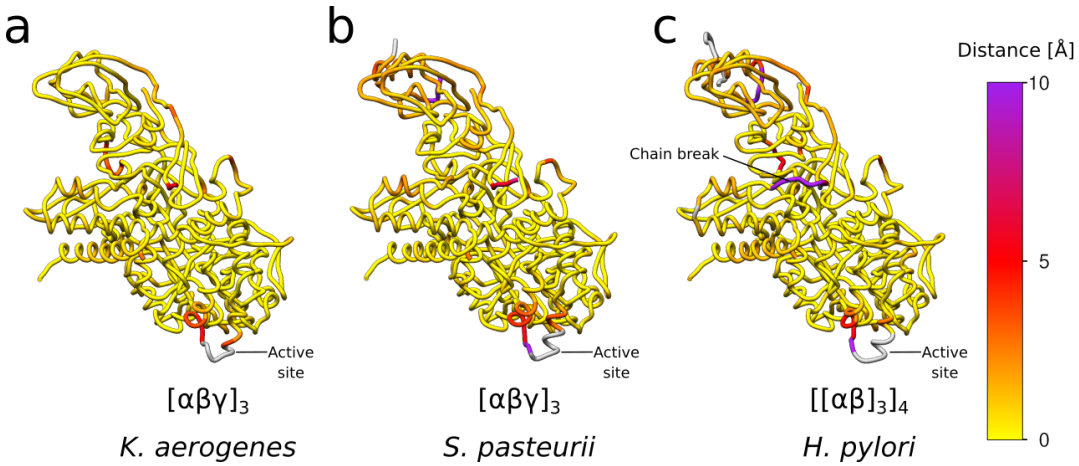

Figure 2.10.6 Distance comparison between residues of *Y. enterocolitica* urease against ureases with different modes of assembly.

Tubes are colored by the pairwise distance of Cα atoms to the corresponding residue in *Y. enterocolitica* urease. Residues without equivalence after sequence alignment are shown in gray color. Segments with particularly high deviations are indicated.

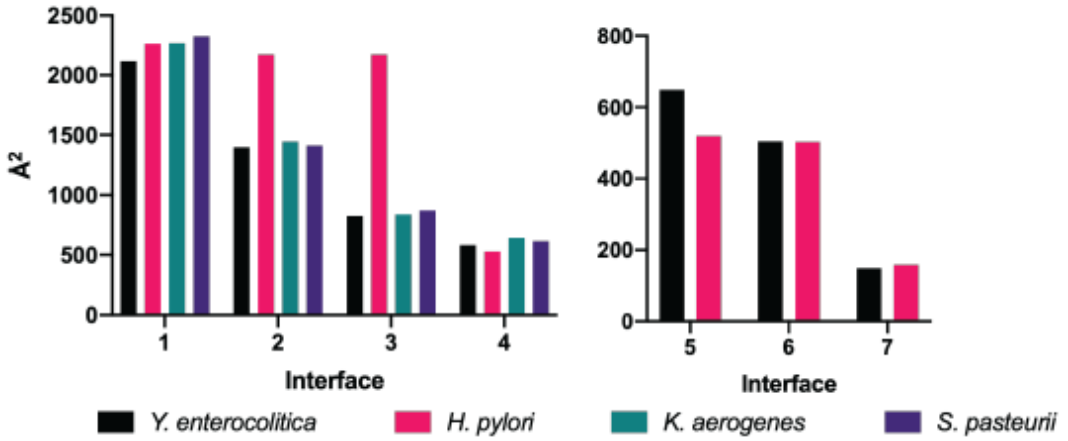

Figure 2.10.7 *Y. enterocolitica* interfaces

Area (Å^2) of interfaces 1-7 (numbering as in Figure 4) plotted by organism. Left: intra-trimer interfaces; right: inter-trimer interfaces. The intra-trimer areas of interfaces 2 and 3 in *H. pylori* are higher because they are part of the same chain.

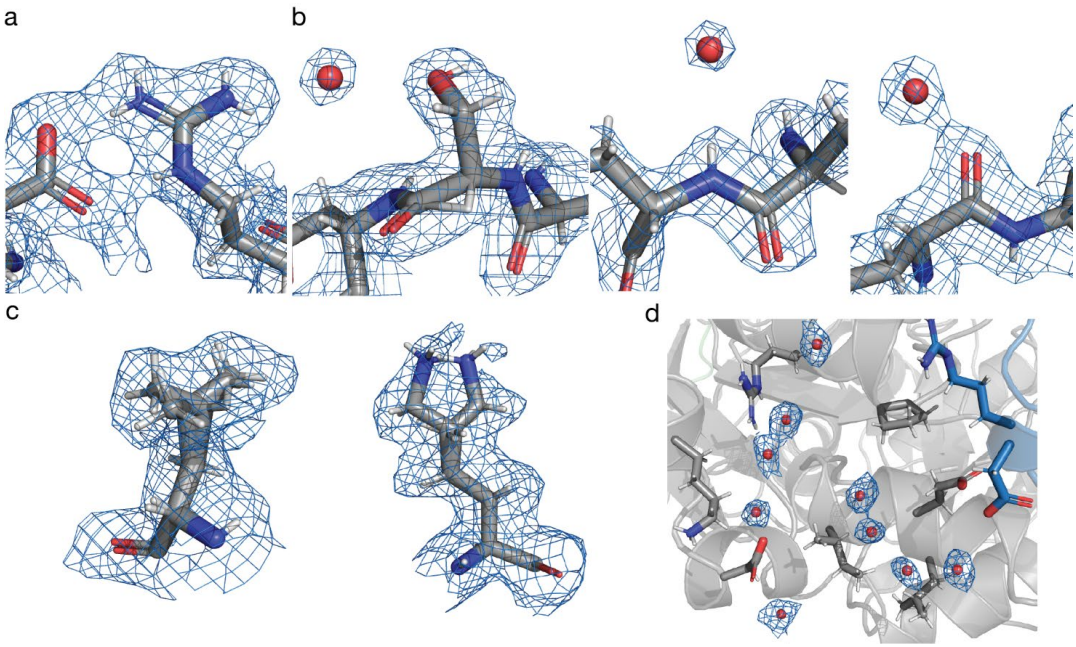

Figure 2.10.8 High resolution structural features

High-resolution structural features can be readily seen in the cryo-EM map. a) Salt bridge between D69 and R63. b) Side chain hydration of S17 and back bone hydration of the oxygen and the nitrogen, respectively. c) alternative side chain conformations of L406 and K2. d) network of waters between two ureC proteins (grey, blue).

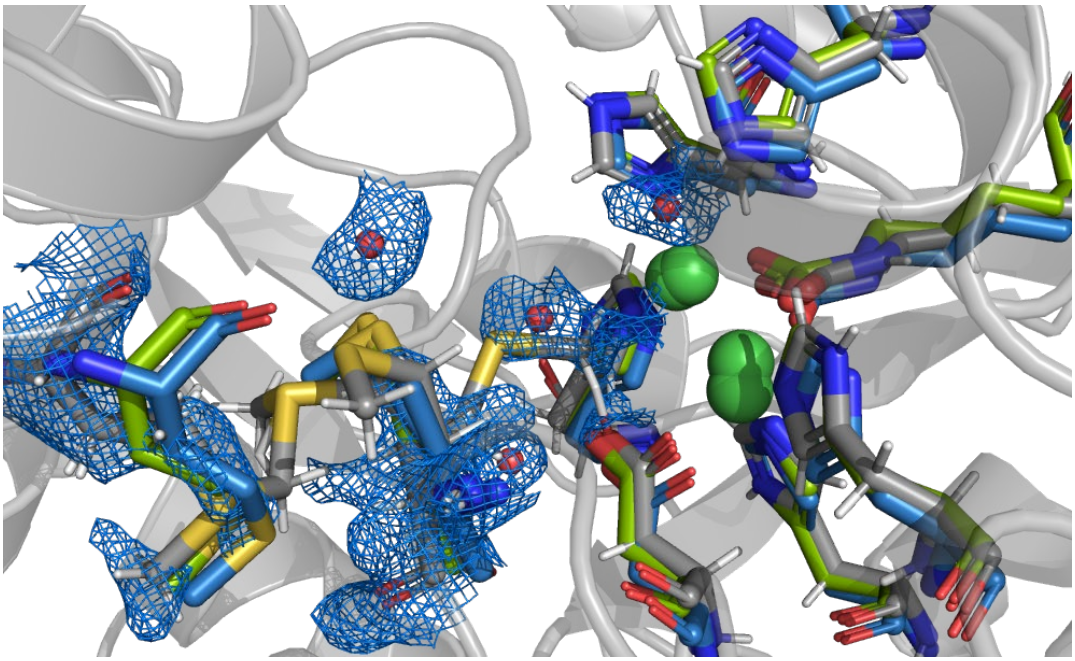

Figure 2.10.9 MET369 can adopt different conformations

MET369 can adopt different conformations in the absence of the substrate or an inhibitor. *Y. enterocolitica* urease (grey), *S. pasteurii* urease (blue), *K. aerogenes* urease (green). One conformation could potentially reach the active site. There is no described function for this amino acid in the catalysis. The conformation close to the active site is only possible because the pocket is empty, which can be seen with an overlay of the active site from *K. aerogenes* urease (PDB: 1EJW) at 1.9 Å and *S. pasteurii* urease (PDB: 5OL4) at 1.28 Å resolution, respectively (Figure 5d,e).

2.11 Supplementary Information

Supplementary Movie 1 Overview and feature highlights of the *Yersinia enterocolitica* urease cryo-EM structure.

Supplementary Movie 2 Morphing between maps and models obtained from different sets of frames along the exposure.

2.12 Supplementary Tables

	Single shot	Multi shot
Data collection		
Microscope	Titan Krios	Titan Krios
Voltage [kV]	300	300
Direct electron detector	Gatan K2	Gatan K2
Zero-loss energy filter	GIF (20 eV slit width)	GIF (20 eV slit width)
Physical pixel size [Å]	0.639	0.639
Super-resolution mode	No	No
Total exposure [$e^-/Å^2$]	42	42
Exposure time [s]	8	8
Number of frames [per movie]	40	40
Movies acquired	2,252	2,243
Beam-image shift?	No	3 shots per hole
Image processing (before merging)		
Movies processed	2,115	2,197
Pixel size [Å]	0.639	0.639
Box size [$pixels^2$]	512	512
Particles picked (Gautomatch w/ templates)	87,204	107,399
Particles after 2D classification	87,032	80,239
Particles after 3D classification	51,173	62,884
Unmasked resolution [Å] (FSC 0.143)	2.52	2.39
Masked resolution [Å] (FSC 0.143)	2.20	2.10
Image processing (after merging)		
Particles in final reconstruction		119,020
Unmasked resolution [Å] (FSC 0.143)		2.20
Masked resolution [Å] (FSC 0.143)		1.98

Table 2.12.1 Cryo-EM data collection and image processing summary

	Multi-shot #1	Multi-shot #2	Multi-shot #3	Single-shot
Beamtilt x *(mrad)*	0.008	-0.218	-0.028	-0.078
Beamtilt y *(mrad)*	-0.091	-0.127	-0.000	-0.065

Table 2.12.2 Beam tilt estimation after merging the two datasets.

	Multi Shot *(before merging)*	**Single Shot *(before merging)***	**Merged**	**Multi Shot #1**	**Multi Shot #2**	**Multi Shot #3**	**Single shot**
Number of particles in training *(approximate)*	5,000	5,000	10,000	10,000	10,000	10,000	10,000
Sigma for Velocity (Å/dose) *(smaller = shorter tracks)*	0.531	0.819	0.779	0.810	0.693	0.636	0.681
Sigma for Divergence (Å) *(higher = more homogeneous tracks across micrograph, i.e. "rigid block")*	3,600	10,455	10,620	10,260	7,350	5,220	8,355
Sigma for Acceleration (Å/dose) *(smaller = straighter tracks)*	2.175	3.735	1.620	1.485	1.425	1.500	0.795

Table 2.12.3 Bayesian polishing training results.

Organism	PDB code	Nr. Of Genes	Stoichiometry α-(β)-(γ)	ureA (%)	ureB (%)	ureC (%)
S. pasteurii	5O4L	3	3-3-3	60.6	46.7	57.5
K. aerogenes	1EJW	3	3-3-3	60.6	52.9	58.7
H. pylori	1E9Z	2	12-12	52.0	50.4	57.6

Table 2.12.4 Sequence identity between different ureases and *Y. enterocolitica* urease.

	Chain A	Chain B	Chain C	Active site
1E9Z	0.860/0.942 (98/100)	0.702/4.424 (107/124)	0.642/1.248 (537/561)	0.634
Alignment Score	339.8	376.8	1874.2	-
5OL4	0.706/0.706 (100/100)	0.756/2.999 (108/120)	0.673/1.171 (542/563)	0.293
Alignment Score	369.3	327.5	2002.7	-
1EJW	0.607/0.607 (100/100)	0.622/1.054 (97/101)	0.611/0.916 (538/559)	0.270
Alignment Score	637.2	269.4	1957.0	-

Table 2.12.5 Comparison of RMSD values of each chain

RMSD values between each chain of our *Y. enterocolitica* urease model versus structures deposited at the PDB. Values are given in Ångstroms. Values in parenthesis indicate the number of Cα atom pairs matched with and without pruning. The program matchmaker from UCSF Chimera was used to calculate sequence alignments and RMSD values between each chain in respective models.

2.13 Supplementary Note

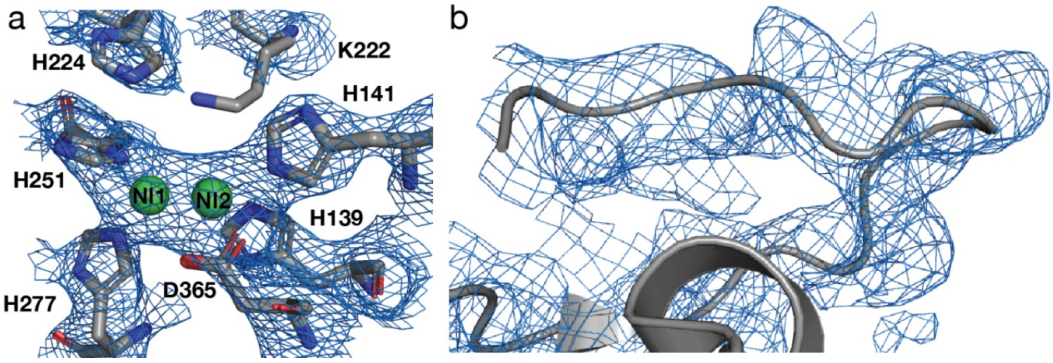

Figure 2.13.1 Low resolution crystal structure of *Y. enterocolitica* urease.

Structure determination by X-ray crystallography yielded a 3.01 Å resolution density, which was sufficient to see the higher oligomeric state of *Y. enterocolitica* urease, but not for detailed analysis of the assembly mechanism and active site.

Comparison of the model based on the high-resolution cryo-EM data *vs.* the previous crystal structure shows a few differences. The previous X-ray model did not include residue 100 of ureA and residues 31-33, 148-162 of ureB. Residues 148-162 form a helix, which was previously not observable. This helix sits right at the interface between three ureB proteins and constitutes the interaction between the trimers.

2.13.1 Supplementary Methods

For X-ray crystallographic analysis of the *Y. enterocolitica* urease, *Y. enterocolitica* was differently purified than described in the Methods for cryo-EM. The protein was precipitated using 40-60% w/v ammonium sulfate, resuspended and dialyzed in 0.15M NaCl, 50mM Tris pH 8.0. It was further purified using a 45 ml self-packed DEAE FF XK26/20 (Sigma, DFF100) and a Superdex 200 10/300 GL column. The SEC was also used for buffer exchange to 20 mM HEPES, 100 mM NaCl pH 7. Urease crystals grew at 20°C at 10 mg/ml in 0.1M CHES pH 9.5; 50% (v/v) PEG 200. Urease crystals belonged to space group H32 with unit cell parameters of a= 157.2 Å, b = 157.2 Å and c= 774.6 Å, with four molecules per asymmetric unit. The structure was determined by molecular replacement with PHASER [200] using the urease crystal structure from *Klebsiella aerogenes* (PDB: 1FWB) [201]. Model building and structure refinement were performed with Coot [192] and Buster-TNT [202]. The atomic coordinates for this model have been deposited in the Protein Data Bank under the accession code 4Z42.

3 The mechanism of citrate-induced filamentation of acetyl-CoA carboxylase 1

Unpublished manuscript

Leonie Anton, Moritz Hunkeler, Anna Hagmann, Edward Stuttfeld, Timm Maier

The mechanism of citrate-induced filamentation of acetyl-CoA carboxylase 1

3.1 Contributions to the manuscript

I cloned, expressed and purified the point mutants and did all experiments pertaining to their folding and stability. I did sequence alignments and analyzed the conservation. I conducted all radiolabeled activity and mass photometry assays. I prepared grids for cryo-EM, screened for best grid candidates, assisted in data collection and processed and refined all the data in cryoSPARC and RELION. I placed the citrate in the model and refined it against the cryo-EM map. I wrote the manuscript and prepared the figures. Contributions of the other authors is indicated in author contribution section.

3.2 Abstract

Acetyl-CoA carboxylase (ACC) catalyzes the first and rate-limiting step in fatty acid biosynthesis. Acetyl-CoA is carboxylated in two distinct and consecutive reactions resulting in malonyl-CoA, the committed substrate of fatty acid synthase. ACC is a multienzyme, which contains both catalytic sites on a single polypeptide chain and assembles a dimer. Over 50 years ago, it was observed that the feed-forward allosteric activator citrate leads to filament formation of higher eukaryote ACC (ACC-Cit). However, the mechanism behind the citrate-dependent assembly of ACC filaments has remained an enigma. A 3.8 Å resolution cryo-EM reconstruction of ACC-Cit enabled the identification of a positively charged pocket in the non-catalytic central domain (CD) as the place of allosteric activation by citrate. Point mutants in the pocket disrupt citrate-dependent ACC activation and filament formation. Presence of citrate causes a cascade of domain rearrangements in the CD, creating a docking platform for the next ACC dimer to bind, resulting in ACC-Cit filament formation.

3.3 Introduction

Eukaryotic acetyl-CoA carboxylase (ACC) is a dimeric multienzyme catalyzing the committed and rate-limiting first step in fatty acids synthesis. Under the consumption of ATP, ACC carboxylates acetyl-CoA to produce malonyl-CoA in a two-step reaction [63]. The first reaction is the transfer of a carboxyl group from bicarbonate to the biotin cofactor of the biotin carboxyl carrier protein (BCCP) domain by the biotin carboxylase (BC) domain of ACC. The BCCP domain shuttles the carboxy-biotin intermediate to the carboxyl transferase (CT) domain, which transfers the carboxyl group onto acetyl-CoA producing malonyl-CoA [67, 68]. Besides the catalytic domains, eukaryotic ACC also contains a central domain (CD). The CD connects the BC and CT domains and is comprised of four domains CD_N (N-terminal), CD_L (linker), CD_{C1} (C-terminal 1), CD_{C2} (C-terminal 2) (Figure 3.3.1 a). Human ACC exists in two isoforms (ACC1 and ACC2), which share a sequence identity of 74.6%. ACC1 is the main metabolic enzyme and is expressed in lipogenic tissues, including adipose tissue and mammary glands. The malonyl-CoA produced by ACC1 is a substrate of fatty acid synthase (FAS), which produces long chain fatty acids. ACC2 acts as a regulator of beta-oxidation in oxidative tissues, such as heart and skeletal muscle, by producing malonyl-CoA for inhibition of carnitine palmitoyl-transferase (CPT1) [78, 203].

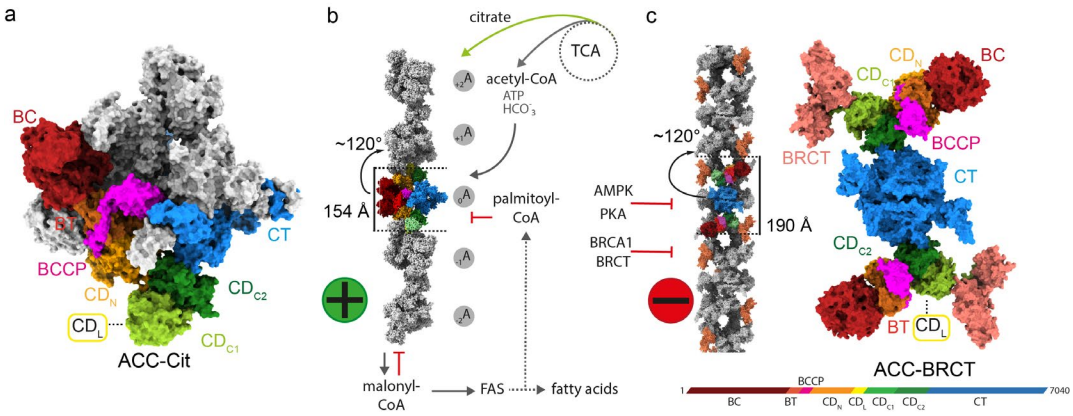

Figure 3.3.1 Filament formation of ACC upon activation and inhibition.

a) Conformation of ACC dimer upon allosteric activation by citrate. Surface representation of one protomer is shown in grey and the other color coded according to domains. b) ACC-Cit filament with one dimer in color. Rise and twist are indicated. Regulation scheme of ACC by feedforward activators and feedback inhibitors. c) Regulation by phosphorylation and protein interaction partners. ACC-BRCT filament with one dimer in color and all BRCT domains in color. Focus on ACC-BRCT dimer with domain names indicated. Linear domain architecture of ACC at the bottom

ACC1 expression in humans has been associated with manifestations of the metabolic syndrome such as obesity, type 2 diabetes, cardiovascular disease and arteriosclerosis [108, 109] and a knockout of the coding ACACA gene is embryonically lethal [103]. ACC is overexpressed in liver, breast and prostate tumors and is a cancer drug target [83, 115, 116, 204].

Regulation of human ACC1 occurs transcriptionally and post-translationally. The transcriptional regulation is controlled by the master regulators of metabolism, the mTOR complex 1 and 2, [83, 84] through SREBP1 (Sterol Regulatory Element Binding Proteins) and ChREBP [80-82]. The AMP-activated kinase (AMPK) and cAMP-dependent kinase (PKA) have been shown to phosphorylate ACC1. AMPK phosphorylates Ser80, Ser1201 and Ser1216, while PKA phosphorylates Ser78 and Ser1201. Phosphorylation of Ser80 and Ser1201 have a strong inhibitory effect on activity of ACC1 [88, 89]. ACC1 is further regulated by direct protein-protein interactions. One study reported that Mig12 can promote ACC activation by lowering the threshold for required citrate [94]. Another study shows that in complex with its paralog Spot14, they attenuate filament formation [95, 96]. A distinct type of filament is formed by ACC1 when interacting with BRCT, the C-terminal tandem domains of tumor-suppressor BRCA1. Binding of BRCT locks ACC in an open and inactive conformation (Figure 3.3.1 c) [141].

Feedback inhibition by long acyl-CoA chains is an essential regulatory mechanism of ACC, signaling a surplus of fatty acids [205]. It is not clear how the inhibition is mediated and for the inhibitor palmitoyl-CoA the specificity of this inhibition is a topic of debate [86, 87] (Figure 3.3.1 b).

The allosteric activator citrate induces filament formation in ACC1, which significantly increases activity [63, 77]. The formation of ACC filaments upon addition of citrate was first observed over 50 years ago [51]. However, only recently, progress towards the understanding of ACC1 activity and filament formation has been made. The cryo-EM structure of the ACC1 filament (ACC-Cit) was previously reported by our group and shows a helical assembly of dimers in a closed and perpetually active conformation (Figure 3.3.1 b) [141]. This assembly locks BC domains in dimer form, which is essential for activity [134]. The closed arrangement also ensures accessibility of the BCCP to both catalytic sites leading to high activity [141]. While this reconstruction at 5.4 Å-resolution allowed unambiguous placement of crystal structures of all domains, identification of small molecules and their binding sites was not possible. Until now the mechanism with which citrate induces these massive conformational changes is unknown.

S. cerevisiae ACC (SceACC) adopts a similar closed conformation when crystallized in presence of citrate [135]. SceACC is not activated by citrate and filaments formation has not been reported, however the closed conformation is stabilized by citrate [206]. Comparison of the CD interfaces of SceACC and human ACC from ACC-Cit filaments, shows that the CD conformation of SceACC is incompatible with filament formation [141]. Several regulatory phosphosites have been identified in SceACC in phosphoproteomic studies. Ser1157 gets phosphorylated (pSer1157) by Snf1, a homologue of AMPK, which leads to a stark reduction in activity [91-93]. Structural analysis showed that pSer1157 is located in the CD and is coordinated by Arg1173 and Arg1260 in a positively charged pocket between the CD_{C1} and CD_{C2}. Mutational studies of these residues showed reduced capacity to inhibit SceACC activity [206].

3.4 Results and discussion

3.4.1 3.8 Å resolution of ACC filament enables visualization of positively charged pocket

Previously, we have determined the cryo-EM structure of citrate-induced ACC filaments at a resolution of 5.4 Å using a single particle-approach [141] (PDB ID:62GD). Collecting additional data, as well as employing recent algorithms for the correction of aberrations introduced by the optical system, resulted in an overall resolution of 3.84 Å (Figure 3.9.1). The highest resolution is present in the highly ordered CT domain and the resolution decrease towards the edges of the map (Figure 3.9.1). The biotin moiety attached to a lysine in the BCCP is buried at the active site of the CT, which shows the potential of the map to visualize small molecules (Figure 3.9.2 a). 3D classification revealed that the angles between the dimers within the filament are not identical, which is in part responsible for the strongly decreased resolution of the CD (Figure 3.9.1; Figure 3.9.2). The difference in angles can be caused by the filament being inherently flexible and dynamic at the interfaces. The varying angles could also represent static states of filaments that adopt distinct conformations. The partial or full occupancy by citrate, can be an influencing factor as well.

Large conformational changes are necessary to go from an open inactive conformation to a closed active conformation. The CD has already been described as capable of undergoing large-scale hinge-like motions, capable of rotating (CD_N can be in positions at least 160° rotated from each other [139]) to accommodate BC dimerization (Figure 3.4.1 a). Citrate could directly bind somewhere in the CD domains, bind at the interface between two dimers in a filament or bind between two domains of one dimer to fix them in place. Comparison of CD domains of ACC-Cit and ACC-BRCT show highly similar folds for CD subdomains, but their relative arrangement changes (Figure 3.4.1 b).

Citrate is a negatively charged molecule, suggesting it could be recognized via a cluster of positive charges. The pocket identified for the binding of pSer1157 in SceACC meets these requirements (Figure 3.4.1 c). Closer inspection of the same positively charged cluster in the full length SceACC X-ray structure reveals a large unassigned electron density in between the arginines, which could potentially correspond to citrate (Figure 3.9.3). This cluster is highly conserved among eukaryotes including human ACC (Figure 3.9.4). Residues Arg1232 and Arg1326 in human ACC are Arg1173 and Arg1260 in yeast, respectively.

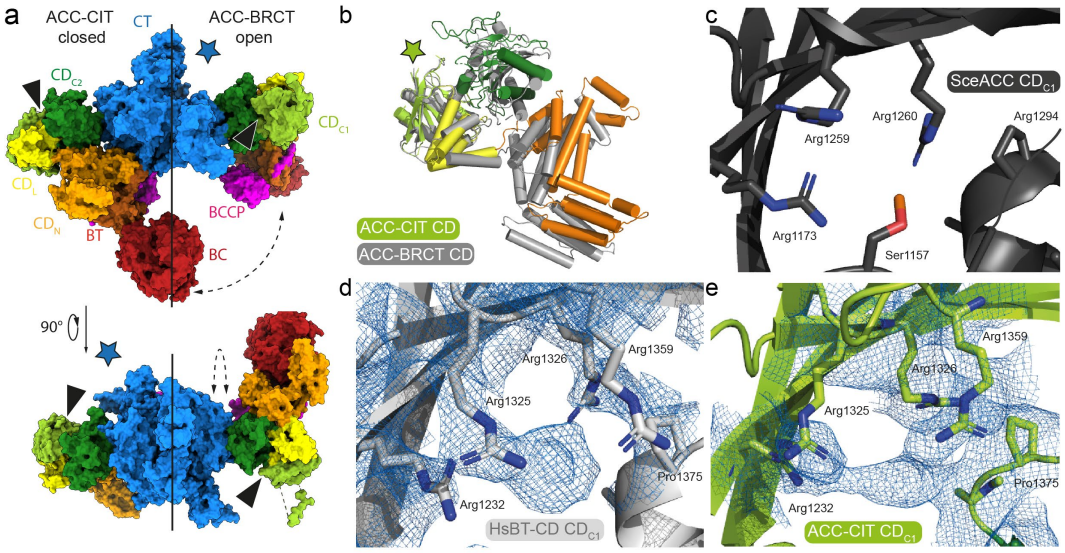

Figure 3.4.1 Conserved positively charged pocket is located in CDC1 domain.
a) Surface rendering of aligned ACC-Cit and ACC-BRCT protomers, both colored according to domains. Protomers have been aligned by their CT domains, which is indicated by a star in color of CT. Black triangles indicate location of conserved pocket. Arrows indicate the domain movements needed for the other conformation to be adopted. b) Overlay of ACC-Cit and ACC-BRCT CD aligned at the CD_{C1} domain in cartoon with cylindrical helices. c) Conserved pocket of SceACC with phospholoop. Part of the loop is omitted for clarity (PDB ID:5I6E). d) Shows conserved pocket in all-atom model of HsBT-CD crystal structure with the electron density as mesh. e) Same pocket as in d), but in ACC-Cit filament and with cryo-EM density shown as mesh.

An X-ray structure of the human CD fused to the BT (HsBT-CD) was previously solved to 3.7 Å resolution in presence of citrate and was interpreted as a poly-alanine model [139]. Further rounds of iterative model-building and refinement now provide an all-atom model and allow more detailed investigation of the conserved pocket, where extra electron density in a cluster of positive residues is observed (Figure 3.4.1 d). Because of the low resolution in the HsBT-CD crystal structure, the electron density cannot be unambiguously identified as citrate. Notably, crystallization of this construct was only possible in the presence of citrate and the other components of the crystallization buffer seem unlikely candidates for binding. Investigating the same pocket in cryo-EM reconstruction of ACC-Cit filament, extra density is seen in the same location (Figure 3.4.1 e, Figure 3.9.4). However, like for the crystal structure of the BT-CD construct, the resolution does not allow for unequivocal placement and orientation of citrate. Besides citrate, the sample buffer additionally contains also β-mercapoethanol (β-ME), n-Dodecyl β-D-maltoside (DDM) and HEPES. While DDM and β-ME can be excluded as binding in the

positively charged pocket, HEPES at pH 8 is anionic and could bind the positively charged sidechains. Additional experiments are necessary to confirm that it is in fact citrate acting on the pocket.

3.4.2 Citrate affects a conserved pocket in the central domain region

Six residues of the cluster are highly conserved across species, indicating evolutionary pressure to maintain this pocket. Specifically, residues Arg12, Arg1326, Tyr1158, Tyr1161, Arg1325 and Pro1375 show high conservation (Figure 3.9.4). To test for a role of this pocket in citrate mediated activation, Arg1232 and Arg1326 were mutated to alanines (R1232A/R1326A) (Figure 3.4.2 a).

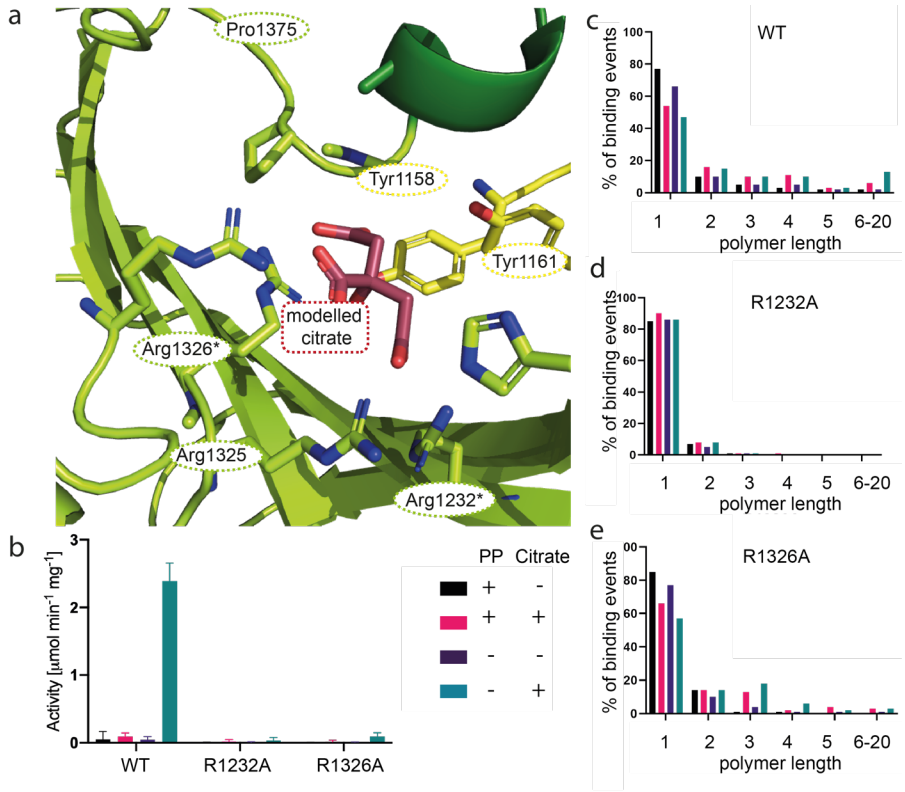

Figure 3.4.2 Point mutations impair citrate-dependent activation and filament formation.

a) Manually modelled citrate in the positive cluster with conserved residues labeled and shown as sticks. Mutated residues are indicated with a star. b) Results of four separate activity assays done in triplicates. Error bars show standard deviation. c-e) Mass photometry results displayed as bar graphs. Polymer length refers to number of dimers. Color legend for graphs indicated at the bottom right. PP=phosphorylated.

Purified wild-type (WT) and mutant ACC are all present as more than 70 % dimers in inactivated state as can be seen by mass photometry (Figure 3.9.7) and run as a single band on a SDS-PAGE gel (Figure 3.9.5). Thermal shift assay of all proteins resulted in equivalent curve shapes of raw fluorescence. Melting points were calculated from the first derivative and were at 48°C for WT ACC, 47.5°C and 48.5°C for R1232A and 46 and 45 °C for R1326A. The two melting points refer to phosphorylated and dephosphorylated sample, respectively. It was concluded that the mutations do not affect expression and folding of the protein (Figure 3.9.5).

To evaluate the impact of the mutations on citrate-induced activation, an assay measuring the incorporation of radiolabeled carboxyl groups into malonyl-CoA was used. In four independent experiments with three repeats each, a reduction of activity for the mutants was seen (Figure 3.4.2 b). Dephosphorylated WT ACC in presence of citrate had the highest activities at 2.4 µmol min^{-1} mg^{-1}. The activities of phosphorylated WT ACC with citrate and dephosphorylated WT ACC without citrate were lower, while phosphorylated WT without citrate had almost no activity. R1232A and R1326A had very low activity in all conditions, with between 0 and 0.09 µmol min^{-1} mg^{-1} (Figure 3.4.2 b). These results indicate, that mutations in the described pocket lead to low activity also in the presence of citrate.

To investigate whether the mutations also affect filament formation we used mass photometry to assess the particle size distribution based on single molecule measurements of ACC variants in the presence of citrate and phosphorylation. The ACC dimer has a molecular mass of about 520 kDa and polymer length is indicated by calculated number of dimers per particle (Figure 3.4.2 c, d, e).

WT ACC displays 77 % dimers for phosphorylated condition without citrate compared to 47% dimers in dephosphorylated/with citrate. All conditions show progressively lower percentage of binding events per polymer length until polymer lengths of 6-20 dimers. Phosphorylated and dephosphorylated ACC in presence of citrate show 6 and 13% binding events at high polymer lengths, respectively (Figure 3.4.2 c). R1232A shows 85-90% polymers of single dimer weight, 5-8% with two dimers and <1 % for the other polymer lengths in all conditions. The mutation of R1232 to alanine appears to completely abrogate citrate-induced filament formation (Figure 3.4.2 d). 18% of dephosphorylated citrate induced R1326A are present as polymers of 3 dimer length compared to the 10% in WT. The mutant also assembles polymers of 6-20 dimers length, but only 3% of the total. This could hint at the filament formation being impaired but not completely abrogated as for R1326A and

the mutant could still assemble shorter filaments (Figure 3.4.2 e) (For all plots see Figure 3.9.7, Figure 3.9.8). More repetitions of the mass photometry experiments are needed to confirm these findings.

In summary, the results of the activity and filamentation assays demonstrate that mutant R1232A is indeed insensitive to activation by citrate. R1326A shows reduced sensitivity to citrate but is still able to assemble shorter polymers.

3.4.3 Repositioning of loop leads to rearrangement of domains in CD

The key factors contributing to ACC activity are the dimerization of BC domains and allowing the BCCP to reach both catalytic sites [141]. In yeast ACC as well as human ACC, the CD domain acts as the hinge allowing for flexible movement of the BC domains [139, 141, 206]. To analyze the effect of putative citrate binding on structural rearrangements, the open ACC-BRCT dimer and the closed ACC-Cit dimer are overlaid [141] (Figure 3.4.3 a-c). The CD_N is an alpha-helical domain with 11 helices with a four-helix bundle and a helical hairpin. and the CD_L is a linker domain containing a 4 helix bundle. The CD_{C1} and CD_{C2} adopt an alpha-beta fold with a six stranded beta-sheet with two long helices on one side. CD_{C2} is C-terminally extended by a b-strand and b-hairpin (Figure 3.9.6) [139]. The six stranded beta-sheet (beta 1-6) of CD_{C1} contains the positively charged pocket coordinating citrate. When overlaying the CD_{C1} domains of ACC-Cit and ACC-BRCT filament there is a 33° turn of the CD_{C2} relative to the CD_{C1} (Figure 3.4.3 a). This rotation occurs at the small helix after the loop connecting the domains. Binding of citrate could potentially interfere with the helix and reposition it inducing the rotation (Figure 3.4.3 a). CD_{C1} and CD_{C2} rotating against each other changes the position of the CD relative to CT. CD_N is now close to the CT and BT and needs to rotate another 37° to interact with the domain and close the dimer (Figure 3.4.3 b, c). The BC domains would be close enough to each other to dimerize. The repositioning of the whole CD creates the interface needed for filament formation (Figure 3.4.3 d).

The CD domain creates a platform for docking of another ACC dimer in a similar conformation. It is also possible that the citrate alone does not yet lead to a completely closed dimer but pushes the conformational change far enough so that another ACC dimer can bind and complete the dimerization.

One of the hurdles in the investigation of ACC is the separation of activity and filamentation. With a suggested binding site for citrate and a 3.8 Å reconstruction,

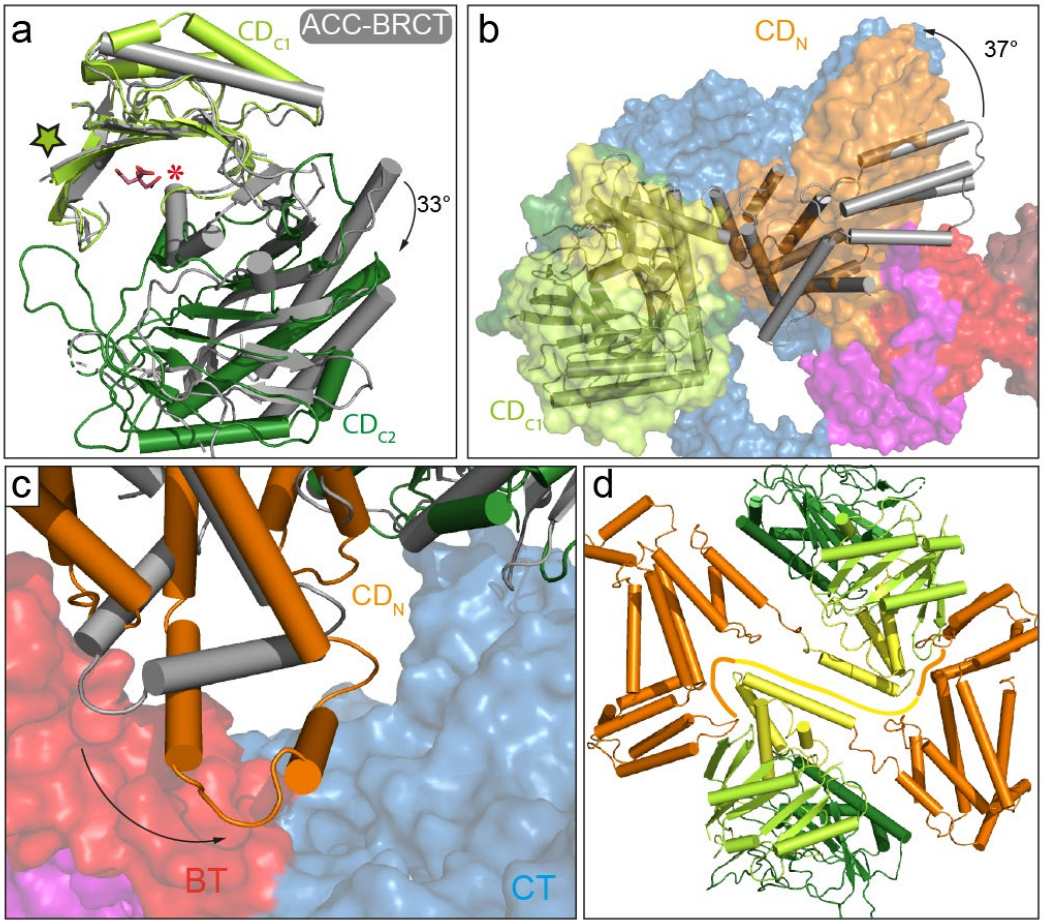

Figure 3.4.3 Citrate binding induces domain rearrangements.

a) ACC-Cit (color) and ACC-BRCT (grey) CD_{C1} and CD_{C2} aligned at CD_{C1}. Shown in cartoon representation with cylindrical helices. Repositioned helix indicated with a red star. Rotational movement shown with arrow. b) Transparent ACC-Cit protomer with surface representation aligned at CD_{C1} with ACC-BRCT CD in grey cartoon style with cylindrical helices. Domain rotation indicated with arrow. c) Close up of CD_N with cylindrical helices and the interaction with CT and BT domains as transparent surface rendering. ACC-BRCT is in grey. d) Interface between two CD domains in filamentation competent state indicated with colored lines.

future studies can be designed to uncouple those events. Direct citrate binding assays are required to confirm the cluster as the responsible site for citrate-mediated activation. To exclude other potential sites, mutated ACC must also be investigated for its ability to directly bind citrate. Protein constructs of the CD_{C1} and CD_{C2} domains can be used to uncouple the binding event from its effects on activity and filamentation. A $CD_{C1}CD_{C2}$ protein construct in presence of citrate is potentially more stable than BT-CD and merits another crystallization trial. Based on the higher-resolution cryo-EM data, mutations of the interface can be designed to

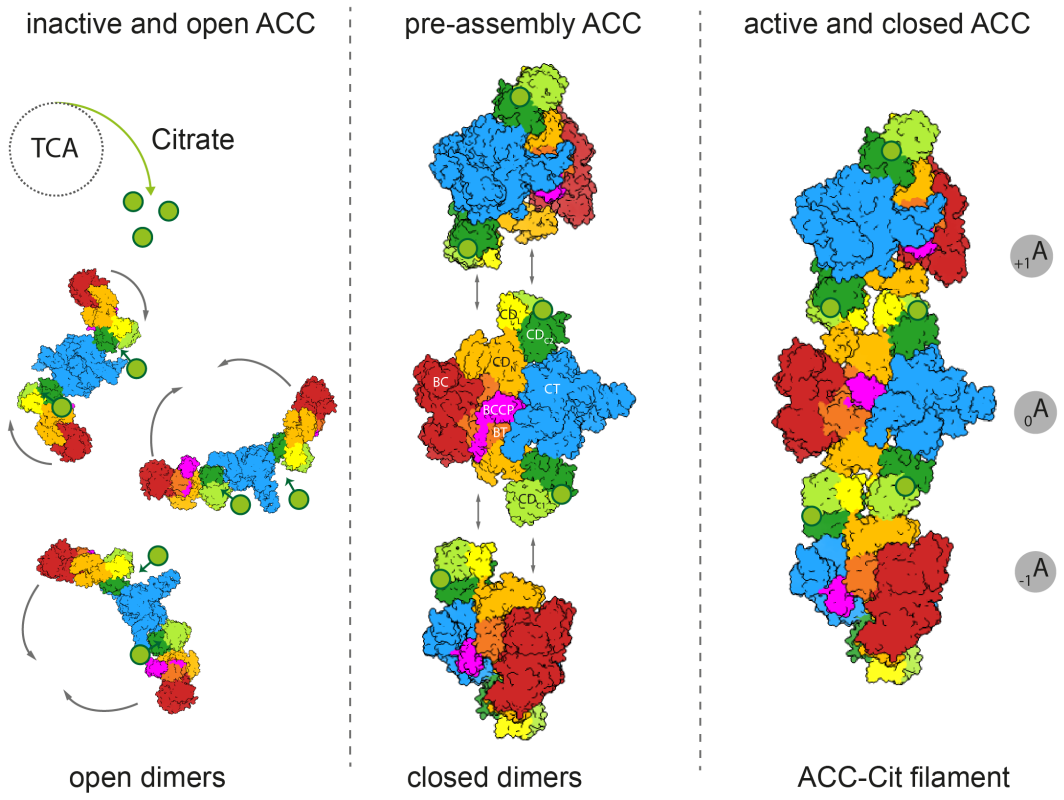

Figure 3.4.4 Mechanism of citrate-induced filament formation.
Citrate allosterically activates ACC by binding to the conserved pocket in the CD. This causes large scale domain rearrangements making ACC dimers competent to stack into a filament. The filament locks ACC in a perpetually active and closed conformation.

disrupt filament formation revealing the contributions of filament formation to increased activity.

Because of limitations in resolution the citrate could not be unambiguously oriented in the pocket. The proposed location is a conceivable model, but molecular dynamics simulations, in absence of higher resolution experimental data, will be ideal to gain deeper insights into the dynamics of the CD region in dependence of citrate binding. The changes for the CD_{C1} and CD_{C2} and the whole ACC dimer could be simulated.

3.5 Conclusion

Improved cryo-EM reconstructions from 5.4 Å to 3.8 Å enable the visualization of a conserved pocket in the non-catalytic CD region of human ACC filaments. Mutations of residues forming the cluster lead to abrogated activity and reduced filament formation in presence of citrate. In SceACC, this positive cluster can be

tightly bound by a pSer on a loop, which leads to inhibition of SceACC activity. pSer prevents movement of the CD_{C1} and CD_{C2}, adopting a conformation very similar to CD_{C1} and CD_{C2} in the inactive ACC-BRCT filament. Potential citrate binding in the same pocket could change the position of the adjacent helix. The repositioning of the helix then causes a cascade of domain rearrangements. The resulting ACC dimers adopt a closed and active conformation where the CD domains are compatible with filament assembly (Figure 3.4.4).

After 50 years of research into the mechanism of citrate-induced filament formation, we present the first tangible model for the assembly of ACC higher oligomers. This mechanistic model allows for the design of future studies with the goal of furthering our understanding of enzyme regulation through oligomerization and the potential of the allosteric site as a drug target.

3.6 Author contributions

L.A. collected and processed cryo-EM data and performed biochemical and biophysical assays. M.H. cloned and purified ACC WT protein. A.H. collected cryo-EM data sets. E.S solved the X-ray crystal structure and built an all-atom model of human CD. T.M. conceptualized the project. L.A. and T.M. wrote the manuscript and prepared figures.

3.7 Acknowledgement

We would like to thank the staff from the Swiss Light source beamline and Biozentrum BioEM facility for their excellent support. Cryo-EM data processing calculations were performed at sciCORE (http://scicore.unibas.ch/) scientific computing center at the University of Basel. L.A. and A.H. acknowledge funding from the Fellowships for Excellence program sponsored by the Werner-Siemens Foundation and the University of Basel.

We would like to thank the Biozentrum proteomics core facility for their support and help with measurement and analysis of phosphorylation states of ACC. We would also like to thank Katharina Häussermann from Refeyn and all of Refeyn staff for the demonstration and use of the instrument.

Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at UCSF.

3.8 Materials and Methods

3.8.1 Cloning and protein purification

The published ACC vector was used for site directed mutagenesis using primers containing the desired mutation followed by blunt end ligation. Primers used to clone R1232A mutation: forward (fw) TCACTCCACCTTGTCAGGCAATG GGCGGAATGGTCTC, reverse (rv) GAGACCATTCCGCCCATTGCCTGACAAG GTGGAGTGA; R1326A mutation: fw GTTGACCATGGGATCCGGGCACTTACTTT CCTGGTTGCA; rv TGCAACCAGGAAAGTAAGTGCCCG GATCCCATGGTCAAC. Mutants were confirmed by sequencing.

Purification, *in vitro* biotinylation and dephosphorylation was conducted as described [141]. Biotinylation was confirmed by using streptavidin antibody from abcam. Dephosphorylation was confirmed by mass spectrometry (Method 3.10).

3.8.2 Sample preparation and cryo-EM collection

ACC WT protein was dialyzed overnight against a buffer containing 50 mM Hepes/KOH pH 7.5, 10 mM K_3citrate, 0.1 mM EDTA, 5 mM β-ME. Protein was diluted in the buffer without β-ME to a concentration of 400 µg/ml and 17 µM DDM were added. C-Flat 1.3/1.2 200 mesh grids from Electron Microscopy Sciences were prepared using a FEI Vitrobot MarkIV with 4 µl of protein solution (4°C, 100% humidity, double blotting paper, 10 sec wait time, 3 sec blot time). Samples were imaged using an FEI Titan Krios equipped with a Gatan image filter (Quantum-LS GIF, 20 eV zero loss filtering) and a post-GIF K2 summit direct electron detector (Gatan). Movies were recorded at 300 kV with a pixel size of 1.05 Å and 50 frames at ~1 e^- $Å^{-2}$ per frame (yielding a total dose of 50 e^- $Å^{-2}$). Sample preparation and collection parameters of the previous 13,671 movies (here called datasets 1-4) can be found in publication [141].

3.8.3 Data processing

Collection parameters of the previous 13,671 movies (datasets 1-4) can be found in publication [141]. For the new data collection 3’746 movies were recorded using SerialEM (dataset 5) [166]. Processing for all datasets (DS) was done in cryoSPARCV2 and the same strategy was used unless indicated otherwise [152]. Using cryoSPARCs patch motion beam-induced motion was corrected for and patch CTF was used to correct for differing defocus values across one micrograph.

Poor quality micrographs were discarded at this step. Particles were picked using a 2D template made from previous ACC-Cit reconstructions and employing the cryoSPARC template picker. The picked particles were classified into 2D classes and low quality particles were sorted out, which resulted in a total of 467'027 useful particles (DS1: 146'013; DS2: 154'994; DS3: 43'698; DS4: 30'153; DS5: 92'169). Previous ACC-Cit reconstruction (EMD-4342) was as used as a starting model for refinement. DS1-4 were strongly affected by optical aberrations of the microscope [141]. In recent publications, algorithms for the correction of the symmetrical and antisymmetrical aberrations and magnification anisotropy were described [153, 167]. These algorithms were implemented in cryosparcV2.1[207]. Through iterative local and global CTF refinements the resolutions of the 4 datasets were improved by 0.2-0.8 Å (Figure 3.9.1). DS1-5 were merged together and refined as one data set. 3D classifications showed heterogeneity in the angles of filaments (Figure 3.9.2 b). Number of particles in the separated classes were too low, to separately calculate another reconstruction. Using all particles non-uniform refinement and local masking of one dimer, an overall resolution of 3.84 Å was reached, based on the FSC 0.143 threshold criterion [185].

3.8.4 Model refinement cryo-EM

Phenix_real_space_refine (version 1.17) was used to refine the model against the new map. Citrate molecule was manually place into the binding pocked and oriented based on the best fit into the density and our assumed interactions with sidechains. This new model containing citrate was again refined in phenix. The final model of ACC-Cit contains residues 102–511; 524–543; 556–1188; 1230–1256; 1284–1333; 1352–1518 and 1525–2338.

3.8.5 Crystallization and X-ray structure

Purification, crystallization and data collection of HsBT-CD can be found in publication [139].

3.8.6 Activity assay

The catalytic activity of ACC was measured by following the incorporation of radioactive ^{14}C into acid-stable non-volatile product. The assay was preformed according to protocol from [141]. Measurement was carried out four times in three

replicates each and catalytic activities were calculated using a standard curve derived from measurements of varying concentrations of $NaH^{14}CO_3$ in reaction buffer (established in [141]).

3.8.7 Mass Photometry

In mass photometry the size of the molecule or complex is determined by the scattered light of a molecule at the moment of impact on a glass slide. The scattered light correlates linearly with mass and can therefore be used to weigh complexes [208]. The impact on the glass slide is called a binding event and they can be separated by their different masses. The limit of accuracy of the instrument at the moment is around 3 MDa, therefore from 6 dimers onwards there are still binding events detected but they cannot be properly separated anymore. The highest recorded masses were around 10 MDa corresponding to 20 dimers in one molecule so all binding events from 6 to 20 dimers are reported in one category.

ACC WT and mutant proteins were diluted to 1 µM in buffer containing 20 mM bicine pH 8, 150 mM NaCl, 5% Glycerol, 5 mM TCEP and 8 mM K_3citrate if indicated. Samples were incubated for 30 min-2h and the diluted 1:50 for measurement. Except dephosphorylated WT with citrate, which could only be diluted 1:10. This was to get comparable amounts of binding events per sample. The number of binding events should be similar between different samples to better compare them. Too many binding events crowd the glass slide and limit accuracy of the detection. Too few binding events will limit the amount of information you get out of the sample. The Refeyn One^{MP} instrument was used for measurements [209]. The presented data is from a single experimental set up (see Figure 3.9.7 and Figure 3.9.8).

3.8.8 Data analysis and presentation

Structural figures were prepared using ChimeraX [199] and PyMol [210], and graphs were made in GraphPad Prism version 8 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.

3.9 Supplementary Data

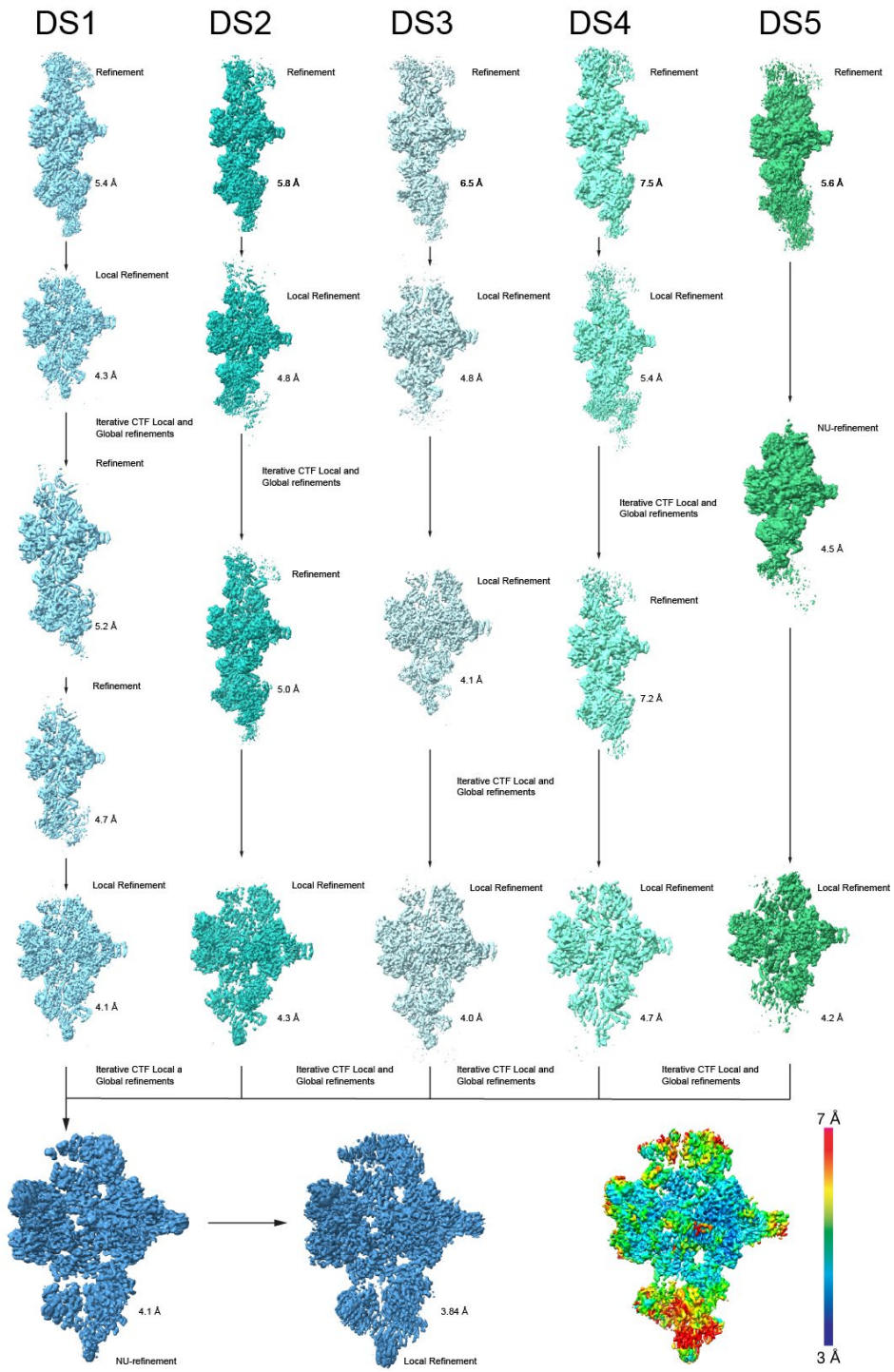

Figure 3.9.1 Processing scheme for ACC-Cit.

All processing was done in cryoSPARCV2.1-2.15.

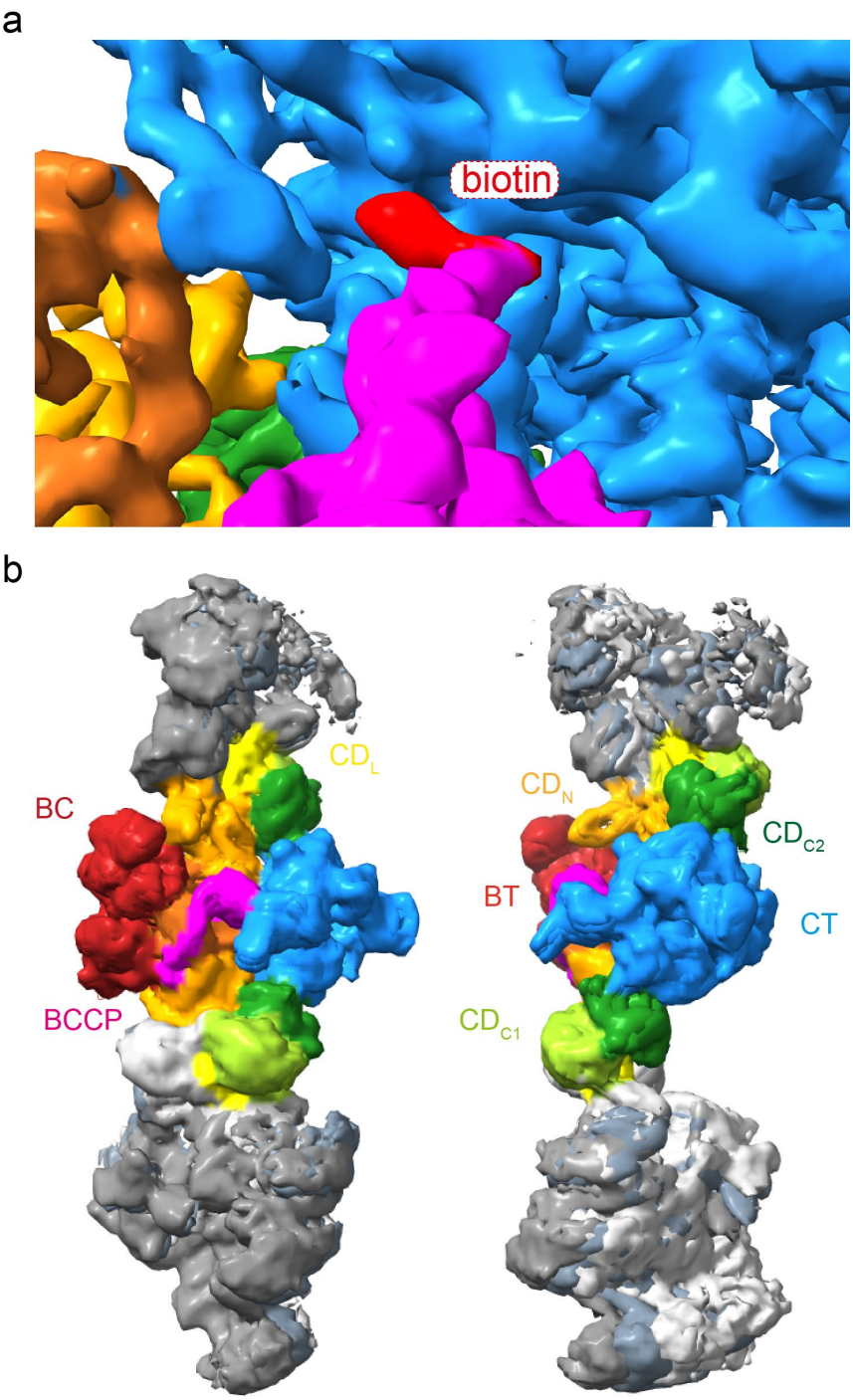

Figure 3.9.2 Visualization of biotin and filament heterogeneity.
a) Biotin moiety (red) visualized in the CT of the ACC-Cit filament. b) 3D-classification shows heterogeneity in classes and distinct angles between two ACC dimers with the filament.

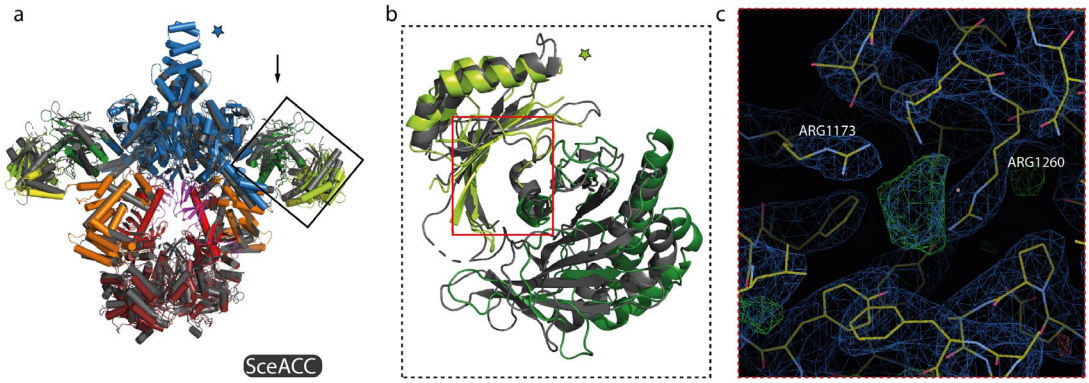

Figure 3.9.3 Extra density in conserved pocket in yeast ACC crystal structure.

a) Overlay of ACC-Cit dimer (color) and SceACC dimer (black) aligned at the CT domain (indicated with blue star) in cartoon representation with cylindrical helices. b) Cartoon representation of CD_{C1} and CD_{C2} domains aligned at CD_{C1} indicated by the green star. c) Image from coot from red highlighted region in b). Electron density map (blue) and model (yellow) of SceACC crystal structure (PDB ID:5CSK). Positive difference density shown in green. Contour level is 0.12 e/A^3.

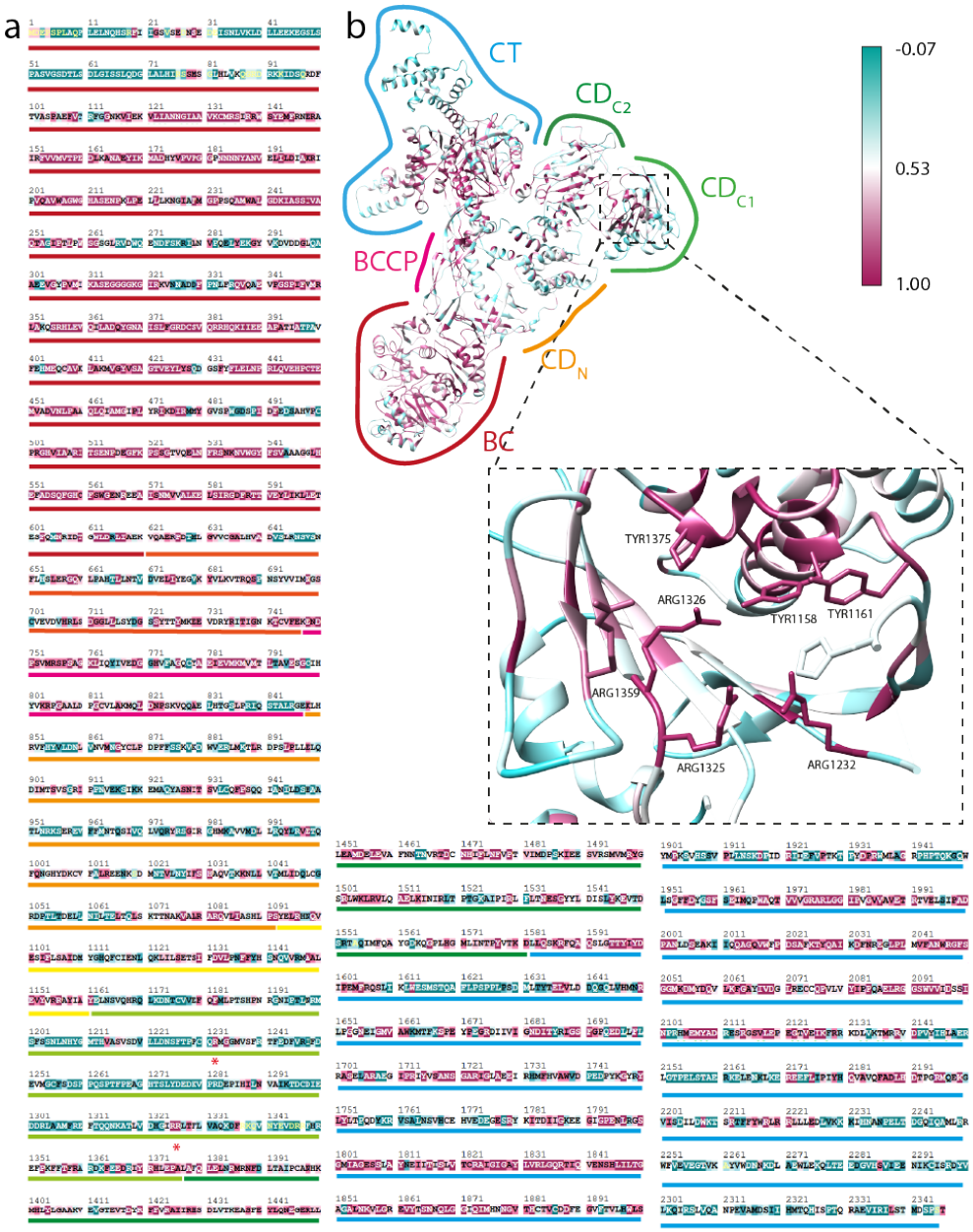

Figure 3.9.4 Sequence alignment shows highly conserved citrate binding pocket.

a) Sequence alignment with domains indicated in color coded lines. Stars indicate mutation sites. b) Conservation mapped onto ACC protomer with domains indicated. Bar shows level of conservation with 1 being highly conserved and inset of citrate binding pocket.

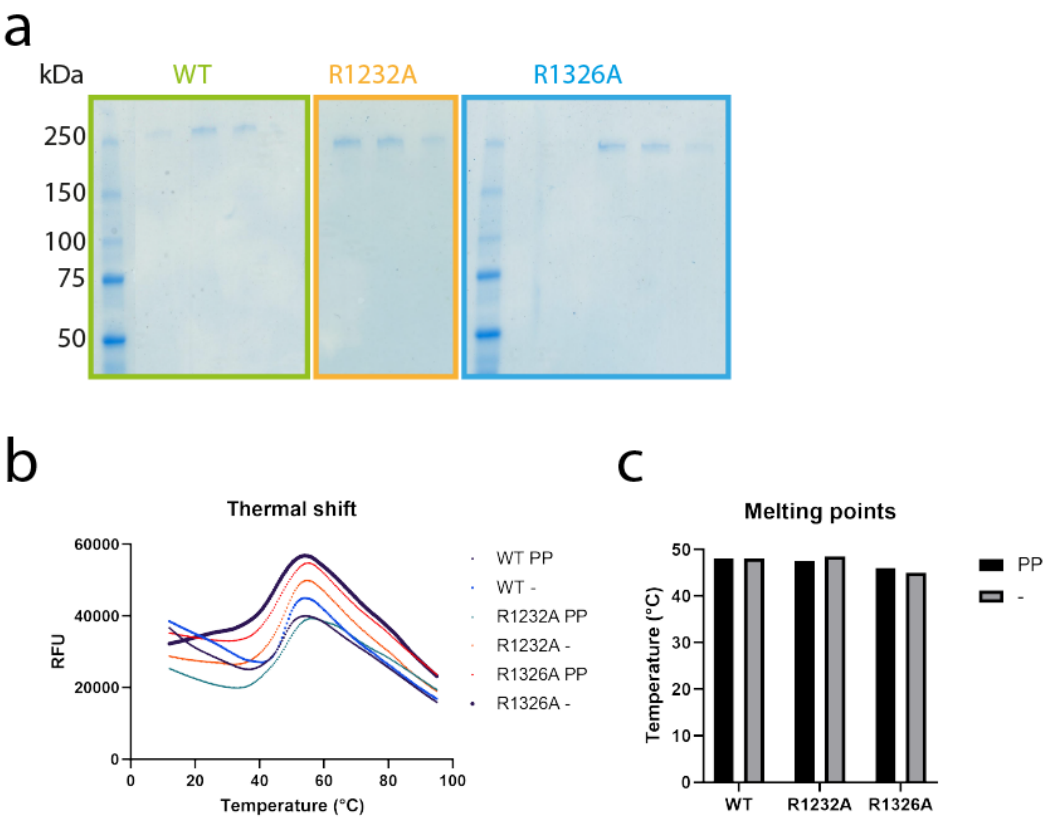

Figure 3.9.5 Expression and folding of ACC WT and point mutants.

a) Coomassie stain of SDS-PAGE of WT and mutants. b) Fluorescence readout of thermal shift assay plotted into temperature axis. RFU = raw fluorescence. c) Melting points determined by the first derivative of thermal shift assay plotted as bar graphs for visualization (PP=phosphorylated).

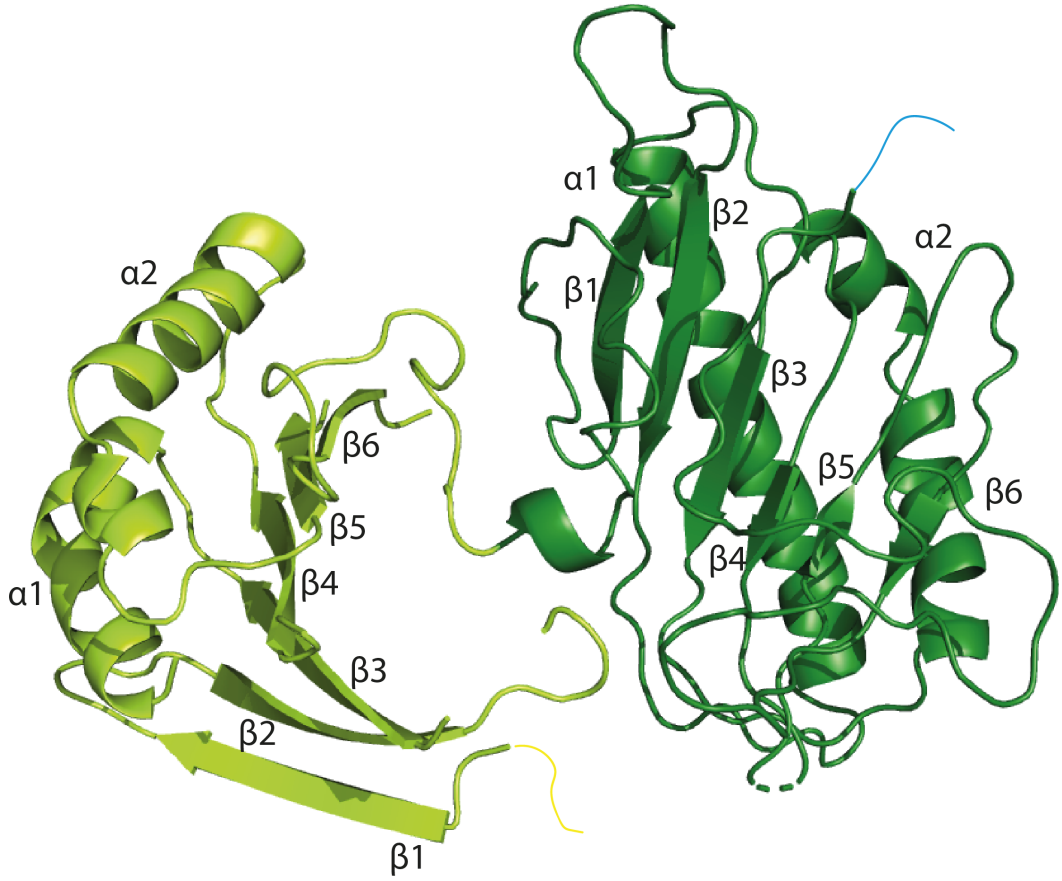

Figure 3.9.6 Human ACC-Cit CDC1 and CDC2 domains with labeled α-helices and β-sheets.

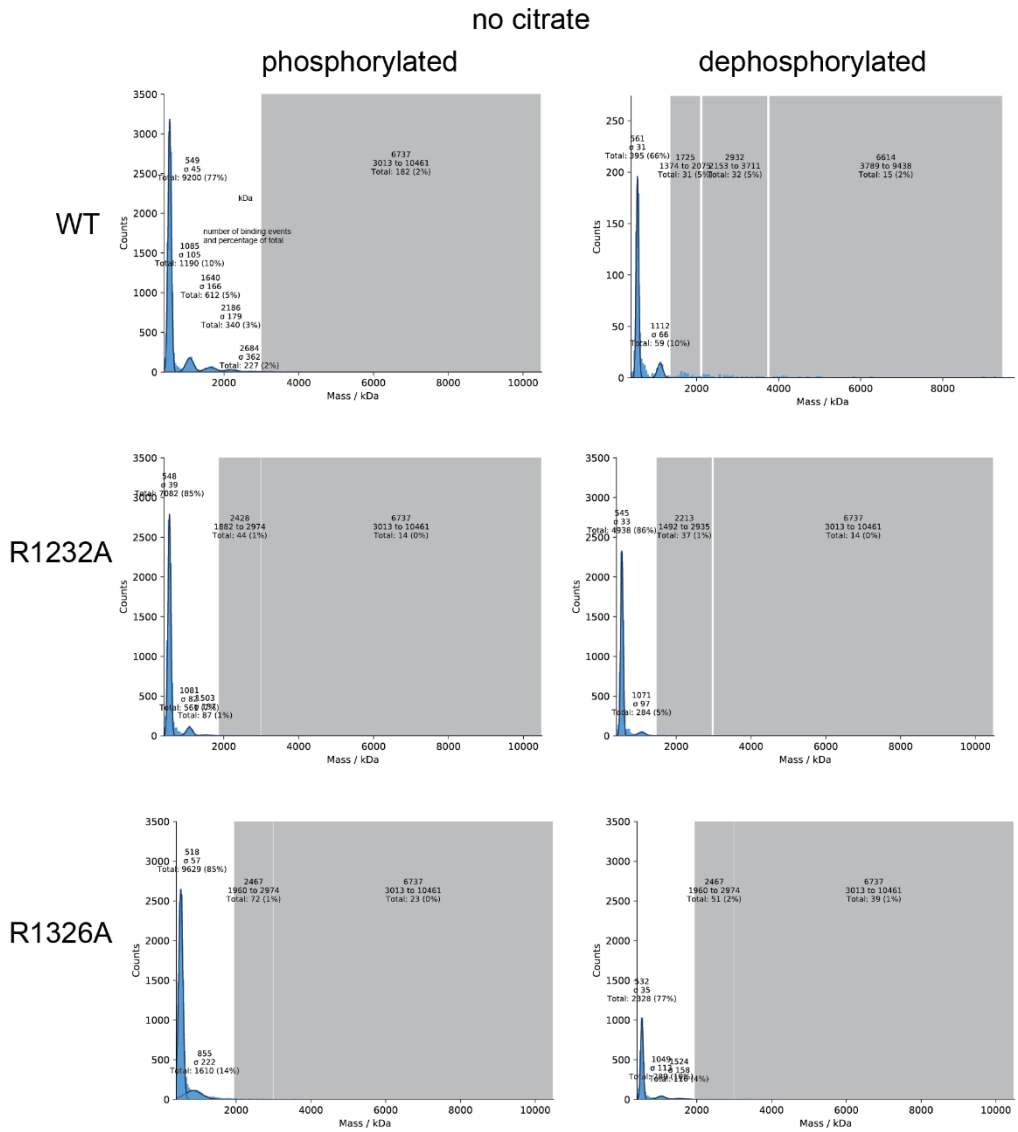

Figure 3.9.7 Mass photometry of samples without citrate show low amount of high molecular weight species.

Number of binding events plotted as a function of molecular weight from start of first peak. Label: top: Mean peak value or integral of region (grey box); middle: Width of the fitted Peak (σ represents the standard deviation); bottom: Sum of the number of counts under the fitted peak/region (expressed as the number of counts and the proportion of counts– compared to the total number of events).

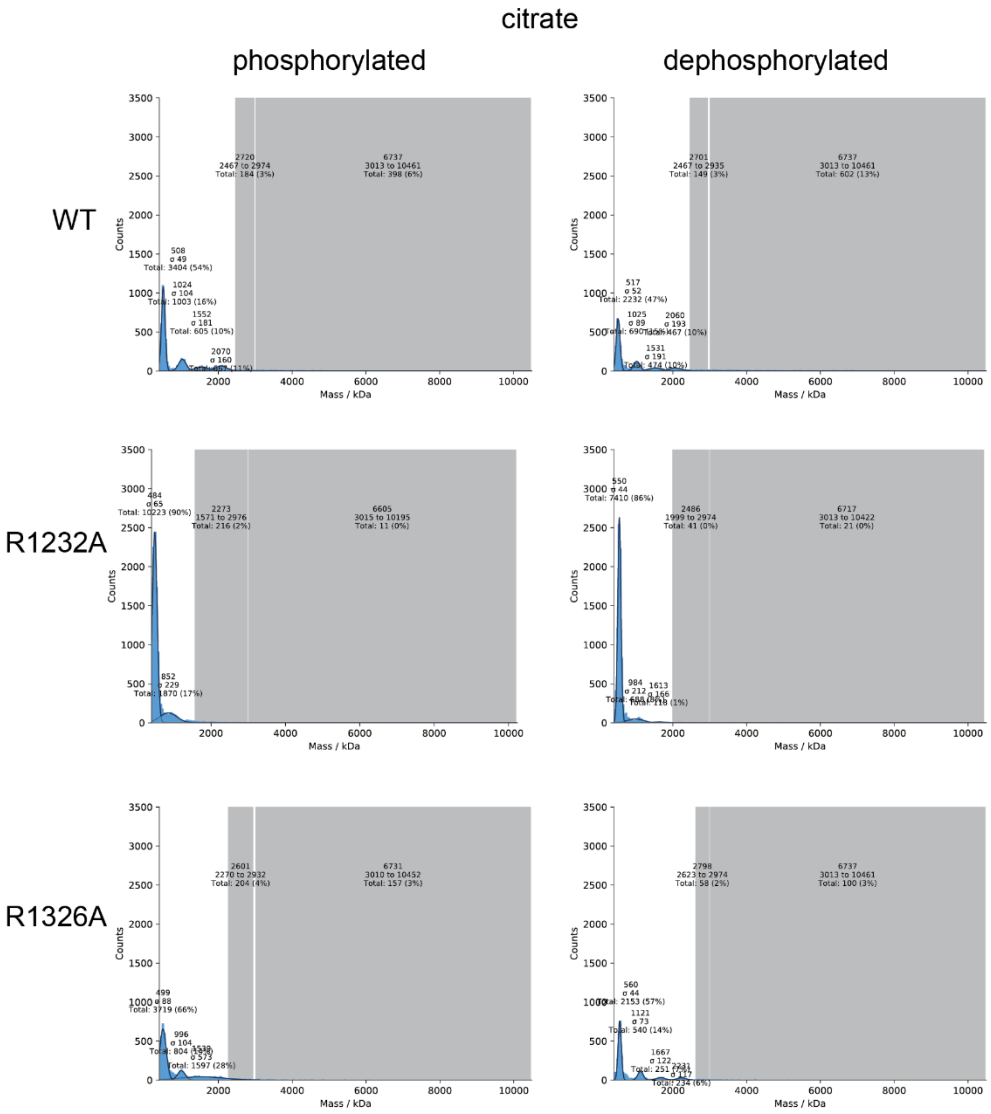

Figure 3.9.8 Mass photometry of WT with citrate show high molecular weight species.

Number of binding events plotted as a function of molecular weight from start of first peak. Label: top: Mean peak value or integral of region (grey box); middle: Width of the fitted Peak (σ represents the standard deviation); bottom: Sum of the number of counts under the fitted peak/region (expressed as the number of counts and the proportion of counts– compared to the total number of events).

3.10 Supplementary method

3.10.1 Assessment of phosphorylation by mass spectrometry

To assess levels of phosphorylation before and after in vitro dephosphorylation we preformed mass spectrometry. Dephosphorylation was preformed independently three times for WT and both mutants using λ-phosphatase (see Methods, Cloning and protein purification).

Proteins were diluted to 50 ug total and were prepared for digestion with 1 ug of trypsin and digested overnight at 37°C. Solid phase extraction was done by using C18 columns. 2.5ug where removed from the sample as Input and was vacuum dried. Resuspension in 20 ul LC-MS buffer (0.1 % formic acid in H_2O) was followed by concentration measurement by nanodrop. 4 ul of sample were injected into Q Exactive HF orpitrap LC-MS/MS (QE-HF) system by Thermo Fisher.

The remaining sample was subjected to phosphopeptide enrichment by AssayMAP Bravo Platform (Agilent). Eluted sample was vacuum dried and peptides dissolved in 20 ul LC-MS buffer and 4ul injected into QE-HF.

Analysis of results was done using Scaffold4 (Scaffold 4.11.0, Proteome software)

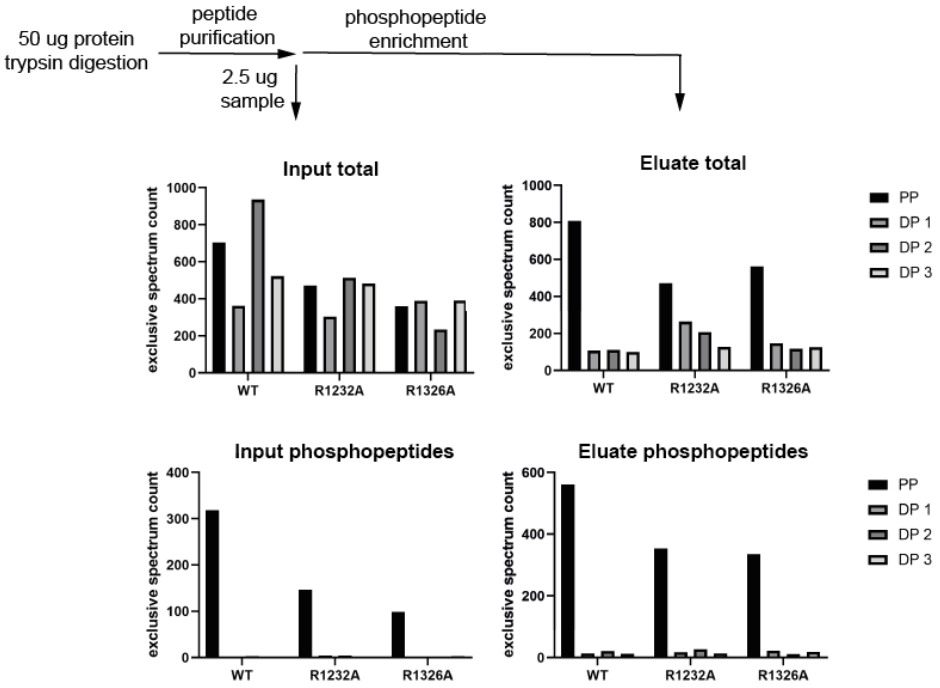

Figure 3.10.1 Mass spectrometry workflow.

Strong reduction in phosphopeptides in dephosphorylated samples.

4 Discussion

In this thesis, we explored the role of oligomerization of urease and acetyl-CoA carboxylase in metabolism. The assembly state of urease, a virulence factor in *Y. enterocolitica,* has been linked to increased survival of pathogens. For more than 50 years, the mechanism behind citrate-induced filament formation of ACC and the coupling to increased activity was an unsolved enigma.

After a general summary of the results in this discussion, the topics of oligomerization interfaces as drug targets, regulation of filament formation as an intervention strategy and the potential of cryo-EM in drug development will be analyzed.

4.1 Summary of Results

The aim 1.1 was to understand how *Y. enterocolitica* urease assembles, by collecting cryo-EM data and processing them with state-of-the-art single particle analysis algorithms. While half of the data was acquired by recording a single movie per grid hole, the other half resulted from collection at three movies from one grid hole, the advantage of which includes shorter acquisition time and reduced image drift. Data processing revealed that the particles displayed preferential orientation and the presence of a small fraction of monomers and broken assemblies. The presence of strong preferential orientation was a limiting factor in resolution, which could not be compensated by applying tetrahedral symmetry. The final resolution of 1.98 Å was achieved using the higher-order aberration corrections for CTF estimation. The reconstruction revealed a tetramer-of-trimers assembly structure for *Y. enterocolitica* urease. Model building was initiated from urease crystal structures. The assembly is similar to *H. pylori* urease which also forms a dodecamer with tetrahedral symmetry.

To understand the determinants for oligomeric assembly of urease, as described in aim 1.2, structure-guided sequence alignments of different ureases were crucial. Through analysis of per residue conservation combined with the atomistic model, relevant differences could be identified. All ureases share a similar overall fold, but detailed investigation of the interfaces revealed variances that may be connected to the various oligomeric assemblies. The areas contributing to oligomerization

show low sequence identity between species, while the active site is highly conserved. A central helix and oligomerization loop identified at the inter-trimer interface have a potential role in assembly of oligomeric urease in *Y. enterocolitica*. The loop binds into a pocket of the neighboring trimer and the helix interacts with another part of the next trimer. Comparison to other ureases show that these structural features are rather unique to *Y. enterocolitica* urease. The helix and loop are absent in some trimeric bacterial urease assemblies and plant ureases seem to have a similar helix but no loop. *H. pylori*, the only other urease known to assemble a dodecamer, has a similar helix and loop. Structural features like these could be determining the oligomeric capacity of a urease.

The broader aim 1.3 of exploring cryo-EM potential in drug discovery was heavily dependent of the first aim. Building of a complete atomistic model allowed for the assessment of previously poor resolved regions at the interfaces between protomers. The interfaces were crucial for the understanding of how oligomerization is mediated. The high level of structural detail allowed for visualization of the active site residues and ions, water networks, salt bridges, alternative side chain conformations and hydration of residues and backbone. For drug discovery and design, these high-resolution features are essential. The coordination of metal ions in the active site is crucial for describing the catalytic mechanism, but also to design targeted intervention. Active site waters are involved in the catalyzed reaction and their exact position is pivotal. Alternative side chain conformation can inform about local flexibility or different states of the protein. The protonation of residues can indicate the hydrogen bonds that are being formed and which atoms interact with each other. This further establishes the capacity of cryo-EM as a tool for structure-based drug discovery for bacterial virulence factors.

A limiting factor in the previous cryo-EM reconstruction of ACC filaments was a low number of particles. To understand the binding mechanism of citrate as outlined in aim 2.1, additional data was acquired and merged with the previous. Through the increase in cryo-EM movies and using automated picking software, the number of particles was increased two-fold, a decisive factor for reconstruction of large dynamic assemblies. Previous cryo-EM maps of ACC were strongly affected by the heterogeneity of the filament, but also by the conditions of sample prep and microscope alignment. ACC filaments can be up to one micron in length and tend to extend from the carbon into the middle of the grid hole. As a consequence, the different regions of the filament are located in varying ice thickness, which influences CTF. By employing the new CTF refinement algorithms, the CTF of each

particle in the filament could be estimated. For the first three data collections, the microscope was not well aligned, which was known at that time. Higher order corrections accounting for aberrations introduced through misalignment lead to further improved CTF estimations. Improved CTF refinements and aberration correction yielded a reconstruction with high enough resolution where visualization of small molecules is possible. The biotin moiety of the BCCP can be seen clearly at the CT. In the CD domain, between the CD_{C1} and CD_{C2}, a region with positive residues was observed to stretch into the center of a pocket. The reconstruction showed an additional density in the center, which is consistent with the presence of a citrate molecule but the current resolution prevents and unambiguous assignment.

To validate the potential binding site as per aim 2.2, the candidate residues involved in the binding of citrate were mutated. Two arginines in the presumptive binding pocket were mutated to alanines. Further experiments with a mutant not affecting citrate binding are ongoing. The proper folding and behavior of mutants was analyzed using thermal shift assays and gel filtration. The mutants were then compared to WT ACC in activity assays with and without citrate. In the activity assay the formation of malonyl-CoA is quantified through incorporation of a C^{14} labelled carboxyl group. The mutants showed either lower activity than WT ACC or complete abrogation of production of malonyl-CoA. To analyze the impact of the mutants on filamentation, we used mass photometry to distinguish different sizes of oligomers assembled by WT and mutant ACC. WT ACC forms larger oligomers than the mutants. One mutant showed almost no higher oligomeric assemblies, while the other had a reduced potential to form oligomers. Both mutants clearly affect activity of ACC, but to estimate the degree to which they abolish citrate binding more data is required.

In this thesis we show the potential of cryo-EM to further the understanding of large oligomeric assemblies. Detailed description of the oligomeric state and active site of *Y. enterocolitica* urease was achieved by using the latest in cryo-EM methods. Additionally, the high resolution enabled us to visualize the effects of radiation damage on negatively charged side chains and nickel ions in the active site. Frame-wise analysis revealed the dynamic movement of the residues during image acquisition.

The improved 3D reconstruction of ACC in combination with biochemical assays revealed a potential citrate binding site in the CD of ACC. However, limitations in resolution do not allow for unequivocal placement of citrate. Further refinements

using different masking to improve resolution of the CD and the citrate binding pocket itself have so far not been successful. Focusing on a part of the complex that is too far away from the center of mass during refinement, which is the case for the CD region, can lead to a distortion of the reconstruction. A possible solution would be to place the fulcrum close to or at the CD domain. Using symmetry to improve resolution of ACC dimers within the filament, has not yet yielded any insightful reconstructions. Using helical reconstruction and applying helical symmetry could help in improving representation of the CD and the filament interface. Classification of 3D reconstruction visualized the different angles adopted by the filaments, but there were not enough particles in the separated states to again obtain a high-quality reconstruction showing the citrate binding pocket clearly. Establishing an assay to directly quantify the binding of citrate to WT and mutant ACC has not been possible thus far and needs to be investigated further. The size of full-length ACC and the filament formation associated with citrate binding poses a challenge for many methods normally used for such binding assays. Truncations of the protein containing only the CD_{C1} and CD_{C2} domains and the mutations, could be extremely useful for standard binding assays like isothermal titration calorimetry.

4.2 Oligomerization interfaces as drug targets

Bacterial urease is a potent virulence factor in pathogens and therefore an attractive drug target. They are very suitable for treatment of infections in humans, because animals don't have ureases. The structure of the bacterial urease active site has been known for some time, but so far, no efficient and low side-effect inhibitory antibiotic drug has been developed. Urease's oligomeric assemblies could present possibilities for a new intervention strategy. While further investigation using mutagenesis studies into the interfaces is needed, a potential mechanism for inhibition can be envisioned. The interaction within the trimers have large interfaces and would probably be more difficult to disrupt. However, the interface between trimers in the tetrahedral assembly are much smaller. The dodecameric assembly of *Y. enterocolitica* urease shows a central helix and oligomerization loop as possible drivers of oligomerization. The loop binds into a specific pocket of the neighboring trimer. The interaction is mediated by few residues and could be targeted with a chemical compound. As we see from comparison with *H. pylori* urease, this loop is not conserved. Therefore, a molecule designed to out-compete binding of the loop for *Y. enterocolitica* urease may not be effective against *H. pylori*

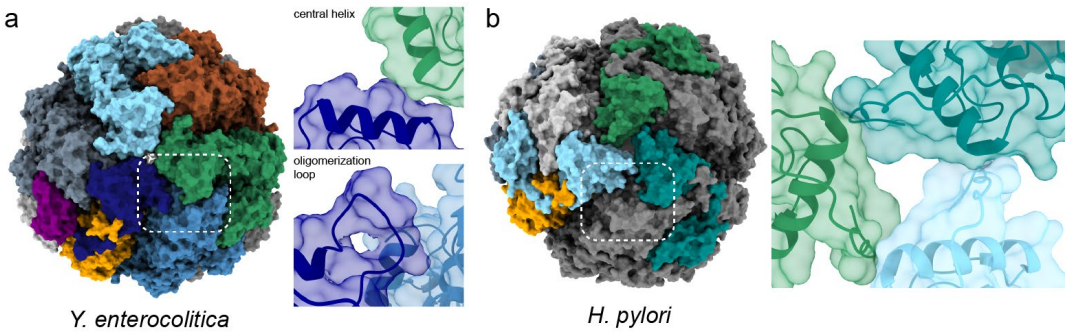

Figure 4.2.1 Interfaces in dodecameric ureases as drug targets.

Surface representation of a) *Y. enterocolitica* urease (PDB ID:6YL3) and b) *H. pylori* (PDB ID: 1E9Z) assemblies with region of interfaces highlighted. Insets show zoomed in interfaces.

urease. In the case of *H. pylori,* the loop directly interacts with the central helix of the neighboring trimer, so the helix could be a more general target. The search for similar motifs in urease sequences could also potentially lead to the identification of additional higher oligomeric ureases. Investigation of other pathogens where urease is associated with virulence, could potentially identify a more universal mechanism for the oligomerization of urease and provide new avenues for therapeutic interventions. The disruption of interfaces could have a lower potential for antibiotic resistance, as it would require parallel mutation of two regions of the protein under strong selective pressure.

Protein-protein interfaces (PPI) have been a proposed target for inhibition for a long time, as approval rates of novel molecules affecting the active sites of enzymes decrease [211]. The challenges associated with development of PPI inhibitors are still large. There needs to be detailed structural information for the complete description of an interface and not all PPI are suitable for interference [212, 213]. Peptides mimicking the interaction partners may provide high target specificity but are usually not stable enough to be used as drugs [214]. PPIs usually have large buried surface areas, but a small molecule would not have to cover the entire interface to perturb the interaction. These interfaces have grooves and patches with high binding energy, so only few residues would need to be covered by an inhibitor [212]. So far there are only few examples of successful interference with PPI by small molecules. The most prominent example of PPI inhibitors are nutlins, which prohibit binding of the oncogene murine double minute 2 (MDM2) to p53. By prohibiting this interaction p53 can continue its tumor suppressing function in DNA repair, cell, cycle regulation, angiogenesis and apoptosis [214].

4.3 Regulation of filament formation as a therapeutic intervention strategy

Human ACC is intimately connected with obesity and cancer, two of the most widespread ailments in society. High activity of the enzymes is associated with disease states. Therefore, the interface of filamentation in ACC presents an alternative target for therapeutic intervention instead of the more commonly explored active sites. Inhibition of filament formation has the potential to prevent the associated increase in activity, although more research is needed. In similar fashion to the peptides mentioned in the previous section, a bulky strongly binding peptide could make filamentation impossible. However, for the design of such a peptide, residue level resolution is necessary, which is not yet available. Directly targeting the citrate binding site also presents a potential site for blocking increase in activity. Designing a compound that prevents the domain rearrangements within the CD, would prevent a closed and active conformation. The phospholoop of fungal ACC binds in the same pocket as citrate in human ACC, but inhibits activity. It is highly likely that the phospholoop achieves a stabilization of the domains, preventing the conformational changes needed for efficient catalysis. Mimicking the inhibition by the phospholoop presents a potential path to an inhibitor.
A further possibility is to modulate the oligomer equilibrium by stabilizing the inactive state over the active one [5]. BRCT domains of BRCA1 completely change ACC conformation and assemble a distinct inactive type of filament. The isomerase Pin1 stabilizes ACC1 and prevents the lysosomal pathway by binding to the CT of ACC, but the structural basis for this stabilization is unclear. Both proteins lock ACC in a conformation, either promoting or inhibiting activity. These examples show an already evolved strategy to regulate ACC, which could be harnessed as a drug target. The mechanism by which Mig12 reduces citrate threshold for ACC activity and how Spot14 then reverses that effect is unclear. Further structural characterization of the protein interactors is required to understand how they can stabilize inactive rather than active ACC.

More than 60 filament forming enzymes have been identified thus far and there are likely many more [215]. New examples of such enzymes will probably increase as the availability and quality of in-cell visualization of proteins advances [216, 217]. For most of these enzymes, high-resolution structural information is not yet available. The structural and functional investigation of these large assemblies will provide crucial insights in to the role of filaments in regulation [218]. While some

enzymes seem to be activated by filament formation, others are inhibited [219]. In-depth knowledge of the regulatory mechanism of these enzymes will be required to identify potential sites for chemical and therapeutic intervention.

4.4 The potential of cryo-EM in drug development

We outline a possible path to high-resolution structure determination for drug target discovery in bacteria. Conditions for this were a stable, largely static sample and the use of state-of-the-art cryo-EM acquisition and processing algorithms. Urease has high symmetry which also aided in the processing of the data. For cryo-EM to be established as a tool for routine structure determination, several steps in the process need more automatization. New developments are being made in terms of grip preparation and realignment of microscopes after switching of samples [220, 221]. Processing of data during image collection has recently become available, as well as algorithms assessing data quality [222, 223]. Like X-ray crystallography now, cryo-EM is on track to become a fully automated technique and become a routine method for structure determination and drug target discovery in the near future.

While writing this thesis two cryo-EM structures of apoferritin at atomic resolution of below 1.3 Å were released. Newly developed electron sources, energy filters and detectors permit to go beyond the previous resolution limit of 1.5 Å for cryo-EM [144, 224]. The reconstructions enable visualization of separated atoms including hydrogen atoms. Because cryo-EM maps represent a coulomb potential, the information on atom bonds, charged states and hydrogens is available at lower resolutions compared to X-ray crystallography [225]. So far this information could only be extracted from electron density maps and is crucial for drug design. Precise location information of individual atoms is essential to understand catalytic reactions and develop small molecules to interfere with them.

Understanding the behavior of protein complexes in their physiological context is crucial to move forward with drug development. As for all structural methods available, the inherent properties of the sample in terms of size, symmetry and flexibility pose challenges for cryo-EM. For large and static protein assemblies, near atomic resolution can be achieved. Resolving heterogeneity in smaller regions of larger and more dynamic proteins is still difficult. Using user-designed masks to focus 3D refinement on single regions has already lead to improved interpretation of these data [155, 156]. However, if the flexibilities occur on a continuum rather

than discrete states, currently established programs cannot resolve this. This problem has been recognized and algorithms tackling this problem are being developed. However, they are still in their early stages [226, 227]. The possibility of analyzing structures moving on a continuum would open up the potential for time-resolved cryo-EM.

Electron microscopy has a long history of imaging large samples like cells and tissues. Cryo-electron tomography can also be used to image large cellular complexes, but high-resolution reconstructions from tomograms are still beyond reach. Single-particle cryo-EM has the ability to computationally separate information from different complexes from the same grid gives rise to a potential alternative approach to structure determination. Instead of the often-laborious process of protein purification needed for X-ray crystallography and NMR, whole cell lysates could be imaged. This would make identification of new and rare complexes, which only occur transiently under certain cellular conditions, possible. Algorithms using this kind of approach are already available, but still depend heavily on prior knowledge of the sample [228]. Targeting oligomerization needs near atomic resolution structure description to understand the interface and design drugs accordingly.

In this thesis, we show that enzyme oligomerization is an essential part of metabolic pathways in all branches of life. As more and more enzymes are revealed to form higher order assemblies, it becomes evident that the evolutionary advantage of such complexes must be significant. Oligomeric enzymes can also be the source of diseases, which furthers the need to investigate these assemblies. To be able to develop efficient treatment, the enzyme, its interactions with itself and other proteins, and their impact on metabolism need to be understood.

Metabolism involves all chemical reactions occurring in an organism and is responsible for extracting energy from food, producing building blocks for the cell to grow and divide, and excreting toxic waste. Any type of cell receives a myriad of chemical cues from their surroundings, indicating the metabolic state of the organism or the environmental conditions. These inputs need to be analyzed and integrated into an appropriate output to maintain cell homeostasis. Metabolic pathways need strict regulation and need to be able to react to changing environments at a moment's notice. By evolving different regulatory check points at the enzyme level, the cell can manage the interplay between its catabolism and anabolism. The way enzymes are usually thought about, as rather static units adrift in the cell, does not reflect reality. Enzymes not only depend on their amino acid sequence and three-dimensional structure, but also on the transient or permanent interactions they form with other proteins.

5 Acknowledgements

First, I would like to thank Timm for the possibility to work on both of these great projects. After my first major setback in my scientific career one and a half years into my PhD, you immediately went into action to make sure I had not only a back up, but a truly fantastic new research question. I also want to thank you for always supporting me in my goals and especially in the last few months when I was applying for post-doc positions. I want to thank Petr for his support already during my masters. I always appreciated your input and advice and am grateful to have you on my PAC. I also want to thank Sebastian for being on my committee and always asking the right questions, making me think about science in a broader light. I will never forget how you told me to go and have a beer after my first PAC meeting and I sure took your advice. I would like to thank Marek for coming on board on such short notice as the co-referee. I always enjoyed our chats at apero's and your advice about science and career.

I want to thank all the core facilities at the Biozentrum. Alex and Ulrike from the proteomics core facility for their great support. Tim and Leico from the biophysics facility for taking time to talk to me about my project and coming up with new ideas of how to solve a problem. I also want to thank Mohamed and Carola from the BioEM lab, who taught me everything about cryo-EM sample preparation and the microscopes. I also want to thank everyone from the Stahlberg group in their great support in all things cryo-EM. I want to thank Ricardo Righetto and Ricardo Adaixo for the great collaboration on the urease project. It was fantastic working with you and I wish you all the best for your future.

I want to thank Francesca for always bringing a light to my otherwise very boring days. You showed me what it means to fight for your dreams and stand up for yourself. To Matze, I just want to say I am so glad that I had you during all of the up and downs of the PhD. You were always there with supporting words, but most of all you were always there. The both of you made my time at the Biozentrum not just a good experience but a truly fantastic one and I could not have done it without you. I wish the both of only the best for your future and I know whatever you pursue, it will be fantastic.

I want to thank Malik for all the support and advice you gave me concerning science, career and New York. I will miss our conversations in the office, the lab, the park or the rhine about politics, society, the universe and life in general. Elsa, you only came

to the lab a year ago, but to me it feels like you have been there for my hole PhD. I will miss having breaks with you and airing our frustrations about our proteins. I know you will do fantastic, I have no doubts about it. Yves, I always enjoyed and appreciated our deep conversations we had at aperos or in Saas Fee. You were always there to help me when I had a question about cryo-EM processing or when I fell on the slope. I am looking forward to what you will accomplish in your future. Shubham, you were the other person in the lab with a dark and weird sense of humor and I cannot express how much I appreciated you being there. You always knew when I needed to hear a sarcastic comment and you regularly made my day better.

To Karo, I have not known you for long, but I always like chatting to you about science, career and your rabbits. I enjoy the energy you bring to the group and I am sure you will do great in all your endeavors. To Florian aka captain column, you were always there when I had a problem in the lab and I truly appreciate you taking so much initiative in the group. I want to thank Yana for being my tea drinking buddy amongst all the coffee addicts and I enjoyed chatting with you about the vegetable basket and sustainability. To Stefan, it was truly fantastic being in the lab with you. You were always there with your advice and humor and the lab became a little bit empty without you there. Roman, you were the first person I worked with in the lab and taught me everything I know about protein purification and crystallography. You were always there with advice and never too busy to help. I want to thank Niko for helping me trying set up radioactive assays and for his advice in cryo-EM processing. I want to thank Hugo for your enthusiasm at our game nights, it was always fantastic. To Louis-Marie, welcome and all the best for your future in the Maier lab. I also want to thank Alexandra for her support with all administrative work, you made everything go smoothly.

I also want to thank all other former lab members that I had the pleasure of working with. Janine and Dominik you were always there with your valuable insights and help. Freddy, you always impressed me with your determination and I appreciate all the times we had together and the help and support you provided for my future career. I want to thank the people of team ACC Eddie, Anna and Moritz. I am so grateful that I could build on such a fantastic scientific foundation. Eddie, you were always ready for one of my many questions about anything regarding science and always took time out of your busy day to help me. Anna and Moritz, even across oceans, you were there with your advice and amazing insights. I could not have done it without you. I am looking forward to only being a four-hour train ride away from you.

I want to thank Beat, Barbara, Adela and all the labs from the 3rd floor for providing a great work environment. I especially want to thank the Schirmer and Perez lab for your support and usually having the exact chemicals, I was missing.
To the lizards/ladies Anna, Francesca, Anja, Johanna and Freddy. I appreciate all you have done for me physically and mentally. Especially in the last few months our 24/7 chat was a true life saver. Even when we are spread around the world in three different time zones, I can always count on you. Whether it be scientific, political or personal, our conversations enrich my life and I would not miss it for the world. I am looking forward to seeing all the fantastic things all of you will accomplish in your life.

An all mini Fründe usserhalb vom Biozentrum: Ihr sind mi Halt und mini Stärke. Ohni euch hätt ich es nit bis do ahne gschafft. Ich bi so dankbar das ich Fründe wie euch ha wo mich immer wiedr dra erinnere, dass es au e Läbe usserhab vo dr Wüsseschaft git.
Ich möchte meiner ganzen Familie danken, die mich immer unterstützt haben. Durch alle Höhen und Tiefen, habt ihr immer an mich geglaubt. Ohne euch, wäre ich nicht die Person, die ich heute bin. Auch ein grosses Dankeschön an Jans Eltern, die mich so herzlich als Teil ihrer Familie aufgenommen haben.

And last but not least, Jan. Du bisch immer für mich do gsi, egal was d umständ gsi sind. Du bisch mi Fels in dr Brandig uff dä ich mich ka stütze. Ich kas kum erwarte dr Rescht vo mim Läbe mit dir zvbringe.

6 References

1. Wang P, R.M., Wang Y, , et al., *p53 domains: structure, oligomerization, and transformation.* Molecular and Cellular Biology, 1994. 14: p. 5182-5191.
2. Eisenstein, E. and D. Beckett, *Dimerization of the Escherichia coli Biotin Repressor: Corepressor Function in Protein Assembly.* Biochemistry, 1999. 38(40): p. 13077-13084.
3. Yun, M.-K., et al., *Crystal Structure and Allosteric Regulation of the Cytoplasmic Escherichia coli l-Asparaginase I.* Journal of Molecular Biology, 2007. 369(3): p. 794-811.
4. Thulin, E., T. Kesvatera, and S. Linse, *Molecular Determinants of S100B Oligomer Formation.* PLOS ONE, 2011. 6(3): p. e14768.
5. Gabizon, R. and A. Friedler, *Allosteric modulation of protein oligomerization: an emerging approach to drug design.* Frontiers in chemistry, 2014. 2: p. 9-9.
6. Unwin, N., et al., *Activation of the Nicotinic Acetylcholine Receptor Involves a Switch in Conformation of the α Subunits.* Journal of Molecular Biology, 2002. 319(5): p. 1165-1176.
7. Goodsell, D.S. and A.J. Olson, *Structural Symmetry and Protein Function.* Annual Review of Biophysics and Biomolecular Structure, 2000. 29(1): p. 105-153.
8. Marianayagam, N.J., M. Sunde, and J.M. Matthews, *The power of two: protein dimerization in biology.* Trends in Biochemical Sciences, 2004. 29(11): p. 618-625.
9. Monod, J., J. Wyman, and J.-P. Changeux, *On the nature of allosteric transitions: A plausible model.* Journal of Molecular Biology, 1965. 12(1): p. 88-118.
10. Ali, M.H. and B. Imperiali, *Protein oligomerization: How and why.* Bioorganic & Medicinal Chemistry, 2005. 13(17): p. 5013-5020.
11. Sumner, J.B., *The isolation and crystallization of the enzyme urease: preliminary paper.* Journal of Biological Chemistry, 1926. 69(2): p. 435-441.
12. Dixon, N.E., et al., *Jack bean urease (EC 3.5.1.5). Metalloenzyme. Simple biological role for nickel.* Journal of the American Chemical Society, 1975. 97(14): p. 4131-4133.
13. Zhang, Y., et al., *Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization.* BMC genomics, 2009. 10: p. 78-78.
14. Bioinformatics, S.I.o. *ExPASy*. [cited 2020 31.05]; Available from: https://www.expasy.org/.
15. Maroney, M.J. and S. Ciurli, *Nonredox Nickel Enzymes.* Chemical Reviews, 2014. 114(8): p. 4206-4228.
16. Maroney, M.J., *Structure/function relationships in nickel metallobiochemistry.* Current Opinion in Chemical Biology, 1999. 3(2): p. 188-199.
17. Ragsdale, S.W., *Nickel-based Enzyme Systems.* The Journal of biological chemistry, 2009. 284(28): p. 18571-18575.
18. Withers, P.C., *Urea: Diverse Functions of a 'Waste" Product.* Clinical and Experimental Pharmacology and Physiology, 1998. 25(9): p. 722-727.
19. Blakeley, R.L., et al., *Jack bean urease (EC 3.5.1.5). Demonstration of a carbamoyl-transfer reaction and inhibition by hydroxamic acids.* Biochemistry, 1969. 8(5): p. 1991-2000.

20. Callahan, B.P., Y. Yuan, and R. Wolfenden, *The Burden Borne by Urease.* Journal of the American Chemical Society, 2005. 127(31): p. 10828-10829.
21. Zambelli, B., et al., *Chemistry of Ni2+ in Urease: Sensing, Trafficking, and Catalysis.* Accounts of Chemical Research, 2011. 44(7): p. 520-530.
22. Bremner, J.M. and R.L. Mulvaney, *Urease activity in soils.* New York: Academic Press, 1978: p. 149–96.
23. Modolo, L.V., et al., *A minireview on what we have learned about urease inhibitors of agricultural interest since mid-2000s.* Journal of Advanced Research, 2018. 13: p. 29-37.
24. Mobley, H.L. and R.P. Hausinger, *Microbial ureases: significance, regulation, and molecular characterization.* Microbiological Reviews, 1989. 53(1): p. 85.
25. Kosikowska, P. and Ł. Berlicki, *Urease inhibitors as potential drugs for gastric and urinary tract infections: a patent review.* Expert Opinion on Therapeutic Patents, 2011. 21(6): p. 945-957.
26. Cox, G.M., et al., *Urease as a virulence factor in experimental cryptococcosis.* Infection and immunity, 2000. 68(2): p. 443-448.
27. Mirbod-Donovan, F., Schaller, R., Hung, C.-Y., Xue, J., Reichard, U., Cole, G.T., *Urease produced by Coccidioides posadasii contributes to the virulence of this respiratory pathogen.* Infection and Immunity, 2006. 74: p. 504-515.
28. Jones, B.D., et al., *Construction of a urease-negative mutant of Proteus mirabilis: analysis of virulence in a mouse model of ascending urinary tract infection.* Infection and immunity, 1990. 58(4): p. 1120-1123.
29. Mobley, H.L.T., L.T. Hu, and P.A. Foxall, *Helicobacter pylori Urease: Properties and Role in Pathogenesis.* Scandinavian Journal of Gastroenterology, 1991. 26(sup187): p. 39-46.
30. Uberti, A.F., et al., *Pro-inflammatory properties and neutrophil activation by Helicobacter pylori urease.* Toxicon, 2013. 69: p. 240-249.
31. Kappaun, K., et al., *Ureases: Historical aspects, catalytic, and non-catalytic properties - A review.* Journal of advanced research, 2018. 13: p. 3-17.
32. Maroncle, N., C. Rich, and C. Forestier, *The role of Klebsiella pneumoniae urease in intestinal colonization and resistance to gastrointestinal stress.* Research in Microbiology, 2006. 157(2): p. 184-193.
33. Sangari, F.J., et al., *Characterization of the urease operon of Brucella abortus and assessment of its role in virulence of the bacterium.* Infection and immunity, 2007. 75(2): p. 774-780.
34. Murphy, T.F. and A.L. Brauer, *Expression of urease by Haemophilus influenzaeduring human respiratory tract infection and role in survival in an acid environment.* BMC Microbiology, 2011. 11(1): p. 183.
35. Young, G.M., D. Amid, and V.L. Miller, *A bifunctional urease enhances survival of pathogenic Yersinia enterocolitica and Morganella morganii at low pH.* Journal of bacteriology, 1996. 178(22): p. 6487-6495.
36. Drummond, N., et al., *Yersinia Enterocolitica: A Brief Review of the Issues Relating to the Zoonotic Pathogen, Public Health Challenges, and the Pork Production Chain.* Foodborne Pathogens and Disease, 2012. 9(3): p. 179-189.
37. Giannella, R.A., S.A. Broitman, and N. Zamcheck, *Gastric acid barrier to ingested microorganisms in man: studies in vivo and in vitro.* Gut, 1972. 13(4): p. 251-256.
38. De Koning-Ward, T.F. and R.M. Robins-Browne, *Contribution of urease to acid tolerance in Yersinia enterocolitica.* Infection and immunity, 1995. 63(10): p. 3790-3795.

39. Ha, N.-C., et al., *Supramolecular assembly and acid resistance of Helicobacter pylori urease.* Nature Structural Biology, 2001. 8(6): p. 505-509.
40. Hausinger, R., *Urease.* Biochemistry of nickel, Plenum, New York, 1993: p. 23–57.
41. Ligabue-Braun, R., et al., *3-to-1: unraveling structural transitions in ureases.* Naturwissenschaften, 2013. 100(5): p. 459-467.
42. Jabri, E., et al., *The crystal structure of urease from Klebsiella aerogenes.* Science, 1995. 268(5213): p. 998-1004.
43. Balasubramanian, A. and K. Ponnuraj, *Crystal Structure of the First Plant Urease from Jack Bean: 83 Years of Journey from Its First Crystal to Molecular Structure.* Journal of Molecular Biology, 2010. 400(3): p. 274-283.
44. Turbett, G.R., et al., *Purification and characterization of the urease enzymes of Helicobacter species from humans and animals.* Infection and immunity, 1992. 60(12): p. 5259-5266.
45. Benini, S., et al., *A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from Bacillus pasteurii: why urea hydrolysis costs two nickels.* Structure, 1999. 7(2): p. 205-216.
46. Karplus, P.A., M.A. Pearson, and R.P. Hausinger, *70 Years of Crystalline Urease: What Have We Learned?* Accounts of Chemical Research, 1997. 30(8): p. 330-337.
47. Estiu, G. and K.M. Merz, *Competitive Hydrolytic and Elimination Mechanisms in the Urease Catalyzed Decomposition of Urea.* The Journal of Physical Chemistry B, 2007. 111(34): p. 10263-10274.
48. Soriano, A. and R.P. Hausinger, *GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins.* Proceedings of the National Academy of Sciences, 1999. 96(20): p. 11140.
49. Fong, Y.H., et al., *Structure of UreG/UreF/UreH complex reveals how urease accessory proteins facilitate maturation of Helicobacter pylori urease.* PLoS biology, 2013. 11(10): p. e1001678-e1001678.
50. Musiani, F., et al., *Protein Tunnels: The Case of Urease Accessory Proteins.* Journal of Chemical Theory and Computation, 2017. 13(5): p. 2322-2331.
51. Kleinschmidt, A.K., J. Moss, and M.D. Lane, *Acetyl Coenzyme A Carboxylase: Filamentous Nature of the Animal Enzymes.* Science, 1969. 166(3910): p. 1276-1278.
52. Hawkins, A.R. and H.K. Lamb, *The Molecular Biology of Multidomain Proteins Selected Examples.* European Journal of Biochemistry, 1995. 232(1): p. 7-18.
53. Ekman, D., et al., *Multi-domain Proteins in the Three Kingdoms of Life: Orphan Domains and Other Unassigned Regions.* Journal of Molecular Biology, 2005. 348(1): p. 231-243.
54. Hagmann, A., *Regulation of Eukaryotic Acetyl-CoA Carboxylases*, in *Biozentrum*. 2018, University of Basel.
55. Aparicio, J.F., et al., *Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus: Analysis of the enzymatic domains in the modular polyketide synthase.* Gene, 1996. 169(1): p. 9-16.
56. Srere, P.A., *Complexes of Sequential Metabolic Enzymes.* Annual Review of Biochemistry, 1987. 56(1): p. 89-124.
57. Perham, R.N., *Swinging Arms and Swinging Domains in Multifunctional Enzymes: Catalytic Machines for Multistep Reactions.* Annual Review of Biochemistry, 2000. 69(1): p. 961-1004.
58. Maier, T., et al., *Structure and function of eukaryotic fatty acid synthases.* Quarterly Reviews of Biophysics, 2010. 43(3): p. 373-422.

59. de Kok, A., et al., *The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1998. 1385(2): p. 353-366.
60. Tong, L., *Chapter Five - Striking Diversity in Holoenzyme Architecture and Extensive Conformational Variability in Biotin-Dependent Carboxylases*, in *Advances in Protein Chemistry and Structural Biology*, T. Karabencheva-Christova, Editor. 2017, Academic Press. p. 161-194.
61. Tong, L., *Structure and function of biotin-dependent carboxylases.* Cellular and Molecular Life Sciences, 2013. 70(5): p. 863-891.
62. Jitrapakdee, S., et al., *Structure, mechanism and regulation of pyruvate carboxylase.* The Biochemical journal, 2008. 413(3): p. 369-387.
63. Wakil, S.J., J.K. Stoops, and V.C. Joshi, *Fatty Acid Synthesis and its Regulation.* Annual Review of Biochemistry, 1983. 52(1): p. 537-579.
64. Cronan, J.E. and G.L. Waldrop, *Multi-subunit acetyl-CoA carboxylases.* Progress in Lipid Research, 2002. 41(5): p. 407-435.
65. Tong, L., *Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery.* Cellular and Molecular Life Sciences CMLS, 2005. 62(16): p. 1784-1803.
66. Zhang, H., B.A. Boghigian, and B.A. Pfeifer, *Investigating the role of native propionyl-CoA and methylmalonyl-CoA metabolism on heterologous polyketide production in Escherichia coli.* Biotechnology and Bioengineering, 2010. 105(3): p. 567-573.
67. Knowles, J.R., *The Mechanism of Biotin-Dependent Enzymes.* Annual Review of Biochemistry, 1989. 58(1): p. 195-221.
68. Attwood, P.V. and J.C. Wallace, *Chemical and Catalytic Mechanisms of Carboxyl Transfer Reactions in Biotin-Dependent Enzymes.* Accounts of Chemical Research, 2002. 35(2): p. 113-120.
69. Hunkeler, M., *Dynamic architecture of multi-domain carboxylase*, in *Biozentrum*. 2016, University of Basel.
70. Maurice, M.S., et al., *Domain Architecture of Pyruvate Carboxylase, a Biotin-Dependent Multifunctional Enzyme.* Science, 2007. 317(5841): p. 1076-1079.
71. Schär, J., et al., *Pyruvate Carboxylase Plays a Crucial Role in Carbon Metabolism of Extra- and Intracellularly Replicating <em>Listeria monocytogenes</em&gt.* Journal of Bacteriology, 2010. 192(7): p. 1774.
72. Kanamori, T., et al., *Enzymatic characterization of a prokaryotic urea carboxylase.* Journal of bacteriology, 2004. 186(9): p. 2532-2539.
73. Fan, C., et al., *Crystal Structure of Urea Carboxylase Provides Insights into the Carboxyltransfer Reaction.* Journal of Biological Chemistry, 2012. 287(12): p. 9389-9398.
74. Huang, C.S., et al., *Crystal structure of the alpha(6)beta(6) holoenzyme of propionyl-coenzyme A carboxylase.* Nature, 2010. 466(7309): p. 1001-1005.
75. Huang, C.S., et al., *An unanticipated architecture of the 750-kDa α6β6 holoenzyme of 3-methylcrotonyl-CoA carboxylase.* Nature, 2011. 481(7380): p. 219-223.
76. Aguilar, J.A., et al., *Substrate specificity of the 3-methylcrotonyl coenzyme A (CoA) and geranyl-CoA carboxylases from Pseudomonas aeruginosa.* Journal of bacteriology, 2008. 190(14): p. 4888-4893.
77. Munday, M.R., *Regulation of mammalian acetyl-CoA carboxylase.* Biochemical Society Transactions, 2002. 30: p. 1059-1064.
78. Abu-Elheiga, L., et al., *The subcellular localization of acetyl-CoA carboxylase 2.* Proceedings of the National Academy of Sciences, 2000. 97(4): p. 1444.

79. McGarry, J.D. and N.F. Brown, *The Mitochondrial Carnitine Palmitoyltransferase System — From Concept to Molecular Analysis.* European Journal of Biochemistry, 1997. 244(1): p. 1-14.
80. Eberlé, D., et al., *SREBP transcription factors: master regulators of lipid homeostasis.* Biochimie, 2004. 86(11): p. 839-848.
81. Field, F.J., et al., *Polyunsaturated fatty acids decrease the expression of sterol regulatory element-binding protein-1 in CaCo-2 cells: effect on fatty acid synthesis and triacylglycerol transport.* Biochemical Journal, 2002. 368(3): p. 855-864.
82. Foufelle, F. and P. FerrÉ, *New perspectives in the regulation of hepatic glycolytic and lipogenic genes by insulin and glucose: a role for the transcription factor sterol regulatory element binding protein-1c.* Biochemical Journal, 2002. 366(2): p. 377-391.
83. Guri, Y., et al., *mTORC2 Promotes Tumorigenesis via Lipid Synthesis.* Cancer Cell, 2017. 32(6): p. 807-823.e12.
84. Brown, N.F., et al., *The mammalian target of rapamycin regulates lipid metabolism in primary cultures of rat hepatocytes.* Metabolism - Clinical and Experimental, 2007. 56(11): p. 1500-1507.
85. Locke, G.A., et al., *Differential activation of recombinant human acetyl-CoA carboxylases 1 and 2 by citrate.* Archives of Biochemistry and Biophysics, 2008. 475(1): p. 72-79.
86. Taketa, T. and B.M. Pogell, *The Effect of Palmitoyl Coenzym A on Glucose 6-Phosphate Dehydrogenase and Otehr Enzymes.* Biological Chemistry, 1965. 241(3): p. 720-726.
87. Goodridge, A.G., *Regulation of the Activity of Acetyl Coenzyme A Carboxylase by Palmitoyl Coenzyme A and Citrate.* Biological Chemistry, 1972. 247(21): p. 6946-6952.
88. Ha, J., et al., *Critical phosphorylation sites for acetyl-CoA carboxylase activity.* Journal of Biological Chemistry, 1994. 269(35): p. 22162-22168.
89. Davies, S.P., A.T.R. Sim, and D.G. Hardie, *Location and Function of 3 Sites Phosphorylated on Rat Acetyl-CoA Carboxylase by the AMP-activated Protein-Kinase.* European Journal of Biochemistry, 1990. 187(1): p. 183-190.
90. Cho, Y.S., et al., *Molecular mechanism for the regulation of human ACC2 through phosphorylation by AMPK.* Biochemical and Biophysical Research Communications, 2010. 391(1): p. 187-192.
91. Woods, A., et al., *Yeast SNF1 is functionally related to mammalian AMP-activated protein kinase and regulates acetyl-CoA carboxylase in vivo.* Journal of Biological Chemistry, 1994. 269(30): p. 19509-19515.
92. Ficarro, S.B., et al., *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae.* Nature Biotechnology, 2002. 20(3): p. 301-305.
93. Shi, S., et al., *Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1.* mBio, 2014. 5(3): p. e01130.
94. Kim, C.W., et al., *Induced polymerization of mammalian acetyl-CoA carboxylase by MIG12 provides a tertiary level of regulation of fatty acid synthesis.* Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(21): p. 9626-9631.
95. Colbert, C.L., et al., *Crystal structure of Spot 14, a modulator of fatty acid synthesis.* Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(44): p. 18820-18825.
96. Park, S., et al., *Spot14/Mig12 heterocomplex sequesters polymerization and restrains catalytic function of human acetyl-CoA carboxylase 2.* Journal of Molecular Recognition, 2013. 26(12): p. 679-688.

97. German, N.J., et al., *PHD3 Loss in Cancer Enables Metabolic Reliance on Fatty Acid Oxidation via Deactivation of ACC2.* Molecular cell, 2016. 63(6): p. 1006-1020.
98. Ueda, K., et al., *Prolyl isomerase Pin1 binds to and stabilizes acetyl CoA carboxylase 1 protein, thereby supporting cancer cell proliferation.* Oncotarget, 2019. 10(17): p. 1637-1648.
99. Magnard, C., et al., *BRCA1 interacts with acetyl-CoA carboxylase through its tandem of BRCT domains.* Oncogene, 2002. 21: p. 6729.
100. Shen, Y. and L. Tong, *Structural evidence for direct interactions between the BRCT domains of human BRCA1 and a phospho-peptide from human ACC1.* Biochemistry, 2008. 47(21): p. 5767-5773.
101. Mao, J., et al., *Liver-specific deletion of acetyl-CoA carboxylase 1 reduces hepatic triglyceride accumulation without affecting glucose homeostasis.* Proceedings of the National Academy of Sciences, 2006. 103(22): p. 8552.
102. Mao, J., et al., *aP2-Cre-mediated inactivation of acetyl-CoA carboxylase 1 causes growth retardation and reduced lipid accumulation in adipose tissues.* Proceedings of the National Academy of Sciences, 2009. 106(41): p. 17576.
103. Abu-Elheiga, L., et al., *Mutant mice lacking acetyl-CoA carboxylase 1 are embryonically lethal.* Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(34): p. 12011.
104. Abu-Elheiga, L., et al., *Acetyl-CoA Carboxylase 2−/− Mutant Mice are Protected against Fatty Liver under High-fat, High-carbohydrate Dietary and de Novo Lipogenic Conditions.* Journal of Biological Chemistry, 2012. 287(15): p. 12578-12588.
105. Hoehn, K.L., et al., *Phenotypic Discrepancies in Acetyl-CoA Carboxylase 2-deficient Mice.* Journal of Biological Chemistry, 2012. 287(19): p. 15801.
106. Olson, D.P., et al., *Gene knockout of Acc2 has little effect on body weight, fat mass, or food intake.* Proceedings of the National Academy of Sciences, 2010. 107(16): p. 7598.
107. Abu-Elheiga, L., et al., *Continuous Fatty Acid Oxidation and Reduced Fat Storage in Mice Lacking Acetyl-CoA Carboxylase 2.* Science, 2001. 291(5513): p. 2613.
108. Flier, J.S., *Obesity Wars: Molecular Progress Confronts an Expanding Epidemic.* Cell, 2004. 116(2): p. 337-350.
109. Friedman, J.M., *A War on Obesity, Not the Obese.* Science, 2003. 299(5608): p. 856.
110. Harriman, G., et al., *Acetyl-CoA carboxylase inhibition by ND-630 reduces hepatic steatosis, improves insulin sensitivity, and modulates dyslipidemia in rats.* Proceedings of the National Academy of Sciences of the United States of America, 2016. 113(13): p. E1796-E1805.
111. Takagi, H., et al., *A Novel Acetyl-CoA Carboxylase 2 Selective Inhibitor Improves Whole-Body Insulin Resistance and Hyperglycemia in Diabetic Mice through Target-Dependent Pathways.* Journal of Pharmacology and Experimental Therapeutics, 2020. 372(3): p. 256.
112. Gerth, K., et al., *The Soraphens: A Family of Novel Antifungal Compounds from Sorangium cellulosum (Myxobacteria).* The Journal of Antibiotics, 1994. 47(1): p. 23-31.
113. Chen, L., et al., *Acetyl-CoA carboxylase (ACC) as a therapeutic target for metabolic syndrome and recent developments in ACC1/2 inhibitors.* Expert Opinion on Investigational Drugs, 2019. 28(10): p. 917-930.
114. Svensson, R.U., et al., *Inhibition of acetyl-CoA carboxylase suppresses fatty acid synthesis and tumor growth of non-small-cell lung cancer in preclinical models.* Nature medicine, 2016. 22(10): p. 1108-1119.

115. Abramson, H.N., *The Lipogenesis Pathway as a Cancer Target.* Journal of Medicinal Chemistry, 2011. 54(16): p. 5615-5638.
116. Chajès, V., et al., *Acetyl-CoA Carboxylase α Is Essential to Breast Cancer Cell Survival.* Cancer Research, 2006. 66(10): p. 5287-5294.
117. Liu, H., et al., *Biochemistry, molecular biology, and pharmacology of fatty acid synthase, an emerging therapeutic target and diagnosis/prognosis marker.* International journal of biochemistry and molecular biology, 2010. 1(1): p. 69-89.
118. Corbet, C., et al., *Acidosis Drives the Reprogramming of Fatty Acid Metabolism in Cancer Cells through Changes in Mitochondrial and Histone Acetylation.* Cell Metabolism, 2016. 24(2): p. 311-323.
119. Mizojiri, R., et al., *The identification and pharmacological evaluation of potent, selective and orally available ACC1 inhibitor.* Bioorganic & Medicinal Chemistry Letters, 2019. 29(23): p. 126749.
120. Esler, W.P., et al., *Human sebum requires de novo lipogenesis, which is increased in acne vulgaris and suppressed by acetyl-CoA carboxylase inhibition.* Science Translational Medicine, 2019. 11(492): p. eaau8465.
121. Singh, S., et al., *AMP-Activated Protein Kinase Restricts Zika Virus Replication in Endothelial Cells by Potentiating Innate Antiviral Responses and Inhibiting Glycolysis.* The Journal of Immunology, 2020. 204(7): p. 1810.
122. Long, S., et al., *Fatty Acids Regulate Porcine Reproductive and Respiratory Syndrome Virus Infection via the AMPK-ACC1 Signaling Pathway.* Viruses, 2019. 11(12): p. 1145.
123. Jiménez de Oya, N., et al., *Targeting host metabolism by inhibition of acetyl-Coenzyme A carboxylase reduces flavivirus infection in mouse models.* Emerging microbes & infections, 2019. 8(1): p. 624-636.
124. Kluck, G.E.G., et al., *Plasmodium Infection Induces Dyslipidemia and a Hepatic Lipogenic State in the Host through the Inhibition of the AMPK-ACC Pathway.* Scientific reports, 2019. 9(1): p. 14695-14695.
125. Currais, A., et al., *Elevating acetyl-CoA levels reduces aspects of brain aging.* eLife, 2019. 8: p. e47866.
126. Kemal, C. and J.E. Casida, *Coenzyme a esters of 2-aryloxyphenoxypropionate herbicides and 2-arylpropionate antiinflammatory drugs are potent and stereoselective inhibitors of rat liver acetyl-CoA carboxylase.* Life Sciences, 1992. 50(7): p. 533-540.
127. Walker, K.A., et al., *Fluazifop, a grass-selective herbicide which inhibits acetyl-CoA carboxylase in sensitive plant species.* The Biochemical journal, 1988. 254(1): p. 307-310.
128. Sawamoto, A., et al., *Heptamethoxyflavone inhibits adipogenesis via enhancing PKA signaling.* European Journal of Pharmacology, 2019. 865: p. 172758.
129. Shen, J., et al., *Choline and methionine regulate lipid metabolism via the AMPK signaling pathway in hepatocytes exposed to high concentrations of nonesterified fatty acids.* Journal of Cellular Biochemistry, 2019. n/a(n/a).
130. Peng, X., et al., *A novel AMPK activator improves hepatic lipid metabolism and leukocyte trafficking in experimental hepatic steatosis.* Journal of Pharmacological Sciences, 2019. 140(2): p. 153-161.
131. Waldrop, G.L., I. Rayment, and H.M. Holden, *Three-Dimensional Structure of the Biotin Carboxylase Subunit of Acetyl-CoA Carboxylase.* Biochemistry, 1994. 33(34): p. 10249-10256.
132. Chou, C.-Y., L.P.C. Yu, and L. Tong, *Crystal structure of biotin carboxylase in complex with substrates and implications for its catalytic mechanism.* The Journal of biological chemistry, 2009. 284(17): p. 11690-11697.

133. Shen, Y., et al., *A Mechanism for the Potent Inhibition of Eukaryotic Acetyl-Coenzyme A Carboxylase by Soraphen A, a Macrocyclic Polyketide Natural Product.* Molecular Cell, 2004. 16(6): p. 881-891.
134. Janiyani, K., et al., *Function of Escherichia coli Biotin Carboxylase Requires Catalytic Activity of Both Subunits of the Homodimer.* Journal of Biological Chemistry, 2001. 276(32): p. 29864-29870.
135. Wei, J. and L. Tong, *Crystal structure of the 500-kDa yeast acetyl-CoA carboxylase holoenzyme dimer.* Nature, 2015. 526(7575): p. 723-727.
136. Athappilly, F.K. and W.A. Hendrickson, *Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing.* Structure, 1995. 3(12): p. 1407-1419.
137. Reddy, D.V., et al., *High Resolution Solution Structure of the 1.3S Subunit of Transcarboxylase from Propionibacterium shermanii†.* Biochemistry, 2000. 39(10): p. 2509-2516.
138. Lee, C.-K., et al., *Biotinoyl domain of human acetyl-CoA carboxylase: Structural insights into the carboxyl transfer mechanism.* Proteins: Structure, Function, and Bioinformatics, 2008. 72(2): p. 613-624.
139. Hunkeler, M., et al., *The dynamic organization of fungal acetyl-CoA carboxylase.* Nature Communications, 2016. 7: p. 11196.
140. Zhang, H., et al., *Crystal Structure of the Carboxyltransferase Domain of Acetyl-Coenzyme A Carboxylase.* Science, 2003. 299(5615): p. 2064-2067.
141. Hunkeler, M., et al., *Structural basis for regulation of human acetyl-CoA carboxylase.* Nature, 2018. 558(7710): p. 470-474.
142. Kühlbrandt, W., *The Resolution Revolution.* Science, 2014. 343(6178): p. 1443-1444.
143. Lawson, C.L., et al., *EMDataBank.org: unified data resource for CryoEM.* Nucleic Acids Research, 2010. 39(suppl_1): p. D456-D464.
144. Nakane, T., et al., *Single-particle cryo-EM at atomic resolution.* bioRxiv, 2020: p. 2020.05.22.110189.
145. Faruqi, A.R. and R. Henderson, *Electronic detectors for electron microscopy.* Current Opinion in Structural Biology, 2007. 17(5): p. 549-555.
146. Brilot, A.F., et al., *Beam-induced motion of vitrified specimen on holey carbon film.* Journal of Structural Biology, 2012. 177(3): p. 630-637.
147. Campbell, M.G., et al., *Movies of ice-embedded particles enhance resolution in electron cryo-microscopy.* Structure (London, England : 1993), 2012. 20(11): p. 1823-1828.
148. Scheres, S.H.W., *Beam-induced motion correction for sub-megadalton cryo-EM particles.* eLife, 2014. 3: p. e03665.
149. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy.* Nature Methods, 2017. 14(4): p. 331-332.
150. Zivanov, J., et al., *New tools for automated high-resolution cryo-EM structure determination in RELION-3.* eLife, 2018. 7: p. e42166.
151. Cheng, Y., et al., *A Primer to Single-Particle Cryo-Electron Microscopy.* Cell, 2015. 161(3): p. 438-449.
152. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination.* Nature Methods, 2017. 14(3): p. 290-296.
153. Zivanov, J., T. Nakane, and S. Scheres, *Estimation of high-order aberrations and anisotropic magnification from cryo-EM datasets in RELION-3.1.* IUCrJ, 2020. 7(2): p. 253-267.
154. Punjani, A., M.A. Brubaker, and D.J. Fleet, *Building Proteins in a Day: Efficient 3D Molecular Structure Estimation with Electron Cryomicroscopy.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017. 39(4): p. 706-718.

155. Punjani, A., H. Zhang, and D.J. Fleet, *Non-uniform refinement: Adaptive regularization improves single particle cryo-EM reconstruction.* bioRxiv, 2019: p. 2019.12.15.877092.
156. Nakane, T., et al., *Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION.* eLife, 2018. 7: p. e36861.
157. Frank, J. and A. Ourmazd, *Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM.* Methods, 2016. 100: p. 61-67.
158. Mazzei, L., et al., *Urease Inhibition in the Presence of N -(n -Butyl)thiophosphoric Triamide, a Suicide Substrate: Structure and Kinetics.* Biochemistry, 2017. 56(40): p. 5391-5404.
159. Mazzei, L., F. Musiani, and S. Ciurli, *CHAPTER 5 Urease*, in *The Biological Chemistry of Nickel*. 2017, The Royal Society of Chemistry. p. 60-97.
160. Mulinari, F., et al., *Characterization of JBURE-IIb isoform of Canavalia ensiformis (L.) DC urease.* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2011. 1814(12): p. 1758-1768.
161. Mazzei, L., et al., *The Impact of pH on Catalytically Critical Protein Conformational Changes: The Case of the Urease, a Nickel Enzyme.* Chemistry – A European Journal, 2019. 25(52): p. 12145-12158.
162. Mazzei, L., et al., *The Structure of the Elusive Urease–Urea Complex Unveils the Mechanism of a Paradigmatic Nickel-Dependent Enzyme.* Angewandte Chemie International Edition, 2019. 58(22): p. 7415-7419.
163. Podschun, R. and U. Ullmann, *Klebsiella spp. as Nosocomial Pathogens: Epidemiology, Taxonomy, Typing Methods, and Pathogenicity Factors.* Clinical Microbiology Reviews, 1998. 11(4): p. 589-603.
164. Bhagat, N. and J.S. Virdi, *Molecular and biochemical characterization of urease and survival of Yersinia enterocolitica biovar 1A in acidic pH in vitro.* BMC microbiology, 2009. 9: p. 262-262.
165. Cheng, A., et al., *High resolution single particle cryo-electron microscopy using beam-image shift.* Journal of Structural Biology, 2018. 204(2): p. 270-275.
166. Schorb, M., et al., *Software tools for automated transmission electron microscopy.* Nature Methods, 2019.
167. Zivanov, J., T. Nakane, and S.H.W. Scheres, *A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis.* IUCrJ, 2019. 6(1).
168. Naydenova, K. and C.J. Russo, *Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy.* Nature Communications, 2017. 8(1): p. 629.
169. Afonine, V.P., et al., *New tools for the analysis and validation of cryo-EM maps and atomic models.* Acta Crystallographica Section D Structural Biology, 2018. 74(9): p. 814-840.
170. Berezin, C., et al., *ConSeq: the identification of functionally and structurally important residues in protein sequences.* Bioinformatics, 2004. 20(8): p. 1322-1324.
171. Ashkenazy, H., et al., *ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules.* Nucleic Acids Research, 2016. 44(W1): p. W344-W350.
172. Benini, S., et al., *Fluoride inhibition of Sporosarcina pasteurii urease: structure and thermodynamics.* JBIC Journal of Biological Inorganic Chemistry, 2014. 19(8): p. 1243-1261.
173. Benini, S., et al., *The complex of Bacillus pasteurii urease with β-mercaptoethanol from X-ray data at 1.65-Å resolution.* JBIC Journal of Biological Inorganic Chemistry, 1998. 3(3): p. 268-273.

174. Yano, J., et al., *X-ray damage to the Mn4Ca complex in single crystals of photosystem II: A case study for metalloprotein crystallography.* Proceedings of the National Academy of Sciences, 2005. 102(34): p. 12047-12052.
175. Grant, T. and N. Grigorieff, *Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6.* eLife, 2015. 4.
176. Hattne, J., et al., *Analysis of Global and Site-Specific Radiation Damage in Cryo-EM.* Structure, 2018. 26(5): p. 759-766.e4.
177. Fromm, S.A., et al., *Seeing tobacco mosaic virus through direct electron detectors.* Journal of Structural Biology, 2015. 189(2): p. 87-97.
178. Suárez, D., N. Díaz, and K.M. Merz, Jr., *Ureases: quantum chemical calculations on cluster models.* J Am Chem Soc, 2003. 125(50): p. 15324-37.
179. Rokita, E., et al., *Purification of surface-associated urease from Helicobacter pylori.* Journal of Chromatography B: Biomedical Sciences and Applications, 2000. 737(1-2): p. 203-212.
180. Weatherburn, M.W., *Phenol-Hypochlorite Reaction for Determination of Ammonia.* Analytical Chemistry, 1967. 39(8): p. 971-974.
181. Biyani, N., et al., *Focus: The interface between data collection and data processing in cryo-EM.* Journal of Structural Biology, 2017. 198(2): p. 124-133.
182. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from electron micrographs.* Journal of Structural Biology, 2015. 192(2): p. 216-221.
183. Arnold, S.A., et al., *Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts.* Journal of Structural Biology, 2016. 197(3): p. 0-6.
184. Zhang, K.; Available from: https://www2.mrc-lmb.cam.ac.uk/research/locally-developed-software/zhang-software/.
185. Rosenthal, P.B. and R. Henderson, *Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy.* Journal of Molecular Biology, 2003. 333(4): p. 721-745.
186. Harauz, G. and M. van Heel, *Exact filters for general geometry three dimensional reconstruction.* Optik, 1986. 78(4): p. 146-156.
187. Scheres, S.H.W.W. and S. Chen, *Prevention of overfitting in cryo-EM structure determination.* Nature Methods, 2012. 9(9): p. 853-854.
188. Chen, S., et al., *High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy.* Ultramicroscopy, 2013. 135(0): p. 24-35.
189. Cardone, G., J.B. Heymann, and A.C. Steven, *One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions.* Journal of Structural Biology, 2013. 184(2): p. 226-236.
190. Pettersen, E.F., et al., *UCSF Chimera—A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. 25(13): p. 1605-1612.
191. Adams, P.D., et al., *PHENIX: building new software for automated crystallographic structure determination.* Acta Crystallographica Section D, 2002. 58(11): p. 1948-1954.
192. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallographica Section D, 2004. 60(12): p. 2126-2132.
193. Jurrus, E., et al., *Improvements to the APBS biomolecular solvation software suite.* Protein Science, 2018. 27(1): p. 112-128.
194. Madeira, F., et al., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic acids research, 2019. 47(W1): p. W636-W641.

195. Kearse, M., et al., *Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. 28(12): p. 1647-1649.
196. Krissinel, E. and K. Henrick, *Inference of Macromolecular Assemblies from Crystalline State.* Journal of Molecular Biology, 2007. 372(3): p. 774-797.
197. Afonine, V.P., et al., *Real-space refinement in PHENIX for cryo-EM and crystallography.* Acta Crystallographica Section D, 2018. 74(6): p. 531-544.
198. Cock, P.J.A., et al., *Biopython: freely available Python tools for computational molecular biology and bioinformatics.* Bioinformatics, 2009. 25(11): p. 1422-1423.
199. Goddard, T.D., et al., *UCSF ChimeraX: Meeting modern challenges in visualization and analysis.* Protein science : a publication of the Protein Society, 2018. 27(1): p. 14-25.
200. McCoy, A.J., et al., *Phaser crystallographic software.* Journal of Applied Crystallography, 2007. 40(4): p. 658-674.
201. Pearson, M.A., et al., *Structures of Cys319 variants and acetohydroxamate-inhibited Klebsiella aerogenes urease.* Biochemistry, 1997. 36(26): p. 8164-8172.
202. Blanc, E., et al., *Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT.* Acta Crystallographica Section D: Biological Crystallography, 2004. 60(12 I): p. 2210-2221.
203. Thampy, K.G., *Formation of malonyl coenzyme A in rat heart. Identification and purification of an isozyme of A carboxylase from rat heart.* Journal of Biological Chemistry, 1989. 264(30): p. 17631-4.
204. Corbet, C. and O. Feron, *Emerging roles of lipid metabolism in cancer progression.* Current Opinion in Clinical Nutrition & Metabolic Care, 2017. 20(4): p. 254-260.
205. Moule, S.K., et al., *Coenzyme A is a potent inhibitor of acetyl-CoA carboxylase from rat epididymal fat-pads.* Biochemical Journal, 1992. 283(1): p. 35-38.
206. Wei, J., et al., *A unified molecular mechanism for the regulation of acetyl-CoA carboxylase by phosphorylation.* Cell Discovery, 2016. 2(1): p. 16044.
207. Inc, S.B. *cryoSPARC*. 2016 [cited 2020 15.05]; Available from: https://cryosparc.com/.
208. Young, G., et al., *Quantitative mass imaging of single biological macromolecules.* Science, 2018. 360(6387): p. 423.
209. *Refeyn*. 2018; Available from: https://www.refeyn.com/.
210. Schrodinger, L. *The PyMOL Molecular Graphics System, Version 2.0*. Available from: https://pymol.org/2/#page-top.
211. Valkov E., et al., *Targeting Protein–Protein Interactions and Fragment-Based Drug Discovery*, in *Fragment-Based Drug Discovery and X-Ray Crystallography*, Davies T. and H. M., Editors. 2011, Springer, Berlin, Heidelberg.
212. Zinzalla, G. and D.E. Thurston, *Targeting protein-protein interactions for therapeutic intervention: a challenge for the future.* Future Med Chem, 2009. 1(1): p. 65-93.
213. Chène, P., *Drugs Targeting Protein–Protein Interactions.* ChemMedChem, 2006. 1(4): p. 400-411.
214. Chung, C. and M.M. Hann, *Targeting Protein–Protein Interactions Perspective*, in *Structural Biology in Drug Discovery*. 2020. p. 479-502.
215. Park, C.K. and N.C. Horton, *Novel insights into filament-forming enzymes.* Nature Reviews Molecular Cell Biology, 2020. 21(1): p. 1-2.
216. Shashkova, S. and M.C. Leake, *Single-molecule fluorescence microscopy review: shedding new light on old problems.* Bioscience reports, 2017. 37(4): p. BSR20170031.

217. Shen, Q.J., et al., *Filamentation of Metabolic Enzymes in Saccharomyces cerevisiae.* Journal of Genetics and Genomics, 2016. 43(6): p. 393-404.
218. Park, C.K. and N.C. Horton, *Structures, functions, and mechanisms of filament forming enzymes: a renaissance of enzyme filamentation.* Biophysical Reviews, 2019. 11(6): p. 927-994.
219. Lynch, E.M., et al., *Human CTP synthase filament structure reveals the active enzyme conformation.* Nature Structural & Molecular Biology, 2017. 24(6): p. 507-514.
220. Dandey, V.P., et al., *Spotiton: New features and applications.* Journal of Structural Biology, 2018. 202(2): p. 161-169.
221. Tan, Y.Z., et al., *Automated data collection in single particle electron microscopy.* Microscopy, 2015. 65(1): p. 43-56.
222. Inc, S.B. *cryoSPARC Live*. 2020 [cited 2020 24.05]; Available from: https://cryosparc.com/live.
223. Tegunov, D. and P. Cramer, *Real-time cryo-electron microscopy data preprocessing with Warp.* Nature Methods, 2019. 16(11): p. 1146-1152.
224. Yip, K.M., et al., *Breaking the next Cryo-EM resolution barrier - Atomic resolution determination of proteins!* bioRxiv, 2020: p. 2020.05.21.106740.
225. Marques, M.A., M.D. Purdy, and M. Yeager, *CryoEM maps are full of potential.* Current Opinion in Structural Biology, 2019. 58: p. 214-223.
226. Punjani, A. and D.J. Fleet, *3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images.* bioRxiv, 2020: p. 2020.04.08.032466.
227. Zhong, E.D., et al., *CryoDRGN: Reconstruction of heterogeneous structures from cryo-electron micrographs using neural networks.* bioRxiv, 2020: p. 2020.03.27.003871.
228. Ho, C.-M., et al., *Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu.* Nature Methods, 2020. 17(1): p. 79-85.